Diversification of Afrobatrachian Frogs and the Herpetofauna of the Arabian Peninsula

By

Daniel Portik

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Integrative Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Dr. Jimmy A. McGuire, Chair
Dr. Rauri Bowie
Dr. David Blackburn
Dr. Rosemary Gillespie

Fall 2015

Abstract

Diversification of Afrobatrachian Frogs and the Herpetofauna of the Arabian Peninsula

by

Daniel Portik

Doctor of Philosophy in Biology

University of California, Berkeley

Dr. Jimmy A. McGuire, Chair

The identification of biotic and abiotic factors that promote the diversification of clades across Africa and the Arabian Peninsula remains a difficult challenge. A variety of ecological and evolutionary processes can be driving such patterns, and clade-specific traits may also play a role in the evolution of these groups. Comparative evolutionary studies of particular clades, relying on a phylogenetic framework, can be used to investigate many of these topics. Beyond these mechanisms there are abiotic factors, such as geological events, that can drive vicariance and dispersal events for large sets of taxa. The investigation of historical biogeography in a comparative phylogenetic framework can be used to detect such patterns. My dissertation explores these topics using reptiles and amphibians as study systems, and I rely on the generation of molecular sequence data, phylogenetics, and the use of comparative phylogenetic methods to address a variety of questions. I provide the abstract for each chapter below.

Chapter 1: The reproductive modes of anurans (frogs and toads) are the most diverse among all the terrestrial vertebrates, and a major challenge is identifying selective factors that promote the evolution or retention of reproductive modes across clades. Terrestrialized anuran breeding strategies, in which breeding is partially or completely independent of water, have evolved repeatedly from the plesiomorphic fully aquatic reproductive mode, a process which is thought to occur through a series of intermediate reproductive stages. A number of selective forces have been proposed for the evolution of terrestrialized reproductive traits, but factors such as water systems and co-evolution with particular ecomorphologies have not been investigated in a comparative phylogenetic framework. We examined these topics and the evolution of reproductive mode in Afrobatrachian frogs, an ecologically and reproductively diverse clade representing more than half of the total frog diversity found in Africa (~400 species). We find that direct development evolved twice independently from terrestrialized reproductive modes involving protected eggs or larvae, supporting the intermediate stages hypothesis. We detected correlated evolution in specific derived ecomorphologies and reproductive traits, including arboreality and arboreal oviposition, and fossoriality and subterranean oviposition. There is support for a link between lotic water systems and terrestrial

oviposition, and between arboreal ecomorphology and the use of lentic water systems for reproduction. Our findings indicate clade-specific processes driving the evolution of reproductive modes in the Afrobatrachia, and suggest that other processes such as habitat filtering and co-evolution with ecomorphology may generate global patterns of anuran reproductive modes.

Chapter 2: Custom sequence capture experiments are becoming an efficient approach for gathering large sets of orthologous markers with targeted levels of informativeness in non-model organisms. Transcriptome-based exon capture utilizes transcript sequences to design capture probes, often with the aid of a reference genome to identify intron-exon boundaries and exclude shorter exons (< 200 bp). Here, we test an alternative approach that directly uses transcript sequences for probe design, which are often composed of multiple exons of varying lengths. Based on a selection of 1,260 orthologous transcripts, we conducted sequence captures across multiple phylogenetic scales for frogs, including species up to ~100 million years divergent from the focal group. After several conservative filtering steps, we recovered a large phylogenomic data set consisting of sequence alignments for 1,047 of the 1,260 transcriptome-based loci (~630,000 bp) and a large quantity of highly variable regions flanking the exons in transcripts (~70,000 bp). We recovered high numbers of both shorter (< 100 bp) and longer exons (> 200 bp), with no major reduction in coverage towards the ends of exons. We observed significant differences in the performance of blocking oligos for target enrichment and non-target depletion during captures, and observed differences in PCR duplication rates that can be attributed to the number of individuals pooled for capture reactions. We explicitly tested the effects of phylogenetic distance on capture sensitivity, specificity, and missing data, and provide a baseline estimate of expectations for these metrics based on nuclear pairwise differences among samples. We provide recommendations for transcriptome-based exon capture design based on our results, and describe multiple pipelines for data assembly and analysis.

Chapter 3: The anuran family Hyperoliidae is Africa's most diverse group of frogs, with a distribution across sub-Saharan Africa (16 genera, over 200 species), Madagascar (1 genus, 11 species), and the Seychelles Islands (1 monotypic genus). Hyperoliids are found in a variety of habitats and altitudes, but exhibit the highest species richness and co-occurring species in forested regions, including lowland rainforests and montane forests. As many as 14 hyperoliid species have been recorded in near sympatry, and species show high turnover based on both elevational and habitat gradients. Hyperoliids exhibit extensive morphological diversity, ranging from semi-terrestrial forms, to walker-climbers, to largely arboreal forms, and display considerable variation in body size across species and between the sexes. A striking and unique feature of many hyperoliid species is their bright coloration and patterning, which can be highly variable and driven by the presence of color polymorphism or sexual dichromatism. Using phylogenomic data, I reconstructed the evolutionary relationships of the family and investigated the evolution of sexual size dimorphism and sexual dichromatism. My examination of the body sizes of hyperoliid species revealed this family largely exhibits female-biased SSD, and I find evidence for an isometric scaling relationship between body sizes of the sexes. Within the genus *Hyperolius* I found significant differences in the evolutionary trend of male and

female body sizes, namely selection for a single optimal female body size and strong directional selection for males occurring in different subclades. I propose the functional relationship between female body size and arboreal oviposition may be opposing the selection for increased fecundity associated with larger female sizes. The body sizes of males may be the result of a combination of interspecific resource partitioning, division of acoustic space, and reproductive character displacement. Sexual dichromatism has evolved independently in two main clades of hyperoliids: 1) the common ancestor of the Malagasy and Seychelles Islands species (*Heterixalus*, *Tachycnemis*), and 2) the common ancestor of *Hyperolius, Morerella*, and *Cryptothylax*. A number of subsequent reversals to monochromatism occur in the genus *Hyperolius* and *Heterixalus*, with some cases of secondarily derived dichromatism.

Chapter 4: The Red Sea has had a profound biogeographic effect on organisms with Afro-Asian distributions, resulting in complex patterns of admixture on the Arabian Peninsula. We investigate the phylogenetic affinities of a monitor lizard (*Varanus yemenensis*) restricted to the southwestern Arabian Peninsula by sequencing all African monitor species and several Asian monitor species for the mitochondrial gene *ND2* and the nuclear marker *RAG-1*. We find evidence that *V. yemenensis* is of African origin, being most closely related to the white-throat monitor, *V. albigularis*, an African species complex distributed from the Horn of Africa to southern Africa. Using divergence-dating analyses, we investigate several biogeographic hypotheses to infer the likely mechanism of colonization of the Arabian Peninsula by this species. Our results reveal that both dispersal across a southern landbridge and overwater dispersal are potential explanations. The patterns observed in *V. yemenensis* are contrasted with other taxa having similar Afro-Arabian disjunct distributions to better understand the complex biogeographic history of this region.

Chapter 5: The Arabian Peninsula is home to a unique fauna that has assembled and evolved throughout the course of major geophysical events, including the separation of the Arabian Plate from Africa and subsequent collision with Eurasia. Opportunities for faunal exchanges with particular continents occurred in temporally distinct periods, and the presence of African, Western Eurasian, and South Asian derived taxa on the Arabian Peninsula signifies the complexity of these historical biogeographic events. The six true toad species (family Bufonidae) endemic to Arabian Peninsula present a considerable taxonomic and biogeographic challenge because they are part of a global bufonid radiation, including several genera surrounding the Arabian Peninsula, and difficult to discriminate morphologically. As they could be derived from African, Western Eurasian, or South Asian toad groups, elucidating their evolutionary relationships has important implications for historical biogeography. Here, we analyze a global molecular data set of 243 bufonid lineages, with an emphasis on new sampling from the Horn of Africa, Western Eurasia, South Asia, and the Arabian Peninsula, to reconstruct the evolutionary relationships of the Arabian species. We produce a robust time-calibrated phylogeny to infer the biogeographic history of this group on and around the Arabian Peninsula. Our phylogenetic analyses indicate two of the endemic Arabian toad species, "*Bufo*" *tihamicus* and "*Bufo*" *arabicus*, evolved independently within the African genus *Amietophrynus*. We confirm the Arabian species *Duttaphrynus dhufarensis* is of South

Asian origin, but do not find evidence for the Asian genus *Duttaphrynus* being present in the Horn of Africa, discrediting a previously proposed Asian bufonid dispersal event to Africa. We also do not find evidence of the African genus *Amietophrynus* occurring in South Asia, suggesting that unlike many other vertebrate taxa, toads have not used the Arabian Peninsula as a stepping-stone for trans-continental dispersal. Our divergence dating estimates strongly suggest the formation of the Red Sea drove simultaneous divergences between two of the Arabian species (*A. tihamicus* comb. nov. and *A. arabicus* comb. nov.) and their closest mainland African relatives in the Early Miocene. We estimate the divergence of *D. dhufarensis* with its closest South Asian relatives occurred in the mid to Late Miocene, suggesting the temporary or permanent land connections between the Arabian plate and Eurasia facilitated dispersal of this lineage to the Arabian Peninsula. The Arabian bufonid assemblage, despite being comparatively depauperate with respect to surrounding continents, exemplifies the faunal pattern of the Arabian Peninsula, namely being a complex admixture of African, Western Eurasian, and South Asian elements. The historical biogeographic patterns exhibited by Arabian toads and their allies are concordant with studies of other vertebrate taxa, building support for the role of major geological events in driving simultaneous vicariance and dispersal events around the Arabian Peninsula. Although many taxa or groups exhibiting disjunct Afro-Arabian distributions appear to have dispersed more recently from the Horn of Africa via a southern land bridge or overwater dispersal, both *Amietophrynus tihamicus* and *A. arabicus* likely represent true African relicts resulting from vicariance associated with the Red Sea formation, a pattern that so far is rare among the vertebrate species investigated.

Chapter 6: We summarize the various geologic and climatic-based hypotheses influencing the diversification of Arabian taxa and highlight the role of an understudied biodiversity hotspot, the Arabian Hotspot Area (AHA), in driving patterns of endemism. We evaluate these biogeographic hypotheses and patterns utilizing Arabian amphibians as a study system, and contrast these results with a comprehensive literature review of other Arabian taxa. We investigated biogeographic hypotheses using a temporally explicit phylogenetic framework and by compiling divergence-dating estimates for all available extant Arabian taxa. To examine the effects of climate change, we generated distribution models for Arabian amphibians under current and historical climate regimes to identify historically stable refugial areas for montane and lowland species. We find a striking concordance of divergence times for taxa with similar biogeographic histories, providing strong support for hypotheses derived from geologic events. Most in-situ speciation events occurring on the Arabian Peninsula are concentrated in the late Miocene, with intraspecific divergences of species occurring mainly in the Plio-Pleistocene. Changes in connectivity between the Arabian Peninsula and surrounding continents throughout geologic time allows explicit temporal predictions to be made about biogeographic events for Arabian taxa derived from African, Asian, and Palearctic lineages. We find the AHA is in many ways similar to the montane systems of East Africa, and is characterized by habitat stability and the long-term persistence of localized lineages and the diversification of cyclically fragmented lineages. Historically stable refugial regions for amphibian species mainly occur within the AHA, demonstrating the importance of this region for the persistence of both montane and lowland species.

Table of Contents

List of Figures

List of Tables

Chapter 1

The evolution of reproductive diversity in Afrobatrachia: A phylogenetic comparative analysis of an extensive radiation of African frogs

Abstract

The reproductive modes of anurans (frogs and toads) are the most diverse among all the terrestrial vertebrates, and a major challenge is identifying selective factors that promote the evolution or retention of reproductive modes across clades. Terrestrialized anuran breeding strategies, in which breeding is partially or completely independent of water, have evolved repeatedly from the plesiomorphic fully aquatic reproductive mode, a process which is thought to occur through a series of intermediate reproductive stages. A number of selective forces have been proposed for the evolution of terrestrialized reproductive traits, but factors such as water systems and co-evolution with particular ecomorphologies have not been investigated in a comparative phylogenetic framework. We examined these topics and the evolution of reproductive mode in Afrobatrachian frogs, an ecologically and reproductively diverse clade representing more than half of the total frog diversity found in Africa (~400 species). We find that direct development evolved twice independently from terrestrialized reproductive modes involving protected eggs or larvae, supporting the intermediate stages hypothesis. We detected correlated evolution in specific derived ecomorphologies and reproductive traits, including arboreality and arboreal oviposition, and fossoriality and subterranean oviposition. There is support for a link between lotic water systems and terrestrial oviposition, and between arboreal ecomorphology and the use of lentic water systems for reproduction. Our findings indicate clade-specific processes driving the evolution of reproductive modes in the Afrobatrachia, and suggest that other processes such as habitat filtering and co-evolution with ecomorphology may generate global patterns of anuran reproductive modes.

Introduction

Across terrestrial vertebrates, amphibians have the greatest diversity of reproductive modes, which are often accompanied by a wide range of ecologies (Duellman & Trueb 1986). In particular, the order Anura (frogs and toads) exhibits incredible variation with at least 40 distinct reproductive modes, and further fine-scale divisions resulting in nearly 60 possible modes (Boulenger 1886; Salthe & Duellman 1973; Duellman 1985; Duellman & Trueb 1986; Haddad & Prado 2005; Altig & McDiarmid 2007; Wells 2007; Iskandar et al. 2014; Crump 2015). These modes are defined on the basis of oviposition site, egg and clutch characteristics, larval environment and development, and the degree of parental care. The proposed plesiomorphic reproductive condition of anurans is characterized by aquatic oviposition with a free-swimming exotrophic larval stage (Duellman 1985), which is a widespread mode among extant anurans. This mode is recovered as the ancestral state of anurans using phylogenetic comparative methods, and also appears to be the basal condition for a majority of frog families (Gomez-Mestre et al. 2012). Much of the variation in reproductive mode seems to represent steps towards an independence from water, which many anurans rely on for oviposition, larval development, or both (Lutz 1947, 1948; Salthe & Duellman 1973; Duellman & Trueb

1986; Wells 2007). The conventional view posits that many of these modes represent a long sequence of changes leading to greater terrestriality, with the terminal stage being direct development in which endotrophic larvae develop in terrestrial eggs and hatch as miniature versions of the adults. However, the evolution of direct development and other highly terrestrialized modes through intermediate stages has recently been challenged, and the switch to direct development may also occur from aquatic breeders, thereby circumventing intermediate stages (Gomez-Mestre et al. 2012). The 'intermediate stage' hypothesis has received further exploration with improved sampling across specific clades, although recent studies have found support both for (Meegaskumbura et al. 2015) and against (Pereira et al. 2015) this pattern. An additional surprising result of Gomez-Mestre et al. (2012) is that reproductive modes in which eggs and larvae are placed in locations protected from predators do evolve repeatedly, but do not seem to give rise to direct developers. The prevalence of direct developing species across several anuran families (~25%, Duellman 2007) indicates that the condition is convergent, and has likely evolved from a number of reproductive modes and selective forces (Gomez-Mestre et al. 2012), however these patterns are only recently being explored in a phylogenetic framework across major clades.

There are a number of proposed selective pressures that can promote terrestrialized breeding strategies, including avoidance of aquatic predation (Lutz 1948; Tihen 1960; Crump 1974; Prado et al. 2002; Haddad & Prado 2005) or parasites (Todd 2007), precipitation and humidity levels (da Silva et al. 2012; Gomez-Mestre et al. 2012) especially in forested regions (Poynton 1964; Gomez-Mestre et al. 2012; Müller et al. 2013), and habitat instability (Crump 1974; Magnusson & Hero 1991). The physical environment may also play a role, such as in montane systems. Here, fast flowing water may not be conducive to aquatic oviposition and fertilization (Goin & Goin 1962), and a seasonal lack of water may actually promote complete terrestrial development (Müller et al. 2013). The shift towards terrestrialized breeding strategies may also coevolve with particular morphological features. It has been demonstrated that frogs with terrestrial oviposition and direct development are smaller in body size compared to aquatic breeding species, and they are often miniaturized (Salthe & Duellman 1973; Clarke 1996; Blackburn 2008; Zimkus et al. 2012). In these smaller species, clutch sizes are reduced and ovum size is increased (Salthe & Duellman 1973).

The repeated independent evolution of terrestrialized reproductive modes within and across major clades of anurans may also be tied to particular ecologies or other features of the environment, such as the water systems (lentic vs. lotic) used for egg deposition or tadpole development. The relationships between these traits have not been explicitly investigated for anurans, and several questions remain. Goin & Goin (1961) first proposed that lotic systems are a selective force leading to terrestrial breeding, and although this idea was initially described for montane systems, it could be reasonably extended to include lowland regions. The same problems of fast flowing water for aquatic oviposition (egg loss) and tadpole development (streamlining and other morphological adaptations) would be consistent regardless of elevation, and it can be predicted that a significant proportion of anurans utilizing lotic systems for breeding should exhibit some degree of terrestrialized breeding. In addition, if particular morphological features related to adult ecology either enable or constrain the evolution of reproductive traits, coevolution of particular ecological and reproductive characters

should occur, and the process may be more prevalent among derived ecomorphologies (e.g., fossorial, arboreal ecomorphs). An alternative test of patterns of coevolution is simply whether the repeated evolution of a particular ecomorphology is accompanied by repeated gains in the same key reproductive traits. Beyond observational data and summaries, there have not been explicit phylogenetic tests for correlations between adult ecology and reproductive mode. Explicit correlated trait analyses can also permit estimation of the temporal sequence of changes in co-evolving characters (Pagel et al. 2004; Pagel & Meade 2006), allowing the assessment of whether changes in one trait (e.g., reproductive mode) are contingent upon the background state of another (e.g., ecology).

The anuran clade Afrobatrachia (sensu Frost et al. 2006) contains ~400 species in four families (Arthroleptidae, Brevicipitidae, Hemisotidae, Hyperoliidae) that together represent over half of the total frog species diversity in sub-Saharan Africa (AmphibiaWeb 2015; Frost 2015). The species in this clade are diverse in terms of their ecomorphologies, behaviors, reproductive modes, and general life histories, and include a number of charismatic taxa such as the fossorial 'rain frogs' (genus *Breviceps*; Fig. 1B), the 'hairy frog' (genus *Trichobatrachus*; Fig. 1G), and the colorful hyperoliid reed frogs (Fig. 1S–W). There is great disparity in the number of species contained in genera, ranging from several monotypic genera to over 140 species in the genus *Hyperolius* (AmphibiaWeb 2015). The main diversity of Afrobatrachians is found primarily in lowland and montane forests where they display a number of adult ecologies including arboreality, terrestriality, and fossoriality. The reproductive modes exhibited by Afrobatrachians range from aquatic oviposition with free-swimming exotrophic tadpoles to direct development, and include a number of intermediate modes with both arboreal and terrestrial oviposition. Direct development appears to have evolved at least twice in this group, occurring in the Brevicipitidae (FitzSimons & Van Dam 1929; Wager 1965; Channing 2001) as well as the arthroleptid genus *Arthroleptis* (Guibé & Lamotte 1958; Wager 1986). Though across families and genera there are a number of distinct ecomorphs and reproductive modes, within particular genera these traits are typically conserved (Schiøtz 1999; Channing 2001, Channing & Howell 2006).

The life history characteristics of Afrobatrachians make them an appropriate system for investigating the evolution of reproductive mode and the relationship between terrestrialized reproductive traits, adult ecology, and water systems. While molecular phylogenetic studies provide consensus on the family-level relationships within the Afrobatrachia, the genus-level relationships are not resolved (Bossuyt et al. 2006; Frost et al. 2006; Roelants et al. 2007; Pyron and Wiens 2011) and therefore the sequence of evolution of these traits remains unknown. In this study, we conduct the most comprehensive analysis of genus-level phylogenetic relationships within the Afrobatrachia by generating new sequence data for five nuclear loci and sampling 52 species across 25 genera. We use the resulting time-calibrated phylogeny to explore the evolution of major reproductive modes in Afrobatrachians and assess the intermediate stages hypothesis for the evolution of direct development. We also examine the evolution of adult ecology, and investigate potential correlations between adult ecology, breeding habitat, and reproductive mode. Our key predictions include: 1) direct development evolved from terrestrialized reproductive modes, 2) species using lotic

systems for reproduction have more terrestrialized breeding strategies, and 3) derived ecomorphologies are likely to be correlated with specific reproductive traits.

Methods

Sampling

To resolve higher-level relationships within the Afrobatrachia, we sampled as many genera as possible within the families Hemisotidae, Brevicipitidae, Arthroleptidae, and Hyperoliidae. This resulted in a total of 52 ingroup samples with multiple representatives from each genus, when possible. We sampled all eight arthroleptid genera, thirteen of eighteen hyperoliid genera, two of five brevicipitid genera, and the single hemisotid genus, for a total of 25 genera. Regarding the genera not included, samples were unavailable for five hyperoliid genera (*Alexteroon*, *Arlequinus*, *Callixalus*, *Chrysobatrachus*, *Kassinula;* totaling 7 species), and we did not densely sample the Brevicipitidae because relationships among these genera were recently published (Loader et al. 2014). Within species-rich genera, we chose species that represent major clades and divergent taxa based on published or unpublished mtDNA barcode data (Blackburn 2008; Blackburn, Portik unpubl. data). In addition, to test the monophyly of the genus *Leptopelis*, we sampled each of the subgenera proposed by Laurent (1941). Because of complications related to polyploidy in the genus *Astylosternus* (Bogart and Tandy 1981), we include only one representative of a diploid species. We included one species (*Phrynomantis microps*) of the family Microhylidae as an outgroup because previous studies demonstrate this family to be the sister clade to the Afrobatrachia (e.g., Bossuyt et al. 2006; Frost et al. 2006; Roelants et al. 2007; Pyron and Wiens 2011). The museum and locality information for all 53 specimens is provided in Appendix 1.

Molecular Data

Whole genomic DNA was extracted from ethanol- or RNALater-preserved liver, muscle, or toe clip samples using a high-salt DNA extraction (Aljanabi & Martinez, 1997). We obtained sequence data for five nuclear loci using primers listed in Table 1: *POMC* (624 bp), *RAG-1* (777 bp), *TYR* (573 bp), *FICD* (524 bp), and *KIAA2013* (540 bp). Polymerase chain reactions (PCRs) were carried out in 12.5 µl volumes consisting of: 1.4 µl Roche 10x (500 mM Tris/HCl, 100 mM KCl, 50 mM $(NH_4)_2$ SO_4, 20 mM $MgCl_2$, pH = 8.3), 1.1 µl 25 mM $MgCl_2$, 0.22 µl 2 mM DNTPs, 0.22 µl 10.0 µM forward primer, 0.22 µl 10.0 µM reverse primer, 7.25 µl H_2O, 1.1 µl betaine, 1.1 µl BSA, 0.08 µl Taq, and 1.0 µl DNA. Thermocycling schemes and amplifications of *FICD* and *KIAA2013* followed the specific nested PCR protocol of Shen *et al.* (2013). Remaining genes were amplified using the following thermocycling protocol: an initial denaturation at 94°C for 2 min, followed by 35 cycles of 94°C for 30 s, a gene-specific annealing temp (55°C: RAG-1, TYR; 51°C: POMC) for 30 s, 72°C for 60 s, and a final extension at 72°C for 10 min. The PCR amplifications were visualized on an agarose gel and cleaned using ExoSAP-IT (USB). Gene products were sequenced using BigDye v3.1 on an ABI3730 (Applied Biosystems). All forward and reverse sequences were assembled using Geneious (Kearse et al. 2012) and aligned using MUSCLE (Edgar 2004), and subsequently translated to ensure conservation of reading frame. The final concatenated alignment of the five loci totals 3,038 bp, with 1150 parsimony informative characters (37.9%) and 135 autapomorphic characters (4.4%), with only 10% missing data. The percentage of

parsimony informative characters to total base pairs per gene was relatively consistent and is as follows: *RAG-1*, 30.1%; *FICD*, 34.9%; *TYR*, 36.3%; *KIAA2013*, 37.5%; *POMC*, 39.7%. All sequences will be deposited in GenBank.

Phylogenetic Analyses

We employed both maximum likelihood and Bayesian phylogenetic analyses, and for model selection we used PartitionFinder (Lanfear *et al.*, 2012) to simultaneously determine our best partitioning strategy and models for each partition subset. The greedy search algorithm was employed, and model selection was conducted using the BIC. The optimal partitioning strategy includes five partitions and models are as follows: partition one (*FICD* codon position (cp) 1, *KIAA2013* cp1, *RAG-1* cp1): GTR+G; partition two (*FICD* cp2, *KIAA2013* cp2, *RAG-1* cp2): HKY+G; partition three (*FICD* cp3, *KIAA2013* cp3, *RAG-1* cp3): HKY+G; partition four (*KIAA2013* cp3, *TYR* cp3): GTR+G; and partition five (*POMC* cp1, *POMC* cp2, *TYR* cp1, *TYR* cp2): HKY+G+I. We used MrBayes v3.2 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003) to execute two parallel runs with four MCMC chains, and analyses were run for 20,000,000 generations with sampling every 1,000 generations, resulting in 20,000 trees. Runs were assessed using Tracer v1.5.0 (Rambaut & Drummond, 2009) to ensure key parameters had reached stationarity (ESS values >150). The first 25% of the total number of generations were discarded as burn-in and a 50% majority rule consensus tree was calculated from the combined remaining trees of both parallel runs. We performed maximum likelihood analyses of the partitioned data set using GARLI v2.0 (Zwickl, 2006) to execute 1,000 nonparametric bootstrap replicates using the default stopping criteria. A 50% majority rule consensus tree was generated from the 1,000 replicates.

Divergence Dating Analyses

Dating analyses were carried out using BEAST v1.8 (Drummond et al. 2012) using multiple calibration strategies based on results from Roelants et al. (2007), Kurabayashi and Masayuki (2013), and Loader et al. (2014). We used four secondary calibration points constraining the most recent common ancestors (MRCAs) of (Arthroleptidae + Hyperoliidae + Hemisotidae + Brevicipitidae) to 92.8 Ma ± 5.0 SD, (Hemisotidae + Brevicipitidae) to 65.9 Ma ± 6.5 SD, and Brevicipitidae to 45.4 Ma ± 8.0 SD. Normal distributions were applied to calibrated nodes. To explore the effects of calibration strategy, we used all permutations of the three calibrations, including single calibration analyses, for a total of seven unique calibrated analyses. Each analysis was run 20,000,000 generations with sampling every 2000 generations. For all analyses, we used the Yule model of speciation as the tree prior, applied an uncorrelated relaxed lognormal clock, and unlinked clock and substitution models. Substitution models and partitioning schemes were identical to those used in phylogenetic analyses. Runs were assessed using Tracer v1.5.0 (Rambaut & Drummond, 2009) to examine convergence. A burn-in of 25% was discarded and maximum clade credibility trees were created from the remaining 7,500 trees. For nodes of interest, divergence date means, confidence intervals, and marginal density proportions were obtained from each of the seven calibrated analyses.

Ancestral Character Estimations and Trait Correlations

We collected data for several character sets focused on ecological and reproductive traits, which were coded as multistate discrete characters. These character sets include adult ecology (fossorial, terrestrial, arboreal), oviposition site (aquatic, subterranean, terrestrial, arboreal), tadpole water system (lentic, lotic, or none in the case of direct development), and reproductive mode. There are many ways to categorize the reproductive modes of amphibians (reviewed in Crump 2015), but here we focus on the terrestrialization of breeding strategy, which examines the combination of oviposition site and the site of tadpole development. We categorize reproductive mode to include the following four states: 1) aquatic eggs and aquatic larvae, 2) terrestrial eggs and aquatic larvae, 3) arboreal eggs and aquatic larvae, and 4) terrestrial eggs and terrestrial larvae (direct development). A summary of the character data is provided in Appendix 2.

The ancestral states of adult ecology and reproductive mode were examined using three approaches, including maximum likelihood (ML), stochastic character mapping, and Bayesian inference (BI). Using the ACE function of the APE package in R (Paradis et al. 2004), the likelihoods of three transition models (equal rates [ER], symmetrical rates [SYM], or all rates different [ARD]) were determined by maximum likelihood, and subsequent likelihood ratio tests were conducted to determine the best fitting model for each character set. The best fitting models were used for all subsequent analyses. Maximum likelihood estimation of ancestral states was carried out using ACE to determine the marginal ancestral state reconstruction of the root and the conditional scaled likelihoods of all remaining nodes. To estimate the marginal probabilities for all nodes based on joint sampling, stochastic mapping was implemented using the make.simmap function of the phytools package in R (Revell 2012). To account for branch length and topological uncertainty, mapping was performed using 100 replicates on 100 randomly selected trees from the posterior distribution of trees resulting from BEAST, resulting in a total of 10,000 mapped trees. The number of transitions between character states and the proportion of time spent in each state were summarized using phytools. Finally, the root state was estimated in a Bayesian framework using BayesTraits V2 (Pagel et al. 2004). Alternative root states were compared by fixing the root state using the 'fossil node' implementation for model testing, and an unconstrained analysis was performed to obtain the posterior probability of the root state. Under the multistate model of evolution, analyses were run using reversible jump MCMC for 50,000,000 post-burn-in generations, sampling every 5,000 generations. Marginal likelihoods were calculated using stepping-stone sampling with 100 samples and 10,000 iterations per sample. Alternative hypotheses were compared using log Bayes Factors (logBF= 2(log marginal likelihood[model 1] - log marginal likelihood[model 2]), with logBF > 5 considered strong evidence. Analyses incorporated the topologies of 100 randomly selected trees from the posterior distribution obtained from BEAST to account for topological and branch length uncertainty.

To investigate potential correlations between adult ecology, oviposition site, and water systems, correlations between characters were examined using BayesTraits V2 (Pagel et al. 2004). The discrete dependent and discrete independent models of evolution were used to examine correlations between character states across different character sets. Correlations were investigated between states of adult ecology and oviposition site, adult ecology and tadpole water system, and oviposition site and tadpole water system.

Analyses were performed using maximum likelihood and Bayesian inference. Maximum likelihood analyses were conducted using 200 attempts per tree, across 100 randomly selected trees from the posterior distribution obtained from BEAST. Bayesian analyses were run using reversible jump MCMC for 50,000,000 post-burn-in generations, sampling every 5,000 generations. Marginal likelihoods were calculated using stepping-stone sampling with 100 samples and 10,000 iterations per sample. The alternative hypotheses of dependent vs. independent models of evolution were compared using logBF, where logBF = 2(log marginal likelihood[dependent model] - log marginal likelihood[independent model]). Analyses incorporated the topologies of 100 randomly selected trees from the posterior distribution of trees resulting from BEAST.

For characters states exhibiting dependent evolution, the rates and z-scores of key character transitions (q_{12}, q_{13}, q_{24}, q_{34}) were estimated (Pagel & Meade 2006). The z-scores were calculated from the proportion of the number of occurrences of a value of zero estimated for relevant transition parameters, and are expressed as a percentage. The z-scores and transition rate estimates allow examination of the contingency of evolutionary changes, with high z-scores and low rates signifying unlikely paths for evolutionary transformations of character pairs. In this framework, it can be evaluated if state changes of a particular character are contingent upon the background state of the other character.

Results

Phylogenetic Relationships

The phylogenetic relationships within the Afrobatrachia are concordant across analyses, and both the monophyly of and relationships among the four Afrobatrachian families are resolved with high support (Fig. 1). A sister relationship is recovered with high support for Hemisotidae and Brevicipitidae, as well as for Arthroleptidae and Hyperoliidae. Within the family Arthroleptidae, four major clades are consistently recovered: the genus *Leptodactylodon*, the genus *Leptopelis*, a clade containing *Nyctibates, Scotobleps, Astylosternus,* and *Trichobatrachus*, and a clade containing *Cardioglossa* and *Arthroleptis* (Fig. 1). The relationships among these four clades are not fully resolved, though some analyses recover the genus *Leptodactylodon* as sister to the other two major clades (Fig. 1). The sister relationship between the genus *Leptopelis* and the clade consisting of *Cardioglossa* and *Arthroleptis* is recovered in most analyses. The species rich genera (*Arthroleptis, Cardioglossa,* and *Leptopelis*) are each recovered as monophyletic in all analyses. The species *Leptopelis parkeri* is found to be sister to and highly divergent from a clade containing all other *Leptopelis* (Fig. 1).

Within the Hyperoliidae, there is strong support for two subclades comprising, respectively, the "kassinoids" and all other hyperoliids (Fig. 1). The kassinoid genera represented in our data set include *Semnodactylus, Paracassina, Phlyctimantis,* and *Kassina*. Within the kassinoids, we find that *Kassina maculata* is more closely related to *Phylctimantis* than to other species of *Kassina*. The genus *Acanthixalus* and *Opisthothylax* form the earliest diverging lineages of the remaining hyperoliids. *Heterixalus* and *Tachycnemis*, endemic to the islands of Madagascar and the Seychelles respectively, are recovered as sister taxa, and are in turn sister to the species rich genus *Afrixalus*, which is recovered as monophyletic. The genera *Cryptothylax* and *Morerella* form a strongly supported clade that is sister to the last clade comprising the genus

Hyperolius and the monotypic genus *Chlorolius*, which is embedded well within *Hyperolius*. Relationships within the species-rich genus *Hyperolius* are not strongly resolved.

Divergence Dating Estimates

The dating analyses resulting from the seven calibration strategies produced identical topologies and similar mean estimates for dates (Fig. 1, Table 2). In general, the analyses incorporating fewer calibration points resulted in dating estimates with larger 95% highest posterior density regions (95% HPD), with little difference in means (Table 2). We therefore present results from the analysis incorporating all calibrations (Afr, [H+B], B; Table 2; Fig. 1). The estimated branch-specific clock rates varied across the phylogeny, but the highest clock rates are inferred for the branch leading to the genus *Leptopelis* (excluding *L. parkeri*), and the branch leading to the clade containing *Arthroleptis* and *Cardioglossa*.

The estimates for the TMRCA of the families Hyperoliidae and Arthroleptidae are similar, being approximately 56.6 Ma (95% HPD: 47.0–66.6 Ma) and 53.7 Ma (95% HPD: 43.9–63.9 Ma), respectively (Fig. 1, Table 2). Early diversification events in the species-rich genera appear to be concentrated in the Late Oligocene and Early Miocene (Fig. 2). Diversification began to occur in the genus *Arthroleptis* approximately 16.2 Ma (95% HPD: 11.6–21.2 Ma) and in *Cardioglossa* approximately 14.6 Ma (95% HPD: 9.6–19.9 Ma). *Leptopelis parkeri*, endemic to the Eastern? Arc Mountains, is estimated to have diverged from all other *Leptopelis* approximately 39.5 Ma (95% HPD: 31.0–48.7 Ma), with diversification of the remaining species beginning around 10.3 Ma (95% HPD: 7.0–14.0 Ma). Within hyperoliids, the TMRCA for the genus *Afrixalus* is estimated to be 17.6 Ma (95% HPD: 13.2–22.4 Ma), and diversification in *Hyperolius*, the largest Afrobatrachian genus, appears to have begun approximately 23.3 Ma (95% HPD: 18.3–28.3 Ma).

The Malagasy and Seychelles Island hyperoliids, comprised of the genera *Heterixalus* and *Tachycnemis*, appear to have diverged from their closest mainland African relatives (*Afrixalus*) approximately 26.1 Ma (95% HPD: 20.4–32.1 Ma). This divergence represents a major overseas dispersal event, which is supported by favorable oceanic currents in the Mozambique Channel prior to the Miocene (Ali & Huber 2010; Samonds et al. 2012; Tolley et al. 2013). Pyron (2014) estimated this divergence to be 51 Ma, and although our results are much younger, both fall in the expected oceanic current time interval predicted (> 24 Ma). *Heterixalus* and *Tachycnemis* are estimated to have diverged from one another 7.6 Ma (95% HPD: 3.8–11.6 Ma).

Evolution of Adult Ecology and Reproductive Mode

Comparisons of the transition rate models revealed that more complex models (SYM, ARD) were not significantly better fits than a simple one-rate model (ER); therefore the ER transition model was used for subsequent analyses (Table 3). Maximum likelihood ancestral character estimations of adult ecology estimate that the state of the common ancestor of the Afrobatrachia was terrestrial (marginal probability: 0.62) (Table 4, Fig. 2). Bayesian hypothesis testing performed by fixing root states revealed no significant differences in marginal likelihoods across models (fossorial: -34.17, arboreal: -37.01, terrestrial: -36.01; logBF < 5), though the unconstrained multistate analysis recovered the

root state as fossorial with a posterior probability of 0.81. Accounting for topological uncertainty, the average number of transitions to fossoriality is approximately one (common ancestor of hemisotids and brevicipitids), with one reversal back to terrestriality (brevicipitid genus *Callulina*) (Fig. 3A). The most common ecological transition is from terrestriality to arboreality, and it has occurred independently at least four times (Fig. 3A). The transitions have taken place in the families Arthroleptidae (*Leptopelis*) and Hyperoliidae (*Phlyctimantis*, *Kassina arboricola*, mrca of non-kassinoid hyperoliids) (Fig. 2). Transitions between arboreality and fossoriality do not appear to have occurred within Afrobatrachia (Figs. 2, 3A).

Ancestral character estimates of reproductive mode provide strong evidence for a terrestrialized breeding strategy present in the common ancestor of the Afrobatrachia, specifically terrestrial oviposition with aquatic larvae (Fig. 2). Comparisons of the marginal likelihoods of fixed root states inferred by Bayesian analysis support a terrestrial oviposition/aquatic larvae root state (terrestrial/aquatic: -36.87, aquatic/aquatic: -39.48, terrestrial/terrestrial: -39.90, arboreal/aquatic: -38.88; logBF: 5.22, 4.02, 6.06, respectively), and maximum likelihood analyses estimate this root state with a marginal probability of 0.57 (compared to 0.18, 0.17, and 0.08 of other states, Table 4). Likewise, the unconstrained multistate Bayesian analysis recovered the root state as terrestrial oviposition/aquatic larvae with a posterior probability of 0.62. The number of transitions between these four reproductive modes varies (Fig. 3B), but few transitional paths occur frequently. The complete terrestrial reproductive mode (direct development) arises primarily from the terrestrial/aquatic reproductive mode (1.7 transitions), rather than other states (< 0.3), and is confirmed as originating independently in the family Brevicipitidae and the arthroleptid genus *Arthroleptis* (Figs. 2, 3B). Transitions to the arboreal/aquatic reproductive mode and aquatic/aquatic mode also seem to occur disproportionately from the terrestrial/aquatic reproductive mode (0.7 transitions and 1.3 transitions, respectively). Reversals back to terrestrial oviposition and aquatic larvae from the other three reproductive modes are estimated in very low frequency, suggesting they are not likely to have occurred (Fig. 3B). There are at least two reversals from arboreal oviposition to aquatic oviposition (both with aquatic larvae), both of which occur in the diverse hyperoliid genus *Hyperolius* (Fig. 2).

Character Correlations

The exploration of correlations between adult ecology, oviposition site, and larval water system revealed several sets of significantly correlated characters (Table 5). There is a significant correlation between arboreal ecology and arboreal oviposition (ML likelihood ratio = 14.12, p = 0.007; BI logBF = 6.12), and also between fossorial ecology and subterranean oviposition (ML likelihood ratio = 8.84, p = 0.065; BI logBF = 10.56). The transition parameter rates and z-scores for arboreal ecology and arboreal oviposition reveal that the most likely pathway was a gain of arboreality, followed by a gain of arboreal oviposition (Table 6, Fig. 3C). In contrast, it is most likely that the evolution of subterranean oviposition occurred before the gain of fossoriality (Table 6). No significant correlations between terrestrial ecology and a particular oviposition site were detected (Table 5). With regard to lentic and lotic water systems for larval development, there is support for a link between lotic systems and terrestrial oviposition (BI logBF = 5.24), and for an association between lentic systems and an arboreal ecology (ML likelihood ratio =

9.06, p = 0.060) (Table 5). The z-scores for transition parameters for terrestrial oviposition and lotic systems indicate a shift in oviposition before a transition to lotic systems is not well supported (z-score: 96%), but a move to lotic systems followed by a shift in oviposition is likely (z-score: 1%).

Discussion

Evolution of Direct Development in Afrobatrachians

Among the 40 described reproductive modes for anurans, the most commonly observed mode is the plesiomorphic state involving aquatic oviposition followed by an exotrophic free-swimming tadpole stage (Duellman 1985), a condition recovered as the basal state for a majority of frog families (Gomez-Mestre et al. 2012). Our phylogenetic model testing strongly suggests the common ancestor of the 400 species of Afrobatrachians did not possess this condition, but rather had a reproductive mode of terrestrial eggs and a free-swimming exotrophic tadpole (Fig. 2). Perhaps this should not come as a great surprise as the family Microhylidae, the sister taxon of Afrobatrachia, also exhibits substantial reproductive mode diversity with many terrestrial and fossorial lineages. The presence of terrestrial eggs with aquatic larvae in hemisotids and a majority of arthroleptids is therefore best explained by the retention of this state as opposed to convergence. Conversely, the origin of direct development in Afrobatrachians is a clear example of convergent evolution (Fig. 2). The only scenario in which this trait could result from the retention of the ancestral condition calls for direct development being the ancestral state of the Afrobatrachians, and our model testing strongly rejects this hypothesis in favor of terrestrial eggs and aquatic larvae. Our data indicate that the two independent origins of direct development both involved transitions from the ancestral state of terrestrial eggs with aquatic larvae.

The 'intermediate stages' hypothesis posits that the evolution of direct development should proceed in a sequence through prior stages involving progressively more terrestrialized characteristics, though not requiring a linear trajectory through all possible reproductive modes (Salthe & Duellman 1973; Duellman & Trueb 1986; Wells 2007; Crump 2015). The independent evolution of direct development in Brevicipitidae and the genus *Arthroleptis* support this hypothesis, as the sister taxa of these two groups (Hemisotidae and *Cardioglossa*, respectively) each exhibit a terrestrialized reproductive mode. Though we had no temporal predictions for the evolution of direct development, the condition appeared much earlier in Brevicipitidae (~42 Ma) than in the genus *Arthroleptis* (~16 Ma). Another implicit notion of the 'intermediate stages' hypothesis is that reproductive modes in which eggs or larvae are protected (deposited away from main water bodies, protected from predators, parental care, etc.) should also give rise to direct development. However, across global amphibian diversity Gomez-Mestre et al. (2012) demonstrated that modes with unprotected eggs and larvae gave rise to direct development more often than modes with protected eggs or larvae. The reproductive mode present in Brevicipitidae involves the creation of a subterranean nest in which the eggs are deposited and fertilized, and the female often remains in the chamber or resides in a nearby connected chamber (Wager 1965, 1986). The young either develop completely within the egg, or hatch out as non-feeding tadpoles that remain in an associated jelly mass to complete their development (FitzSimons & Van Dam 1929; Wager 1965, 1986; Channing 2001). Hemisotidae, the sister family of Brevicipitidae, has

a similar reproductive mode involving the building of a subterranean chamber. The egg mass is topped with empty jelly capsules, and the female actively attends the eggs by sitting on them (Wager 1965, 1986). Around the time of hatching, the female digs a tunnel to the nearby water source that ultimately floods the nest, allowing the tadpoles to leave the chamber (Wager 1965, 1986). The high degree of the protection of eggs and larvae in this group, both through the creation of a subterranean chamber and parental care, is the context in which direct development evolved in Brevicipitidae. The genus *Arthroleptis* deposits eggs under leaf litter or in shallow terrestrial chambers (Guibé & Lamotte 1958; Wager 1965; Channing 2001; Channing & Howell 2006), and the metamorphosed young emerge from the unattended eggs. The sister genus to *Arthroleptis* is *Cardioglossa*, and although the breeding biology of this genus is less characterized, these species appear to attach eggs to the undersides of rocks in damp areas away from the main water bodies (Amiet 1972). The tadpoles eventually emerge and find their way to the adjacent water source. The evolution of direct development in the arthroleptids occurs in the context of the protection of eggs, here by the use of terrestrial oviposition sites. Our results are therefore consistent with the notion that reproductive modes involving the protection of eggs or larvae give rise to direct development, a pattern that requires further fine-scale testing among other major clades that have evolved this derived reproductive mode.

Evolution of Reproductive Mode and Water System Preference

The terrestrialization of anuran breeding strategy has been proposed to be a result of selective pressures from flowing water associated with lotic ecosystems (Alcala 1962; Goin & Goin 1962), an alternative to the hypothesis that aquatic predation in the egg and larval stage is the main selective pressure driving terrestrialized reproduction (Tihen 1960; Poynton 1964; Magnusson & Hero 1991; Haddad & Prado 2005). Although initially characterizing montane systems (where the prevalence of lotic systems is more pronounced), there are a variety of lotic habitats occurring in the lowlands that present similar challenges for egg deposition, fertilization, and tadpole development. If the exploitation of lotic habitats requires some degree of reproductive terrestrialization, these patterns should emerge through broad scale comparisons of the evolution of reproductive mode and water system preference, regardless of elevation. Based on our analyses of the coevolution of reproductive traits and water systems, we find support for the correlated evolution of terrestrial oviposition and lotic system habitat preference (Table 5, Fig. 2). This type of oviposition includes shallow mud nests, leaf litter nests, and egg-deposition in rock crevices out of main water bodies, but does not include arboreal egg sites (such as leaves overhanging water). Though requiring more exploration, this correlation suggests that oviposition site is linked with flowing water habitats in both low and high elevations, and that habitat filtering may be promoting this reproductive trait. This pattern is most apparent for the family Arthroleptidae, in which species mainly breed in streams and rivers, and a majority exhibit terrestrial oviposition (Fig. 2). Interestingly, the two arthroleptid genera exhibiting aquatic oviposition (*Trichobatrachus*, *Astylosternus*) also use lotic habitats for breeding, and these species display reproductive characteristics that are adaptive for streams, including the attachment of egg masses to rocks in deeper pools (Rödel et al. 2012) and underwater egg attendance?. Regardless of oviposition site, many of the arthroleptid species breeding in stream habitats also exhibit fast-flowing water

adaptations in tadpoles, including streamlining or semi-fossoriality, highlighting the importance of water systems for the evolution of this clade.

Relationship of Ecomorphology and Reproductive Characteristics

Quantitative analyses of anuran reproductive biology reveal important relationships between body size, clutch size, ovum size, and reproductive mode (Salthe & Duellman 1973). Reproductive traits may coevolve with ecomorphologies, but the relationship between adult ecology and reproductive mode has not been examined in a comparative phylogenetic framework. Here, we used a generalized classification of adult ecology (terrestrial, fossorial, arboreal) to investigate whether the evolution of oviposition site or water system preference is correlated with particular ecologies. We find evidence for the dependent evolution of fossoriality and subterranean oviposition, as well as arboreality and arboreal oviposition (Table 5). The former result may seem intuitive, but most fossorial anurans actually exhibit aquatic oviposition, breeding in temporary pools (Rhinophrynidae, Pelobatidae, Scaphiopodidae, species within Myobatrachidae and Odontophrynidae) or streams (Nasikabatrachidae). Other examples of fossorial ecomorphs with subterranean nests are all direct-developing species, including the monotypic myobatrachid genus *Myobatrachus* from the arid regions of Australia (Roberts 1981) and some asterophrynine microhylids and terraranans. The correlation found in Afrobatrachians between fossoriality and subterranean oviposition appears to be a clade-specific pattern rather than a global one, but is an important development in the evolution of hemisotid and brevicipitid frogs and their subsequent diversification.

The arboreal ecomorph has at least four independent origins within Afrobatrachia, offering an opportunity to examine the repeated evolution of any associated reproductive traits. We find evidence for the dependent evolution of arboreality and arboreal oviposition (Table 5), and the most likely transitional pathway first involves a shift to arboreality, followed by the evolution of arboreal oviposition (Fig. 3C). The non-kassinoid hyperoliids exhibit arboreality and arboreal oviposition (Fig. 2), but there are some arboreal groups that retain the plesiomorphic aquatic oviposition site of their common ancestor. These include the two transitions to arboreality in the kassinoids (genus *Phlyctimantis*, *Kassina arboricola*). These species attach their eggs to underwater vegetation (Schiøtz 1999; Channing & Howell 2006), which is the condition reconstructed for the common ancestor of this clade (Fig. 2). The other example is the arthroleptid genus *Leptopelis*, in which species deposit their eggs in a mud depression or just beneath the soil surface adjacent to the main water body. Terrestrial oviposition is common among the arthroleptids and is reconstructed as the ancestral state of the family, and appears to be retained in arboreal *Leptopelis* species (Fig. 2).

Though the evolution of arboreality is not always accompanied by a shift to arboreal oviposition, there are no examples of terrestrial species of Afrobatrachians that have arboreal oviposition. This evidence, in conjunction with the transformational pathway recovered for the coevolution of these traits, suggests that arboreality is a preadaptation for arboreal oviposition in Afrobatrachians. The sheer number of arboreal Afrobatrachian species with arboreal oviposition (close to 190 species) far exceeds the number of arboreal Afrobatrachian species with aquatic (5 species) or terrestrial oviposition (50 species), strongly suggesting arboreal oviposition has played a key role in the diversification of hyperoliids. The influence of this reproductive trait on speciation

and extinction rates is unclear, but trait-dependent diversification can be more appropriately investigated as species-level sampling improves.

Perhaps the most surprising revelation is the correlation between the evolution of arboreality and use of lentic water systems for breeding (Table 5), which appears to be convergent and largely consistent across Afrobatrachians regardless of oviposition sites (Fig. 2). The association of lentic systems for breeding and arboreality could be explained by a number of selective pressures, but may be most heavily driven by aquatic predation of the eggs or young larvae (Lutz 1948; Crump 1974; Magnusson & Hero 1991; Haddad & Prado 2005). In addition to reducing predation on egg masses, the delay in hatching from non-aquatic oviposition sites often produces larger tadpoles (Salthe & Duellman 1973; Duellman & Trueb 1986), which could reduce predation pressure. The availability of suitable habitat for breeding may also help explain this association. Many hyperoliids call from emergent vegetation in ponds, which are often the same sites used to place eggs overhanging the water (Schiøtz 1999; Channing 2001; Channing & Howell 2006). This emergent vegetation can be more prevalent in lentic systems and is often absent on the banks of swiftly moving streams, though some hyperoliid species have adapted to breeding in lotic systems (the use of large bush and tree sites by *Afrixalus laevis* and *A. lacteus*, for example). Further work is required to determine if this association is clade-specific, African-specific, or truly represents a global pattern for anurans. There are other major radiations of arboreal frogs occurring in South America (Hylidae) and Asia (Rhacophoridae), and these clades can provide independent faunal and continental comparisons to test for the association between arboreality and lentic water systems.

Reproductive Resource Partitioning

Reproductive partitioning typically occurs when potentially interacting species create reproductive interference, and the result is the differentiation of reproductive traits among species. The resulting partitioning can be manifested temporally, spatially, or phenotypically, and in the last case divergences in behavioral, morphological, or physiological traits can minimize reproductive interactions (Pfennig & Pfennig 2012). The reproductive mode of anurans is another complex phenotypic trait that can increase or decrease resource competition among species, particularly with regard to breeding habitats. In principle, there should be limits to the similarity of reproductive traits among co-occurring frog species, which could be related to acoustic space, calling site, oviposition site, water body preference, requirements for tadpole development, and/or parental care. It is possible that changes in these reproductive traits result from interactions with heterospecifics, but given that many clades exhibit long-term retention of reproductive modes (our results, Gomez-Mestre et al. 2012), it is more likely that species sorting drives large-scale patterns of anuran communities. In the latter scenario, the coexistence of species is facilitated by their possession of divergent reproductive traits, but these traits did not evolve through competitive heterospecific interactions (such as character displacement), but rather other selective pressures.

Among the African frogs, Afrobatrachians dominate tropical communities in terms of species composition and abundance, often comprising over 60% of the total species occurring in sites (Portik & Blackburn, unpubl. data). Our results suggest the early evolution of divergent reproductive traits in these four families may have played a role in facilitating their co-occurrence. Perhaps the most striking example of

reproductive partitioning is direct development, as interspecific competition for oviposition sites, water resources, and tadpole development would effectively be reduced or eliminated. Beyond the direct developers, the partitioning of particular water bodies for breeding would greatly reduce reproductive interference, and a majority of arthroleptid frogs (with the exception of *Leptopelis*) utilize lotic systems for breeding, whereas hyperoliid frogs mainly breed in lentic systems (Fig. 2). These preferences, in combination with additional divergences in oviposition site, essentially limit the competition for breeding habitat resources at the clade-level. In the case of *Leptopelis*, the genus has shifted to using lentic systems, which are generally dominated by arboreal hyperoliids with arboreal oviposition. The retention of the plesiomorphic terrestrial oviposition condition in *Leptopelis* may therefore be a consequence of trade-offs between the selective pressures for arboreal oviposition at lentic sites and interspecific competition with hyperoliid frogs for such arboreal sites, or alternatively may simply represent a phylogenetically conserved trait.

Though the early divergence in reproductive modes almost certainly played a role in shaping African anuran communities through species sorting and competitive exclusion, species interactions are likely influencing fine-scale patterns of reproductive partitioning in these communities. The levels of co-occurrence of hyperoliid frogs are remarkable for amphibians, with up to 14 species found in near sympatry (Lawson & Collett 2012), and these species have similar reproductive traits. Given the high number of co-occurring species, it is reasonable to expect interspecific competition for reproductive resources among these species likely promoted divergence among traits, including acoustic space, calling sites, oviposition sites, or larval features. Hyperoliids are a promising system for investigating the occurrence of reproductive character displacement resulting from reproductive resource competition, a topic that remains to be explored not only for African frog communities, but also globally.

Systematics and Taxonomy

Our analyses support the family-level relationships found previously, including sister relationships between Brevicipitidae and Hemisotidae as well as Arthroleptidae and Hyperoliidae (Frost et al. 2006; Roelants et al. 2007; Pyron and Wiens 2011). Yet within these families, certain genus-level relationships differ from those previously obtained. For the most part, it is the handful of monotypic or low diversity genera that remain difficult to place with high confidence in the phylogeny, but also extends to the diverse and widespread genus *Leptopelis*.

Previous work on the Hyperoliidae has reconstructed different topologies among both the 'kassinoid' and other genera. Our work extends previous analyses by including two additional 'kassinoid' genera, *Semnodactylus* and *Paracassina*, which are revealed as the earliest diverging lineages of the kassinoid clade (Fig. 1). In contrast to the analysis of morphological characters by Drewes (1984), we find *Paracassina* (his *Tornierella*) to be sister to a clade containing *Kassina* and *Phylctimantis*, and *Semnodactylus* (his *Kassina wealei*) sister to all other 'kassinoids.' Frost et al. (2006) found the hyperoliid *Acanthixalus* to be sister to the 'kassinoids,, whereas our analyses place this genus as sister to the large clade of the remaining hyperoliids (Fig. 1). In contrast to Veith et al. (2009), we find a strong sister relationship between *Cryptothylax* and *Morerella*. Previous analyses have also found discordant patterns in the relationships among the

genera of the Arthroleptidae (Frost et al. 2006; Blackburn 2008; Pyron and Wiens 2011). In contrast to previous analyses (Frost et al. 2006; Blackburn, 2008), the monotypic genus *Scotobleps* is strongly supported as forming a clade with *Astylosternus, Trichobatrachus*, and *Nyctibates* (Fig. 1). These relationships are significant for interpreting phenotypic evolution because the three genera (*Astylosternus*, *Scotobleps*, and *Trichobatrachus*) that have curved terminal pedal phalanges likely used in defense (Blackburn et al. 2008) form a strongly supported clade. The widespread and diverse genus *Leptopelis* is supported as the sister genus of *Arthroleptis* and *Cardioglossa* (similar to Roelants et al. 2007), rather than sister to all arthroleptids (e.g., Frost et al. 2006) or forming a clade with 'astylosternines' such as *Trichobatrachus* (Bossuyt et al. 2006; Pyron and Wiens 2011).

Consistent with other previous analyses (e.g., Pyron and Wiens 2011), we find *Kassina maculata* to be more closely related to species of *Phlyctimantis*. Based on this, we transfer *Kassina maculata* to the genus *Phlyctimantis* (*Phlyctimantis maculata* comb. nov.). Similarly, our phylogenetic analyses support Amiet's (2012) decision to recognize *Chlorolius koehleri* as a member of the genus *Hyperolius*, though evaluation of the relationships among subgenera of *Hyperolius* will require a more extensive sampling of the species diversity in this genus. Our analyses also reveal that the subgenus *Laurentixalus*, recently erected by Amiet (2012) for *Afrixalus laevis* and *A. lacteus*, is embedded within the diversity of other species of *Afrixalus,* but is itself monophyletic.

Within the Arthroleptidae, our analyses reveal substantial divergence within the genus *Leptopelis* (Fig. 1). Laurent (1941) recognized three subgenera within *Leptopelis* (*Elaphromantis*, *Heteropelis*, and *Taphriomantis*) and we sampled the type species of each (*L. bocagii* for *Taphriomantis*, *L. notatus* for *Elaphromantis*, and *L. parkeri* for *Heteropelis*). While four of the *Leptopelis* species sampled appear to have shared a common ancestor in the Miocene, our analyses suggest that *L. parkeri* diverged from the other species sampled in the Eocene. This deep divergence and a number of skeletal differences noted by Laurent (1941) are compelling evidence for recognizing a separate genus, and additional work underway on this species group will clarify its taxonomic status (Loader et al., in prep).

Anticipated coauthorship: David C. Blackburn.

References

Alcala, A.C. 1962. Breeding behavior and early development of frogs of Negros, Phillipine Islands. Copeia 1962:679–726.

Ali J.R., and M. Huber. 2010 Mammalian biodiversity on Madagascar controlled by ocean currents. Nature 463:653–680.

Aljanabi, S., and I. Martinez. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research 25:4692–4693.

Altig, R., and B.I. Crother. 2006. The evolution of three deviations from the biphasic anuran life cycle: alternatives to selection. Herpetological Review 37:321–325.

Amiet, J.-L. 1972. Le Cardioglossa camerounaises.-Science et Nature, Paris 114:11–24.

Amiet, J.-L. 2012. Les Rainettes du Cameroun (Amphibiens Anoures). Saint-Nazaire, France: La Nef des Livres.

AmphibiaWeb. 2015. Information on amphibian biology and conservation. Berkeley, CA. Available from: URL http://www.amphibiaweb.org/. Accessed September 2015.
Boulenger, G.A. 1886. Remarks in connection with the preceding note. Annals and Magazine of Natural History, series 5, 17:463–464.
Blackburn, D.C. 2008. Biogeography and evolution of body size and life history of African frogs: Phylogeny of squeakers (*Arthroleptis*) and long-fingered frogs (*Cardioglossa*) estimated from mitochondrial data. Molecular Phylogenetics and Evolution 49:806–826.
Bogart, J. P., and M. Tandy. 1981. Chromosome lineages in African ranoid frogs. Monitore Zoologico Italiano. Nuova Serie, Supplemento. Firenze 15:55–91.
Bossuyt, F., R.F. Brown, D.M. Hillis, D.C. Cannatella, and M.C. Milinkovitch. 2006. Phylogeny and biogeography of a cosmopolitan frog radiation: Late Cretaceous diversification resulted in continent-scale endemism in the family Ranidae. Systematic Biology 55:579–594.
Channing, A. 2001. Amphibians of Central and Southern Africa. Cornell University Press, Ithaca, New York, USA.
Channing, A., and K.M. Howell. 2006. Amphibians of East Africa. Edition Chimaira, Frankfurt, Germany.
Clarke, B.T. 1996. Small size in amphibians–its ecological and evolutionary implications. Symposia of the Zoological Society of London 69:201–224.
Crump, M.L. 1974. Reproductive strategies in a tropical anuran community. University of Kansas Museum of Natural History Miscellaneous Publication No. 61: 1–68.
Crump. M.L. 2015. Anuran reproductive modes: evolving perspectives. Journal of Herpetology 49:1–16.
da Silva, F.R., M. Almeida-Neto, V.H.M. do Prado, C.F.B. Haddad, D.C. Rossa-Feres. 2012. Humidity levels drive reproductive modes and phylogenetic diversity of amphibians in the Brazilian Atlantic forest. Journal of Biogeography 39:1720–1732.
Drewes, R.C. 1984. A phylogenetic analysis of the Hyperollidae (Anura): treefrogs of Africa, Madagascar, and the Seychelles Islands. Occasional Papers of the California Academy of Sciences 139:1–70.
Drummond, A.J., M.A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29:1969–1973.
Duellman, W.E. 1985. Reproductive modes in anuran amphibians: phylogenetic significance of adaptive strategies. South African Journal of Science 81:174–178.
Duellman, W.E. 2007. Amphibian life histories: Their utilization in phylogeny and classification. Pp. 2843–2892 in H. Heatwole and M.J. Tyler (eds.), Amphibian Biology. Volume 7, Systematics. Surrey Beatty & Sons, NSW, Australia.
Duellman, W.E., and L. Trueb. 1986. Biology of Amphibians. McGraw-Hill, New York, USA.
Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792–1797.
FitzSimons, V., and G. Van Dam. 1929. Some observations on the breeding habitats of *Breviceps*. Annals of the Transvaal Museum 13:152–153.

Frost, D.R., T. Grant, J. Faivovich, R.H. Bain, A. Haas, C.F.B. Haddad, R.O. de Sa, A. Channing, M. Wilkinson, S.C. Donnellan, C.J. Raxworthy, J.A. Campbell, B.L. Blotto, P. Moler, R.C. Drewes, R.A. Nussbaum, J.D. Lynch, D.M. Green, W.C. Wheeler. 2006. The amphibian tree of life. Bulletin of the American Museum of Natural History 297:1–370.

Goin, O.B., and C.J. Goin. 1962. Amphibian eggs and the montane environment. Evolution 16:364–371.

Gomez-Mestre, I., R.A. Pyron, and J.J. Wiens. Phylogenetic analyses reveal unexpected patterns in the evolution of reproductive modes in frogs. Evolution 66:3687–3700.

Guibé, P.J. and M. Lamotte. 1958. Morphologie et reproduction par développement direct d'un anoure du Mont Nimba, *Arthroleptis crusculum* Angel. Bulletin de Muséum National d'histoire Naturelle, Série 2, 2:125–133.

Haddad, C.F.B., and C.P.A. Prado. 2005. Reproductive modes in frogs and their unexpected diversity in the Atlantic forest of Brazil. BioScience 55:207–217.

Huelsenbeck, J.P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Iskandar, D.T., B.J. Evans, and J.A. McGuire. 2014. A novel reproductive mode in frogs: A new species of fanged frog with internal fertilization and birth of tadpoles. PLoS ONE 10(3):e0119988.

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

Kurabayashi, A., and M. Sumida. 2013. Afrobatrachian mitochondrial genomes: genome reorganization, gene rearrangement mechanisms, and evolutionary trends of duplicated and rearranged genes. BMC Genomics 14:633.

Lanfear, R., B. Calcott, S.Y.W. Ho, and S. Guindon. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29:1695–701.

Laurent, R.F. 1941. Contribution à l'ostéologie et à la systématique des rhacophorides africains. Revue de Zoologie et de Botanique Africaines. Tervuren 35:85–111.

Lawson, L. P., and L. Collett. 2011. Herpetofauna of Montane Areas of Tanzania. 1. Results from Two Amphibian Surveys of Malundwe Mountain, Mikumi National Park. Fieldiana Life and Earth Sciences 4:74–80.

Loader, S.P., F.S. Ceccarelli, M. Menegon, K.M. Howell, R. Kassahun, A.A. Mengistu, S.A. Saber, F. Gebresenbet, R. de Sa, T.R.B. Davenport, J.G. Larson, H. Müller, M. Wilkinson, and D.J. Gower. Persistence and stability of Eastern Afromontane forests: evidence from brevicipitid frogs. Journal of Biogeography 41:1781–1792.

Lutz, B. 1947. Trends towards non-aquatic and direct development in frogs. Copeia 1947:242–252.

Lutz, B. 1948. Ontogenetic evolution in frogs. Evolution 2:29–39.

Magnusson, W.E., and J.-M. Hero. 1991. Predation and the evolution of complex oviposition behavior in Amazonian rainforest frogs. Oecologica 86:10–18.

Meegaskumbura, M., G. Senevirathne, S.D. Biju, S. Garg, S. Meegaskumbura, R. Pethiyagoda, J. Hanken, and C.J. Schneider. 2015. Patterns of reproductive-mode

evolution in Old World tree frogs (Anura, Rhacophoridae). Zoologica Scripta 44:509–522.
Müller, H., H.C. Liedtke, M. Menegon, J. Beck, L. Ballesteros-Mejia, P. Nagel, and S.P. Loader. 2013. Forests as promoters of terrestrial life-history strategies in East African amphibians. Biology Letters 9:20121146.
Pagel, M., and A. Meade. 2006. Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. The American Naturalist 167:808–825.
Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. Systematic Biology 53:673–684.
Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analysis of phylogenetics and evolution in R language. Bioinformatics 20:289–290.
Pereira, E.B., R.G. Collevatti, M.N.C. Kokubum, N.E.O. Miranda, and N.M. Maciel. 2015. Ancestral reconstruction of reproductive traits shows no tendency toward terrestriality in leptodactyline frogs. BMC Evolutionary Biology 15:91.
Pfennig, D., and K. Pfennig. 2012. Evolution's Wedge. University of California Press, California, USA.
Poynton, J.C. 1964. Relationships between habitat and terrestrial breeding in amphibians. Evolution 18:131.
Prado, C.P.A., M. Uetanabaro, and C.F.B. Haddad. 2002. Description of a new reproductive mode in *Leptodactylus* (Anura, Leptodactylidae), with a review of the reproductive specialization towards terrestriality in the genus. Copeia 2002:1128–1133.
Pyron, R.A. 2014. Biogeographic analysis reveals ancient continental vicariance and recent oceanic dispersal in amphibians. Systematic Biology 63:779–797.
Pyron, R.A., and J.J. Wiens. 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.
Rambaut, A., and A.J. Drummond. 2009. Tracer v1.5.0 and higher. Available from: http://beast.bio.ed.ac.uk.
Revell, L.J. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.
Roberts, J.D. 1981. Terrestrial breeding in the Australian Leptodactylid frog *Myobatrachus gouldii* (Gray). Australian Wildlife Research 8:451–462.
Rödel, M.-O., M.F. Barej, A. Hillers, A.D. Leaché, N'G.G. Kouamé, C. Ofori-Boateng, N.E. Assemian, B. Tohé, J. Penner, M. Hirschfeld, J. Doumbia, L.N. Gonwouo, J. Nopper, C. Brede, R. Diaz, M.K. Fujita, M. Gil, G.H. Segniagbeto, R. Ernst, and L. Sandberger. 2012. The genus *Astylosternus* in the Upper Guinea rainforests, West Africa, with the description of a new species (Amphibia: Anura: Arthroleptidae). Zootaxa 3245:1–29.
Roelants, K., D.J. Gower, M. Wilkinson, S.P. Loader, S.D. Biju, K. Guillaume, L. Moriau, and F. Bossuyt. 2007. Global patterns of diversification in the history of modern amphibians. Proceeding of the National Academy of Sciences 104: 887–892.
Ronquist, F., and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Salthe, S.N., and W.E. Duellman. 1973. Quantitative constraints associated and reproductive mode in anurans. Pp. 229–249 in J.L. Vial (ed.), Evolutionary Biology of the Anurans. University of Missouri Press, Missouri, USA.
Samonds, K.E., L.R. Godfrey, J.R. Ali, S.M. Goodman, M. Vences, M.R. Sutherland, M.T. Irwin, and D.W. Krause. 2012 Spatial and temporal arrival patterns of Madagascar's vertebrate fauna explained by distance, ocean currents, and ancestor type. Proceedings of the National Academy of Sciences 109:5352–5357
Shen, X.X., D. Liang, Y.J. Feng, M.Y. Chen, and P. Zhang. 2013. A versatile and highly efficient toolkit including 102 nuclear markers for vertebrate phylogenomics, tested by resolving the higher level relationships of the Caudata. Molecular Biology and Evolution. 30: 2235–2248.
Schiøtz, A. 1999. Treefrogs of Africa. Edition Chimaira, Frankfurt, Germany.
Tihen, J.A. 1960. Comments on the origin of the amniote egg. Evolution, 14: 528–531.
Todd, B.D. 2007. Parasites lost? An overlooked hypothesis for the evolution of alternative reproductive strategies in amphibians. The American Naturalist 170:7793–799.
Tolley, K.A., T.M. Townsend, and M. Vences. 2013. Large-scale phylogeny of chameleons suggests African origins and Eocene diversification. Proceedings of the Royal Society B 280:20130184.
Wager, V.A. 1965. The Frogs of South Africa. Purnell & Sons (S.A.) Pty., Ltd., Cape Town, South Africa.
Wager, V.A. 1986. Frogs of South Africa, Their Fascinating Life Stories. Delta Books (Pty) Ltd., Johannesburg, South Africa.
Wells, K.D. 2007. The Ecology and Behavior of Amphibians. University of Chicago Press, Illinois, USA.
Zimkus, B.M., L. Lawson, S.P. Loader, and J. Hanken. 2012. Terrestrialization, miniaturization and rates of diversification in African Puddle Frogs (Anura: Phrynobatrachidae). PLoS ONE 7(4):e35118.
Zwickl, D.J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD thesis. Austin, The University of Texas at Austin.

Figure 1. Bayesian maximum clade credibility chronogram of the Afrobatrachia obtained using BEAST, illustrating divergence date estimates and the 95% highest posterior density region of dates (indicated by grey bars). Filled circles on nodes represent high support (PPB > 0.95); open circles indicate PPB < 0.95. Branches are color coded by the estimated molecular clock rate inferred from all five nuclear loci (*POMC*, *RAG-1*, *TYR*, *FICD*, *KIAA2013*): red, faster clock rate; blue, slower clock rate. Genera with > 15 species are highlighted. Nodes of interest are color coded: purple, Afrobatrachia; light blue, Arthroleptidae; green, Hyperoliidae; yellow, *Leptopelis*; orange, *Arthroleptis*; red, *Cardioglossa*; dark blue, *Afrixalus*; magenta, *Hyperolius*. The TMRCA marginal density proportion for each of these specific nodes, based on the combined results of all calibrations strategies, is illustrated below the chronogram, with distributions matching the same color scheme. Families are represented by the following: Hemisotidae (A); Brevicipitidae (B); Arthroleptidae (C: *Leptodactylodon*, D: *Nyctibates*, E: *Scotobleps*, F: *Astylosternus*, G: *Trichobatrachus*, H: *Leptopelis*, I–J: *Arthroleptis*, K–L: *Cardioglossa*); Hyperoliidae (M–N: *Kassina*, O: *Phlyctimantis*, P–Q: *Afrixalus*, R: *Cryptothylax*, S–W: *Hyperolius*). Photos A–Q, S–W by D.M. Portik, photo R by G. Jongsma.

Figure 1

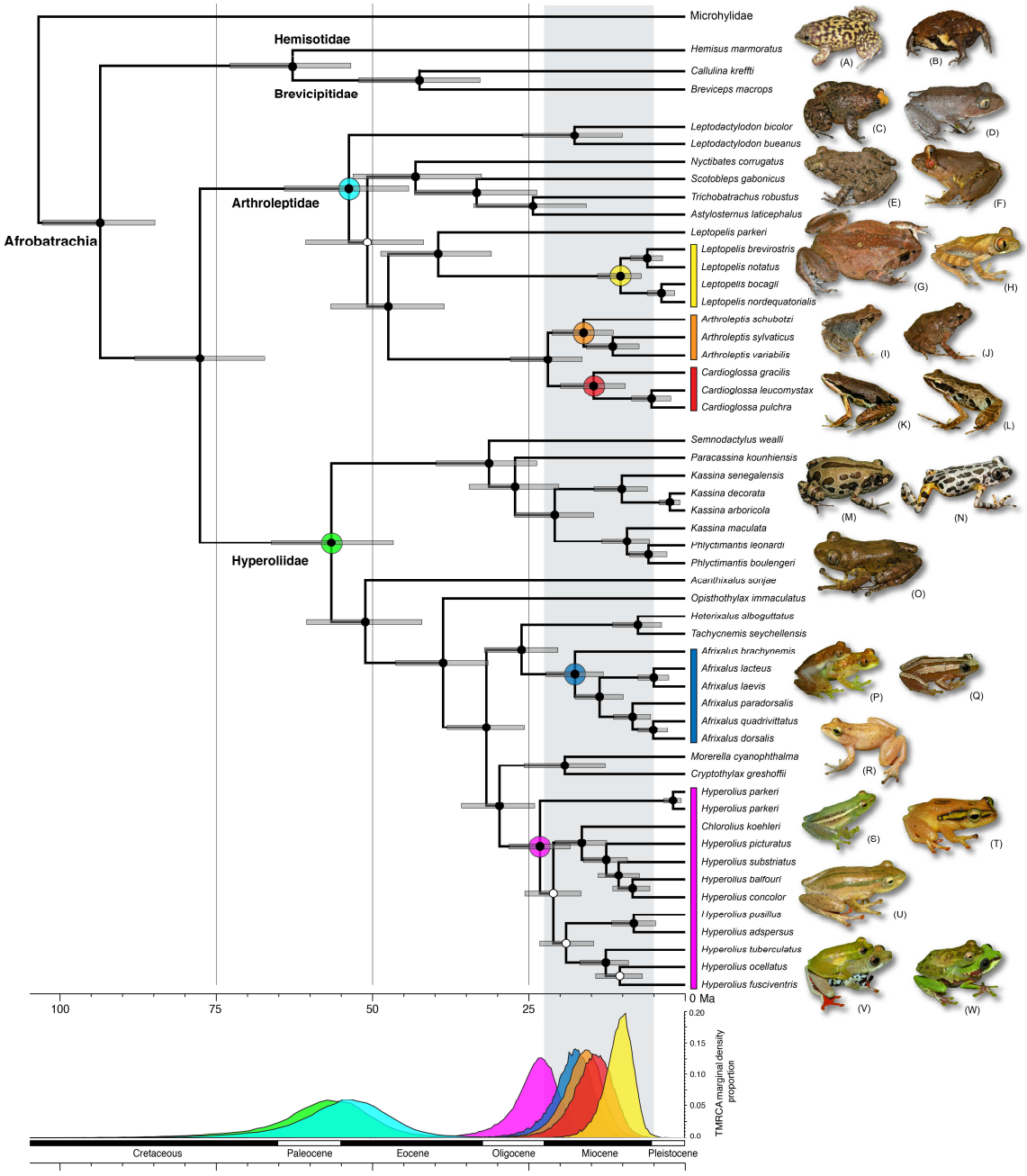
Microhylidae
Hemisotidae
Hemisus marmoratus
Callulina kreffti
Brevicipitidae
Breviceps macrops
Leptodactylodon bicolor
Leptodactylodon bueanus
Nyctibates corrugatus
Scotobleps gabonicus
Arthroleptidae
Trichobatrachus robustus
Astylosternus laticephalus
Afrobatrachia
Leptopelis parkeri
Leptopelis brevirostris
Leptopelis notatus
Leptopelis bocagii
Leptopelis nordequatorialis
Arthroleptis schubotzi
Arthroleptis sylvaticus
Arthroleptis variabilis
Cardioglossa gracilis
Cardioglossa leucomystax
Cardioglossa pulchra
Semnodactylus wealli
Paracassina kounhiensis
Kassina senegalensis
Kassina decorata
Kassina arboricola
Kassina maculata
Phlyctimantis leonardi
Hyperoliidae
Phlyctimantis boulengeri
Acanthixalus sonjae
Opisthothylax immaculatus
Heterixalus alboguttatus
Tachycnemis seychellensis
Afrixalus brachynemis
Afrixalus lacteus
Afrixalus laevis
Afrixalus paradorsalis
Afrixalus quadrivittatus
Afrixalus dorsalis
Morerella cyanophthalma
Cryptothylax greshoffii
Hyperolius parkeri
Hyperolius parkeri
Chlorolius koehleri
Hyperolius picturatus
Hyperolius substriatus
Hyperolius balfouri
Hyperolius concolor
Hyperolius pusillus
Hyperolius adspersus
Hyperolius tuberculatus
Hyperolius ocellatus
Hyperolius fusciventris
(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)
(J)
(K)
(L)
(M)
(N)
(O)
(P)
(Q)
(R)
(S)
(T)
(U)
(V)
(W)
100
75
50
25
0 Ma
0.20
0.15
0.10
0.05
0.0
TMRCA marginal density proportion
Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Pleistocene

Figure 2. Mapping of ecological, habitat, and reproductive character states onto the time-calibrated phylogeny of Afrobatrachians. Boxes represent characters used for ancestral state reconstructions, with pie charts at nodes representing posterior probabilities of character states. Circles represent characters used for correlated evolution analyses. Numbers above boxes or circles match to character legends.

Figure 2

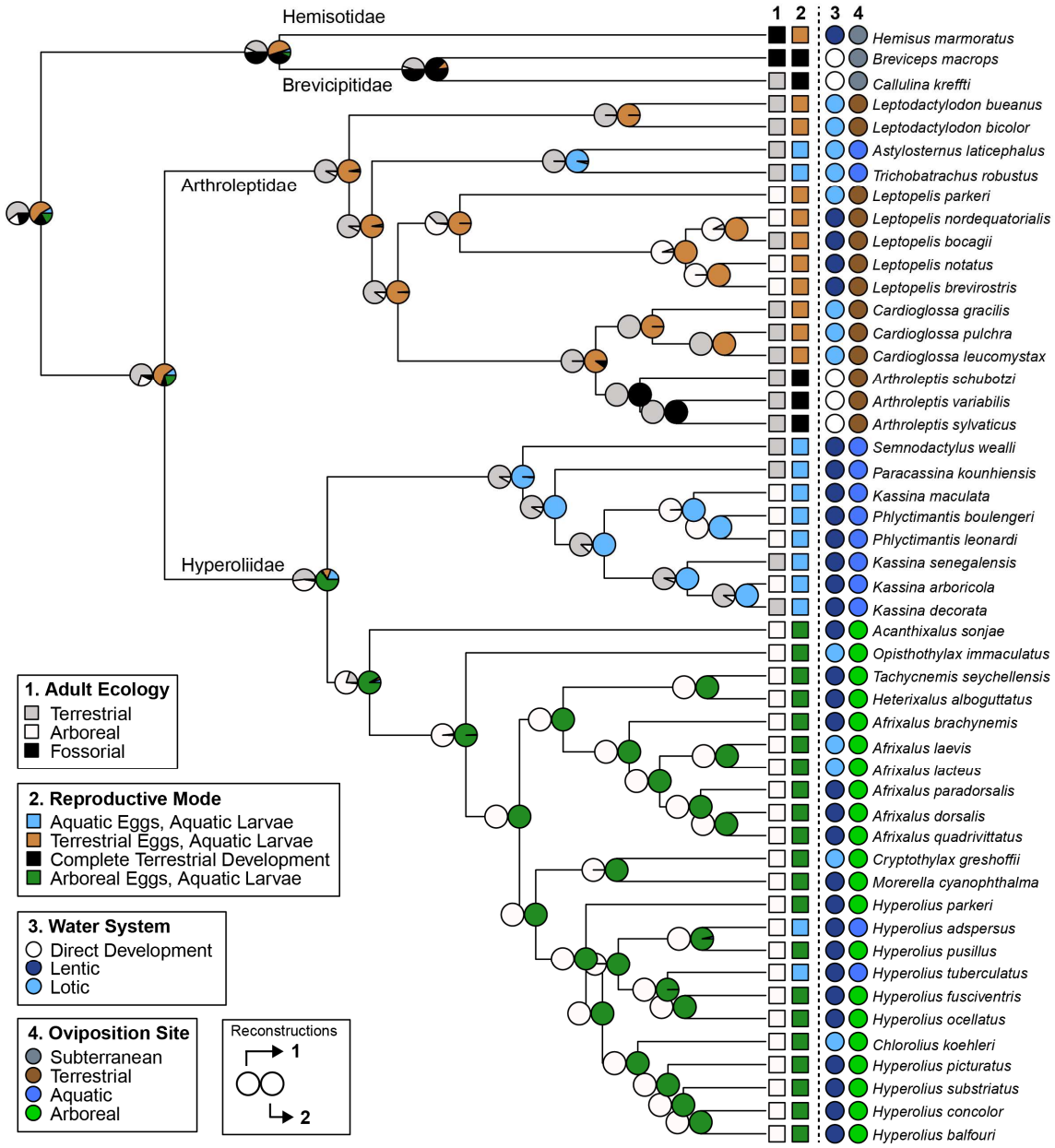

Figure 3. Estimated number of evolutionary changes between ecological (A) or reproductive modes (B) inferred from stochastic character mapping. Width of arrows is proportional to estimated number of changes. (C) The transitional pathway of correlated characters, arboreality, and arboreal oviposition. Rates and z-scores of key transition parameters are shown, and dashed arrows indicate unlikely pathways.

Figure 3

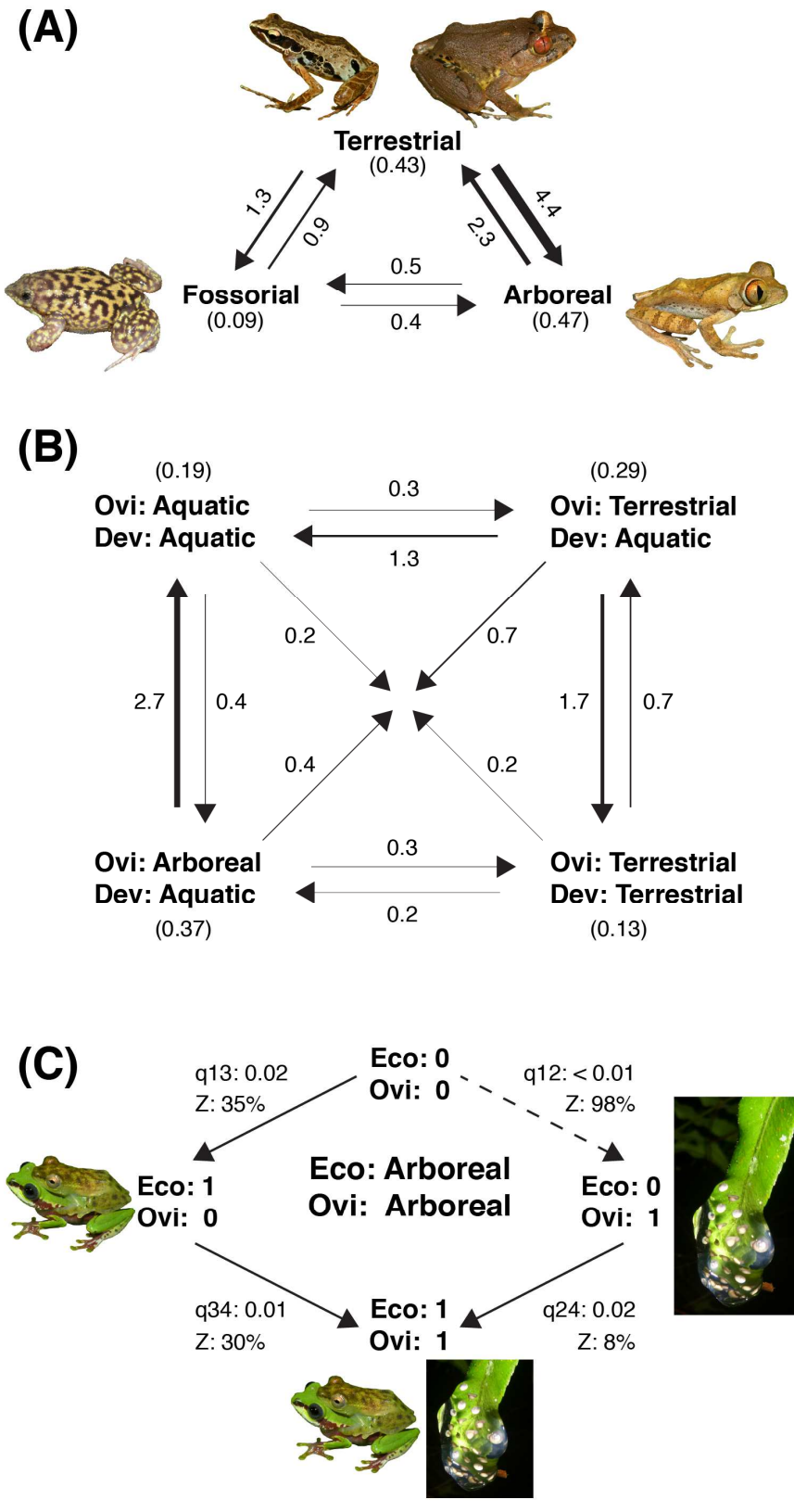
(A)
Terrestrial
(0.43)
1.3
0.9
2.3
4.4
0.5
0.4
Fossorial
(0.09)
Arboreal
(0.47)
(B)
(0.19)
Ovi: Aquatic
Dev: Aquatic
0.3
1.3
(0.29)
Ovi: Terrestrial
Dev: Aquatic
0.2
0.7
2.7
0.4
1.7
0.7
0.4
0.2
Ovi: Arboreal
Dev: Aquatic
(0.37)
0.3
0.2
Ovi: Terrestrial
Dev: Terrestrial
(0.13)
(C)
q13: 0.02
Z: 35%
Eco: 0
Ovi: 0
q12: < 0.01
Z: 98%
Eco: 1
Ovi: 0
Eco: Arboreal
Ovi: Arboreal
Eco: 0
Ovi: 1
q34: 0.01
Z: 30%
Eco: 1
Ovi: 1
q24: 0.02
Z: 8%

Table 1. A list of primers, primer sequences, and sources for genes sequenced in this study.

Table 1

Gene	Primer Name	Sequence (5' to 3')	Source
POMC	POMC-1	GAATGTATYAAAGMMTGCAAGATGGWCCT	Wiens et al. (2005)
	POMC-7	TGGCATTTTTGAAAAGAGTCAT	Smith et al. (2007)
RAG-1	RAG1 DCB1Fi	CTTCCGTGGAACAGGATATGA	Present study
	RAG1 DCB1R	CCAGATTCGTTGCCTTCACT	Present study
TYR	TyrC	GGCAGAGGAWCRTGCCAAGATGT	Bossuyt & Milinkovitch (2000)
	TyrG	TGCTGGCRTCTCTCCARTCCCA	Bossuyt & Milinkovitch (2000)
FICD	FICD F1	CCKCTNGTNGARGARATHGAYCA	Shen et al. (2013)
	FICD R1	TYTCNGTRCAYTTNGCDATRAA	Shen et al. (2013)
	FICD F2	AGGGTTTTCCCAGTCACGACTACTAYCAYCAYATHTAYCAYAC	Shen et al. (2013)
	FICD R2	AGATAACAATTTCACACAGGAARGGCCKVACRTCNCCYTCRTT	Shen et al. (2013)
KIAA2013	KIAA2013 F1	CTSAANTAYGCNGAYCAYTGYTT	Shen et al. (2013)
	KIAA2013 R1	CCNGGNCCRCARTAYTCRTTRTA	Shen et al. (2013)
	KIAA2013 F2	AGGGTTTTCCCAGTCACGACACYATGCAYGCNGAGAAYYTGTGG	Shen et al. (2013)
	KIAA2013 R2	AGATAACAATTTCACACAGGGANGCCACNCTRAACCARAA	Shen et al. (2013)

Table 2. Divergence dating estimates and 95% HPD intervals obtained for groups of interest.

Table 2

Analysis	Afrobatrachia	Hyperoliidae	Arthroleptidae	*Arthroleptis*	*Cardioglossa*	*Leptopelis*	*Hyperolius*	*Afrixalus*
Afr, (H+B), B	93.6 (84.8–102.8)	56.6 (47.0–66.6)	53.7 (43.9–63.9)	16.2 (11.6–21.2)	14.6 (9.6–19.9)	10.3 (7.0–14.0)	23.3 (18.3–28.3)	17.6 (13.2–22.4)
(H+B), B	103.8 (73.9–138.0)	62.8 (43.6–83.8)	59.0 (40.4–79.5)	17.7 (11.3–25.6)	15.8 (9.1–23.3)	11.3 (6.5–16.1)	25.8 (17.5–34.5)	19.5 (12.5–26.8)
(H+B)	100.9 (71.0–133.2)	60.6 (40.5–81.0)	57.5 (39.5–77.6)	17.2 (10.7–24.0)	15.4 (9.0–22.5)	11.1 (6.4–15.9)	24.9 (16.5–33.7)	18.8 (12.1–26.1)
B	96.1 (47.1–151.6)	58.0 (26.4–91.6)	55.0 (25.2–87.0)	16.3 (7.4–26.7)	14.7 (5.8–24.5)	10.5 (4.4–17.3)	23.8 (10.8–37.9)	18.1 (7.8–29.0)
Afr, B	92.7 (82.9–101.9)	56.3 (46.2–66.2)	53.3 (43.4–64.2)	16.0 (11.6–20.9)	14.4 (9.0–19.5)	10.2 (6.8–13.8)	23.0 (18.3–27.8)	17.5 (12.9–22.0)
Afr, (H+B)	93.3 (84.0–102.6)	56.5 (46.6–66.7)	53.2 (43.0–63.2)	16.0 (11.4–20.7)	14.3 (9.4–19.5)	10.2 (6.8–13.8)	23.2 (18.6–28.3)	17.6 (13.0–22.2)
Afr	91.9 (82.2–101.8)	55.1 (44.8–65.2)	52.5 (42.3–62.9)	15.7 (11.3–20.5)	14.0 (9.0–19.0)	10.1 (6.8–13.7)	22.7 (17.8–27.8)	17.1 (12.5–21.8)

Abbreviations for groups used in calibrations: Afr, tmrca of Afrobatrachia (Hemisotidae, Brevicipitidae, Arthroleptidae, Hyperoliidae); (H+B), tmrca of Hemisotidae and Brevicipitidae; B, tmrca of Brevicipitidae.

Table 3. Summary of transition model comparisons for given character sets.

Table 3

Character Set	States	Model Comparison	Likelihood ratio	d.f.	p-value
Adult Ecology	3	ER vs. ARD	9.84	5	0.08
		ER vs. SYM	5.26	2	0.07
		SYM vs. ARD	4.57	3	0.20
Reproductive Mode	4	ER vs. ARD	9.97	11	0.53
		ER vs. SYM	7.09	5	0.21
		SYM vs. ARD	2.87	6	0.82

The number of states per character set is given. Model comparisons, likelihood ratios, degrees of freedom, and significance are given for each character set. The difference in parameter numbers comprises the degrees of freedom, calculated as the parameter number of the more complex model minus number of parameters of the simpler model. More complex models were accepted only if they performed significantly better than the simpler model.

Table 4. The marginal probability of the root state for adult ecology and reproductive mode.

Table 4

Character Set	Character States	Scaled Root Likelihood
Adult Ecology	Terrestrial	**0.62**
	Arboreal	0.17
	Fossorial	0.21
Reproductive Mode	Aquatic/Aquatic	0.08
	Terrestrial/Aquatic	**0.57**
	Arboreal/Aquatic	0.18
	Terrestrial/Terrestrial	0.17

The probabilities are based on maximum likelihood analyses using ACE. The highest probability state is bolded.

Table 5. Results of character correlations based on maximum likelihood and Bayesian analyses.

Table 5

		Maximum likelihood					Bayesian Inference		
Character 1	Character 2	Dependent	Independent	Likelihood Ratio	d.f.	p-value	Dependent	Independent	logBF
Arboreal Ecology	Aquatic Oviposition	-33.93	-36.76	5.66	4	0.226	-46.13	-45.14	-1.98
	Arboreal Oviposition	-27.91	-34.97	14.12	4	**0.007	-40.25	-43.31	**6.12**
	Terrestrial Oviposition	-25.70	-29.03	6.66	4	0.155	-37.36	-39.20	3.68
Terrestrial Ecology	Aquatic Oviposition	-35.32	-37.30	3.96	4	0.411	-47.70	-46.42	-2.56
	Arboreal Oviposition	-29.80	-30.50	1.40	4	0.844	-42.22	-43.45	2.46
	Terrestrial Oviposition	-26.10	-27.30	2.40	4	0.663	-36.80	-38.18	2.76
Fossorial Ecology	Subterranean Oviposition	-3.29	-7.71	8.84	4	*0.065	-10.31	-15.59	**10.56**
Aquatic Oviposition	Lentic system	-32.31	-34.42	4.22	4	0.377	-42.67	-43.72	2.10
	Lotic system	-32.82	-34.39	3.14	4	0.535	-43.68	-42.97	-1.42
Arboreal Oviposition	Lentic system	-30.39	-32.63	4.48	4	0.345	-40.17	-40.59	0.84
	Lotic system	-30.78	-32.6	3.64	4	0.457	-40.76	-42.31	3.10
Terrestrial Oviposition	Lentic system	-25.29	-26.68	2.78	4	0.595	-35.35	-34.95	-0.80
	Lotic system	-23.97	-26.65	5.36	4	0.252	-35.00	-37.62	**5.24**
Arboreal Ecology	Lentic system	-36.66	-41.19	9.06	4	*0.060	-49.19	-48.94	-0.50
	Lotic system	-39.44	-41.16	3.44	4	0.487	-50.57	-48.82	-3.50
Terrestrial Ecology	Lentic system	-37.98	-41.73	7.50	4	0.112	-50.27	-49.87	-0.80
	Lotic system	-41.04	-43.69	5.30	4	0.258	-51.63	-52.16	1.06

Model values are log-likelihood values for ML analyses, Bayesian models are marginal likelihoods calculated by stepping stone sampling. Significant values, for which the dependent model of evolution is significantly better than the independent model, are expressed in bold (**p-value < 0.05, *p-value < 0.10; logBF > 5).

Table 6. Summary of transition parameters from correlated characters.

Table 6

		Arboreal Ecology, Arboreal Oviposition		Fossorial Ecology, Subterranean Oviposition	
Transition Parameter	Transition	Rate	Z-score	Rate	Z-score
q12	[0,0] to [0,1]	< 0.01	98%	0.02	14%
q13	[0,0] to [1,0]	0.02	35%	< 0.01	75%
q24	[0,1] to [1,1]	0.02	8%	2.40	5%
q34	[1,0] to [1,1]	0.01	30%	1.60	33%

To assess contingency, the rates and z-scores are given for key transition parameters. High z-scores (> 80%) and low rates suggest the evolutionary pathway of character transformations is unlikely. Character transitions are represented using 0 (absent) and 1 (present).

Appendix 1. List of species and associated catalogue numbers included in this study.

Family	Genus	Species	Museum Number
Arthroleptidae	*Arthroleptis*	*schubotzi*	CAS 250728
		sylvaticus	MH0314
		variabilis	CAS 207824
	Astylosternus	*laticephalus*	MVZ 244909
	Cardioglossa	*gracilis*	NCSM 78888
		pulchra	0951N
		leucomystax	MK067
	Leptodactylodon	*bicolor*	MCZ FS 34647
		bueanus	MCZ FS 34491
	Leptopelis	*bocagii*	CAS 250770
		brevirostris	CAS 207830
		nordequatorialis	MCZ FS 34688
		notatus	CAS 253554
		parkeri	CAS 168787
	Nyctibates	*corrugatus*	MCZ A 136788
	Scotobleps	*gabonicus*	MH0406
	Trichobatrachus	*robustus*	AMCC 117634
Brevicipitidae	*Breviceps*	*macrops*	CAS 193965
	Callulina	*kreffti*	CAS 168715
Hemisotidae	*Hemisus*	*marmoratus*	MVZ 244947
Hyperoliidae	*Acanthixalus*	*sonjae*	MORAS2
	Afrixalus	*brachynemis*	MVZ 265821
		dorsalis	CAS 253854
		lacteus	CAS 253843
		laevis	CAS 253803
		paradorsalis	CAS 253655
		quadrivittatus	MCZ FS 34546
	Chlorolius	*koehleri*	MHO 536
	Cryptothylax	*greshoffii*	MVZ 234714
	Heterixalus	*alboguttatus*	MVZ 241451
	Hyperolius	*adspersus*	CAS 254256
		balfouri	CAS 253643
		concolor	CAS 253869
		fusciventris	CAS 254006
		ocellatus	CAS 254075
		parkeri	FMNH 274372
		parkeri	MVZ 233914
		picturatus	ADL 3929
		pusillus	MVZ 233894
		substriatus	MVZ 265976

		tuberculatus	CAS 207717
	Kassina	*arboricola*	ADL 3887
		decorata	MCZ FS 34409
		maculata	SAIAB 88587
		senegalensis	MVZ 234142
	Morerella	*cyanophthalma*	Ba04.3
	Opisthothylax	*immaculatus*	MVZ 234815
	Paracassina	*kounhiensis*	TJC 869
	Phlyctimantis	*boulengeri*	ADL 3890
		leonardi	CAS 253978
	Semnodactylus	*wealli*	SANBI 4627
	Tachycnemis	*seychellensis*	
Microhylidae	*Phrynomantis*	*microps*	MVZ 249480

Appendix 2**.** List of ecological and reproductive characters associated with taxa.

Family	Taxon	Adult Ecology	Oviposition Site	Water System	Reproductive Mode
Brevicipitidae	*Breviceps macrops*	Fossorial	Subterranean	Direct Development	Terrestrial/Terrestrial
	Callulina kreffti	Terrestrial	Subterranean	Direct Development	Terrestrial/Terrestrial
Hemisotidae	*Hemisus marmoratus*	Fossorial	Subterranean	Lentic	Terrestrial/Aquatic
Arthroleptidae	*Arthroleptis schubotzi*	Terrestrial	Terrestrial	Direct Development	Terrestrial/Terrestrial
	Arthroleptis sylvaticus	Terrestrial	Terrestrial	Direct Development	Terrestrial/Terrestrial
	Arthroleptis variabilis	Terrestrial	Terrestrial	Direct Development	Terrestrial/Terrestrial
	Cardioglossa gracilis	Terrestrial	Terrestrial	Lotic	Terrestrial/Aquatic
	Cardioglossa pulchra	Terrestrial	Terrestrial	Lotic	Terrestrial/Aquatic
	Cardioglossa leucomystax	Terrestrial	Terrestrial	Lotic	Terrestrial/Aquatic
	Leptopelis bocagii	Terrestrial	Terrestrial	Lentic	Terrestrial/Aquatic
	Leptopelis brevirostris	Arboreal	Terrestrial	Lentic	Terrestrial/Aquatic
	Leptopelis nordequatorialis	Arboreal	Terrestrial	Lentic	Terrestrial/Aquatic
	Leptopelis notatus	Arboreal	Terrestrial	Lentic	Terrestrial/Aquatic
	Leptopelis parkeri	Arboreal	Terrestrial	Lotic	Terrestrial/Aquatic
	Leptodactylodon bicolor	Terrestrial	Terrestrial	Lotic	Terrestrial/Aquatic
	Leptodactylodon bueanus	Terrestrial	Terrestrial	Lotic	Terrestrial/Aquatic
	Trichobatrachus robustus	Terrestrial	Aquatic	Lotic	Aquatic/Aquatic
	Astylosternus laticephalus	Terrestrial	Aquatic	Lotic	Aquatic/Aquatic
Hyperoliidae	*Semnodactylus wealli*	Terrestrial	Aquatic	Lentic	Aquatic/Aquatic
	Paracassina kounhiensis	Terrestrial	Aquatic	Lentic	Aquatic/Aquatic
	Kassina arboricola	Arboreal	Aquatic	Lentic	Aquatic/Aquatic
	Kassina decorata	Terrestrial	Aquatic	Lentic	Aquatic/Aquatic
	Kassina maculata	Arboreal	Aquatic	Lentic	Aquatic/Aquatic
	Kassina senegalensis	Terrestrial	Aquatic	Lentic	Aquatic/Aquatic
	Phlyctimantis boulengeri	Arboreal	Aquatic	Lentic	Aquatic/Aquatic

Phlyctimantis leonardi	Arboreal	Aquatic	Lentic	Aquatic/Aquatic
Acanthixalus sonjae	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Opisthothylax immaculatus	Arboreal	Arboreal	Lotic	Arboreal/Aquatic
Morerella cyanophthalma	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Cryptothylax greshoffii	Arboreal	Arboreal	Lotic	Arboreal/Aquatic
Tachycnemis seychellensis	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Heterixalus alboguttatus	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Afrixalus laevis	Arboreal	Arboreal	Lotic	Arboreal/Aquatic
Afrixalus lacteus	Arboreal	Arboreal	Lotic	Arboreal/Aquatic
Afrixalus paradorsalis	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Afrixalus dorsalis	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Afrixalus quadrivittatus	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Afrixalus brachynemis	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Chlorolius koehleri	Arboreal	Arboreal	Lotic	Arboreal/Aquatic
Hyperolius adspersus	Arboreal	Aquatic	Lentic	Aquatic/Aquatic
Hyperolius balfouri	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Hyperolius concolor	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Hyperolius fusciventris	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Hyperolius ocellatus	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Hyperolius parkeri	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Hyperolius picturatus	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Hyperolius pusillus	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Hyperolius substriatus	Arboreal	Arboreal	Lentic	Arboreal/Aquatic
Hyperolius tuberculatus	Arboreal	Aquatic	Lentic	Aquatic/Aquatic

Chapter 2

An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura)

Abstract

Custom sequence capture experiments are becoming an efficient approach for gathering large sets of orthologous markers with targeted levels of informativeness in non-model organisms. Transcriptome-based exon capture utilizes transcript sequences to design capture probes, often with the aid of a reference genome to identify intron-exon boundaries and exclude shorter exons (< 200 bp). Here, we test an alternative approach that directly uses transcript sequences for probe design, which are often composed of multiple exons of varying lengths. Based on a selection of 1,260 orthologous transcripts, we conducted sequence captures across multiple phylogenetic scales for frogs, including species up to ~100 million years divergent from the focal group. After several conservative filtering steps, we recovered a large phylogenomic data set consisting of sequence alignments for 1,047 of the 1,260 transcriptome-based loci (~630,000 bp) and a large quantity of highly variable regions flanking the exons in transcripts (~70,000 bp). We recovered high numbers of both shorter (< 100 bp) and longer exons (> 200 bp), with no major reduction in coverage towards the ends of exons. We observed significant differences in the performance of blocking oligos for target enrichment and non-target depletion during captures, and observed differences in PCR duplication rates that can be attributed to the number of individuals pooled for capture reactions. We explicitly tested the effects of phylogenetic distance on capture sensitivity, specificity, and missing data, and provide a baseline estimate of expectations for these metrics based on nuclear pairwise differences among samples. We provide recommendations for transcriptome-based exon capture design based on our results, and describe multiple pipelines for data assembly and analysis.

Introduction

Using high throughput sequencing, there are now a variety of approaches available to generate large molecular data sets for the purpose of addressing population genetics or phylogenetics questions. A majority of these approaches fall in the category of reduced representation sequencing, in which orthologous sets of markers from a subset of the genome are obtained across taxa or individuals. A commonly used approach is RAD-seq, which targets anonymous loci adjacent to restriction enzyme sites (Miller et al. 2007), though the probability of obtaining orthologous sets of loci decreases as phylogenetic distance between samples increases (Rubin et al. 2012; Arnold et al. 2013). Other approaches include more targeted selection of loci using DNA or RNA probes, such as ultra-conserved element (UCE) sequencing (Faircloth et al. 2012) and anchored hybrid enrichment (Lemmon et al. 2012). Both approaches rely on short, highly conserved genomic regions for probe design and the subsequent capture of these targets for libraries with large insert sizes containing stretches of flanking sequences. This allows the use of the same set of markers across distantly related taxa, but the function of these loci is generally unknown, and the levels of variation in flanking regions are not predictable.

Other targeted sequence capture approaches allow more control over the level of variation of orthologous markers, including sequence capture using PCR-generated probes (SCPP) (Peñalba et al. 2014), and transcriptome-based exon capture (Bi et al. 2012). The latter approach uses transcriptome sequencing to identify protein-coding exons across populations or species, and is particularly useful for organisms for which no other genomic resources are readily available.

An important step before selecting markers derived from transcriptome sequences involves the identification of intron-exon boundaries to select longer exons, which requires the use of reference genomes (Bi et al. 2012; Bragg et al. 2015). Longer exons are preferred because they exceed the length of capture probes, allowing tiling, and for a given evolutionary rate they should have more informative sites compared to shorter exons. The transcriptome sequences recovered are typically composed of multiple exons, often short in length, making probe design challenging. The intron-exon identification step can be exceedingly difficult if the reference genome is too distantly related, and the direct use of transcriptome sequences for probe design is an alternative that has not been explored. Although this alternative approach ignores the potential presence of intron-exon boundaries, it offers an opportunity to capture exons of a variety of lengths along with their associated non-coding flanking regions. The length of probes, level of divergence between probes and targets, length distribution of genomic library fragments, and the number of and lengths of exons in the transcript sequences are all important factors that could determine the success of this alternative approach.

There are several major challenges for designing a custom sequence capture experiment for a non-model organism, particularly if the experiment involves species with relatively large genome sizes, spans multiple phylogenetic scales, and involves the *de novo* generation of genetic resources for probe design. In addition, wet-lab-specific decisions have the potential to significantly influence the outcome of sequence captures, including the number of genomic libraries to pool per capture reaction and the choice of genomic library blocking oligos. Few studies have focused on the exploration of these topics across a single experiment, yet the availability of baseline information can help inform these decisions and improve the success of sequence capture.

Across terrestrial vertebrates, amphibians exhibit the largest genome sizes. The average genome size of frogs is 5.0 gigabases (Gb) (max = 13.1 Gb, n = 497), whereas the salamander genome averages 34.5 Gb (max = 117.9 Gb, n = 426) (Gregory 2015). These genome sizes are larger than those of birds (1.3 Gb, n =896), mammals (3.1 Gb, n = 777), and squamates (2.1 Gb, n = 344) (Gregory 2015), and the performance of targeted exon capture for amphibians remains largely unexplored (but see Hedtke et al. 2013; McCartney-Melstad et al. 2015). Here, we examine the performance of transcriptome-based exon capture for frogs across multiple phylogenetic scales. The main focal group is the African frog family Hyperoliidae, consisting of 13 genera and 254 samples, which have an average genome size of 4.6 Gb (n = 11) (Gregory 2025). Our sampling also includes species from the sister family Arthroleptidae (7 genera, 7 samples), and a single representative from three more distantly related families (Brevicipitidae, Hemisotidae, Microhylidae). Pairwise comparisons within Hyperoliidae do not exceed 10% nuclear divergence, and the hyperoliid genera shared a common ancestor approximately 56 million years ago (Portik & Blackburn, in prep). The family Hyperoliidae shares a common ancestor with Arthroleptidae approximately 77 Ma, with Hemisotidae and

Brevicipitidae approximately 93 Ma, and with Microhylidae approximately 103 Ma, and uncorrected pairwise nuclear differences between hyperoliids and the outgroups approaches 20%.

We describe our methodological approach for generating and mining transcriptome resources, the selection of orthologous markers and probe design, choice of blocking oligos in capture reactions, the pipeline for assembling and processing capture sequence data, and the overall results of our exon capture experiment. Given the tremendous level of divergence between our focal group and available frog reference genomes (*Xenopus laevis* and *X. tropicalis*, minimum 150 million years divergent), we did not attempt to identify intron-exon boundaries to select longer exons. Rather, we use transcriptome sequences directly for probe design. We evaluate our results given this approach, including characterizing the number of exons in transcript sequences, the lengths of these exons, our ability to recover exons and their flanking regions, and the effects of exon length on sequencing depth. The level of variation within the family Hyperoliidae and the inclusion of highly divergent outgroup taxa allows us to examine the effects of phylogenetic divergence on exon capture performance. Specifically, we examine the relationship between phylogenetic distance on exon capture sensitivity, specificity, and the proportion of missing data in the final sequence alignments. We also examine the effects of library pool size during multiplexed captures on raw data yield, sequencing depth, and read duplication levels.

Methods

Transcriptome Sequencing and Analysis

Four species of hyperoliids representing multiple divergent clades were chosen for transcriptome sequencing: *Afrixalus paradorsalis* (CAS 255487), *Hyperolius balfouri* (CAS 253644), *Hyperolius riggenbachi* (CAS 253658), and *Kassina decorata* (CAS 253990). Whole RNA from a portion of liver sample preserved in RNA Later was extracted using the RNeasy Protect Mini Kit (Qiagen). Samples were evaluated using a BioAnalyzer 2100 RNA Pico chip (Agilent), with RIN scores of 7.0, 7.0, 7.4, and 5.5, respectively. Sequencing libraries were prepared using half reactions of the TruSeq RNA Library Preparation Kit V2 (Illumina), beginning with Poly-A selection for samples with high RIN scores (> 7.0) and Ribo-Zero Magnetic Gold (Epicentre) ribosomal RNA removal for samples with low RIN scores (< 7.0). Libraries were pooled and sequenced on an Illumina HiSeq2500 with 100 bp paired-end reads. Transcriptomic data were cleaned following Singhal (2013). Cleaned data were assembled using TRINITY (Grabherr et al. 2011) and annotated with *Xenopus tropicalis* (Ensembl) as a reference genome using reciprocal BLASTX (Altschul et al. 1997) and EXONERATE (Slater & Birney 2005). We then compared annotated transcripts from the four species to search for orthologs via BLAST (Altschul et al. 1990). We removed mitochondrial loci from the transcripts. We only kept transcripts with a GC between 40%-70% because extreme GC content causes a reduced capture efficiency for the targets (Bi et al. 2012). Orthologous transcripts with a minimum length of 500 base pairs (bp) were identified across all four samples, resulting in the identification of 2,444 shared transcripts. Transcripts exceeding 850 bp were arbitrarily trimmed to this length for probe design, reflecting a trade-off decision between locus length and the total number of loci included in the experiment. The average

pairwise divergence across transcripts among all four samples ranged from 1.4% to 25.9%.

Availability of Transcriptome Tools. All the bioinformatics pipelines for transcriptome data processing and annotation are available at https://github.com/CGRL-QB3-UCBerkeley/DenovoTranscriptome. Pipelines for marker development are available at https://github.com/CGRL-QB3-UCBerkeley/MarkerDevelopmentPylogenomics.

Sequence-Capture Probe Design

The orthologous transcripts were subjected to additional filtering steps before a final gene set was chosen. The initial filtering step applied upper and lower limits on average transcript divergence, eliminating loci with low variation (< 5.0% average divergence) and exceptionally high variation (> 15.0% average divergence), resulting in the removal of 266 genes. The remaining 2,178 genes were examined for repetitive elements, short repeats, and low complexity regions, which are problematic for probe design and capture. The four sets of transcripts per gene (totaling 8,712 sequences) were screened using the REPEATMASKER Web Server (Smit et al. 2015). This step resulted in the masking of repetitive elements or low complexity regions in 929 sequences, with 7,783 sequences passing the filters. To be conservative, if any of the four transcripts for a gene contained masked sites, that gene was removed from the final marker set, which resulted in the removal of an additional 468 markers. From this reduced set of 1,710 markers, 400 markers with the highest divergence were selected (average divergence ranging from 10.4% to 14.9%) followed by 860 randomly drawn markers from the remaining subset. This marker set was supplemented with five positive controls, which consisted of nuclear sequence data generated using Sanger sequencing for five loci: *POMC* (624 bp), *RAG-1* (777 bp), *TYR* (573 bp), *FICD* (524 bp), and *KIAA2013* (540 bp). The final marker set selected for probe design included 1,265 genes from four species and 5,060 individual sequences.

The final filtered gene set was used to design a MYaits-3 custom bait library (MYcroarray), which consists of 60,060 unique probes per reaction and a total of 48 capture reactions. Following the manufacturers recommendation for capturing sequences of species greater than 5% divergent, 120mer baits were selected, rather than 100mer or 80mer baits. For each locus, we included a sequence from each of the four species; the 5,060 sequences included for probe design totaled 3,983,022 bp, which is approximately 995,700 bp for each full set of loci per species. Following a 2x tiling scheme (every 60 bp) resulted in 60,179 unique baits, therefore 119 probes were randomly dropped to achieve the probe limit.

Genomic Library Preparation and Pooling

Genomic DNA was extracted from 264 samples (254 ingroup samples, 10 outgroups) using a high-salt extraction method (modified from Aljanabi and Martinez 1997). The DNA was quantified by Qubit DNA BR assay (Life Technologies) and 1700 ng total DNA was diluted in 110 µl of ultrapure H_2O. A Bioruptor UCD-200 (Diagenode) was used to sonicate the samples on a low setting for 15 minutes, using 30s on/30s off cycling. For each sonicated sample, 4.5 µl of product was run on a 1% gel at 135V for 35 min to ensure fragments were appropriately sized (100–500 bp, average 200–300 bp).

Individual genomic libraries were prepared following Meyer and Kircher (2010), with slight modifications, including the use of at least 1600 ng total DNA for library preparation (rather than 500 ng) to remedy the possibility of decreased library diversity resulting from the larger genome size of frogs. We used 7 cycles of post-adapter ligation PCR to enrich the libraries and incorporate a 7bp P7 index, allowing the combination of up to 96 samples in the same sequencing lane. The resulting 50 µl of amplified library product had an average concentration of 35 ng/µl measured by a Nanodrop 1000 spectrophotometer (Thermo Scientific), producing an average yield of 1,750 ng total library DNA.

Samples were pooled for capture reactions according to phylogenetic relatedness as determined by 16S mtDNA data (Portik, unpublished data). Typical pools contained 5–6 genomic libraries, but ranged from 1–7 libraries. All pools contained 1500 ng of total starting DNA, divided equally among the samples included in the pool.

Sequence Capture Reactions

MYbaits capture reactions were performed following the v2.3.1 manual with some modifications. For each capture reaction library master mix, the pooled libraries were vacuum dried at 45°C for 70 min and re-suspended in ultrapure H_2O, then combined with 1.66 µl each of human, mouse, and chicken COT-1, and choice of blocking oligos. The combined volume of water for DNA resuspension and volume of blocking oligonucleotides totaled 6.5 µl. An initial three capture reactions were performed on the same library pool to assess the performance of three different types of oligonucleotide blockers designed to anneal to the library adapters during hybridization and prevent daisy-chaining. These blockers consist of the universal blocking oligos (included with the MYbaits kit) which use inosine to block the 7bp index sequence, short blocking oligos which leave the index sequence unblocked, and xGEN blocking oligos (Integrated DNA Technologies), which use proprietary modifications to block the index. Their performance was compared using qPCR analysis of amplified post-capture products, examining enrichment of positive controls and depletion of negative controls. The xGEN blocking oligos performed significantly better in these tests (see Results); we assumed this assessment was a good proxy for sequencing results and these blocking oligos were used for all subsequent capture reactions.

Beyond the slight modifications to the hybridization reaction components discussed above, we followed the manufacturer's protocol as written, and the hybridization reaction proceeded at 65°C for 27 hours. Individual capture reactions were purified using streptavidin-coated magnetic beads and post-capture products were PCR amplified using four independent reactions of 14 cycles each. These reactions were resuspended in 12 µl of ultrapure H_2O, and had an average concentration of 15 ng/µl, as measured by Nanodrop. Purified PCR products from the same capture were combined and quantified using a BioAnalyzer 2100 DNA 1000 chip. The combined post-capture amplified products were on average 3.7 ng/µl (range of averages: 1.1–7.2 ng/µl) and the average product size was 398 bp (range of averages: 361–466 bp). Results from Qubit assay were similar, with an average concentration of 4.5 ng/µl (range: 1.0–7.7 ng/µl) for combined post-capture amplified products. The combined post-capture libraries were grouped into three sets (totaling 74, 91, and 92 libraries), pooled in equimolar amounts, and sequenced on three lanes of an Illumina HiSeq2500 with 100 bp paired-end reads.

Sequence Capture Data Processing

Raw sequence data were cleaned following Singhal (2013) and Bi et al. (2012). In brief, raw fastq reads were filtered using TRIMMOMATIC (Bolger et al. 2014) and CUTADAPT (Martin 2011) to trim adapter contaminations and low quality reads. BOWTIE2 (Langmead & Salzberg 2012) was used to align the data to *Escherichia coli* (NCBI: 48994873) to remove potential bacteria contamination. We eliminated exact duplicates as well as low complexity sequences using a custom script. Overlapping paired reads were also merged using FLASH (Magoč & Salzberg) and COPE (Liu et al. 2012) to avoid inflated coverage estimate in the overlapping region. The resulting cleaned reads of each individual specimen were *de novo* assembled using ABYSS (Simpson et al. 2009). We first generated individual raw assemblies using a wide range k-mers (21, 31, 41, 51, 61 and 71) and then used CD-HIT-EST (Li & Godzik 2006), BLAT (Kent 2002), and CAP3 (Huang & Madan 1999) to cluster and merge all raw assemblies into final, less-redundant assemblies. We used BLASTN (evalue cutoff = 1e-10, similarity cutoff = 70) to compare the 5,060 target sequences with the raw assemblies of each individual in order to identify the set of contigs that were associated with targets (in-target assemblies). We also ran a self-BLASTN (evalue cutoff =1e-20) to compare the assemblies against themselves to mask any regions from a contig that matched other regions from other contigs. For each matched contig we used EXONERATE (Slater & Birney 2005) to define protein-coding and flanking regions. We retained flanking sequences if they were within 500 bp of a coding region. Finally, all discrete contigs that were derived from the same reference transcript were joined together with Ns based on their relative BLAST hit positions to the reference. Most of the final in-target assemblies contain multiple contigs, and each includes both coding regions and flanking sequences if captured.

Cleaned sequence data were then aligned to the resulting individual-specific in-target assemblies using NOVOALIGN (Li & Durbin 2009) and we only retained reads that mapped uniquely to the reference. We used Picard (http://broadinstitute.github.io/picard/) and GATK (McKenna et al. 2010) to perform re-alignment. We finally used SAMTOOLS/BCFTOOLS (Li et al. 2009) to generate individual consensus sequences by calling genotypes and incorporate ambiguous sites in the in-target assemblies. We kept a consensus base only when the site depth is above 5X. We masked sites within 5 bp window around an indel. We also filtered out sites where more than two alleles were called. We converted FASTQ to FASTA using seqtk (https://github.com/lh3/seqtk) and masked putative repetitive elements and short repeats using REPEATMASKER (Smit et al. 2015) with vertebrata metazoa as a database. We removed markers if more than 80% of the bases were Ns. We then calculated read depth of each individual marker and filtered out loci if the depth fell outside 1st and 99th percentile of the statistics. We also eliminated markers if the individual heteozygosity fell outside the 99th percentile of the statistics. The final filtered assemblies of each individual specimen were aligned using MAFFT (KAtoh & Standley 2013). Alignments were then trimmed using TRIMAL (Capella-Gutierrez et al. 2009). We removed alignments if more than 30% missing data (Ns) are present in 30% of the samples. We also removed alignments if the proportion of shared polymorphic sites in any locus is greater than 20%.

Availability of Sequence Capture Data Tools. The bioinformatic pipelines of sequence capture data processing are available at https://github.com/CGRL-QB3-UCBerkeley/denovoTargetCapturePhylogenomics.

Sequence Capture Efficiency Evaluation

Sequencing Depth, Duplication Levels, and Pooling Sizes. To evaluate capture efficiency, average per-base sequence depth, or coverage, was calculated separately for the exon sequences and for the flanking sequences of each sample. The coverage at each base pair site for either data set was inferred using the SAMTOOLS (Li et al. 2009). The per base pair coverage estimates for all sequences (exon or flanking) associated with each transcript (up to 1,260 genes) were averaged, resulting in a set of average coverage estimates across loci. The resulting output of the set of average coverage estimates was used to infer the median, upper and lower quartiles, and range of coverage estimates using samples or genes as factors. These calculations were performed and automated across samples using python scripts and the output was visualized in R. Differences in the levels of coverage were examined using pooling size as a factor. To control for differences in coverage possibly resulting from phylogenetic distance, comparisons were only made among pools of the ingroup genus *Hyperolius* (160 samples, 28 captures).

Duplication refers to the number of non-unique sequencing reads, which were eliminated from our sequence capture data processing pipeline. The level of duplication among reads, expressed as a percentage, was estimated by dividing the number of duplicate reads by the total number of raw reads. Differences in the levels of duplication were examined using pooling size as a factor, compared across the genus *Hyperolius*. The amount of raw data (in bases) was also compared across pool sizes using the genus *Hyperolius*.

Sequence Capture Sensitivity. Sensitivity refers to the percentage of bases of target sequences that are covered by at least one read, and here the target refers to the exons of each gene. To calculate this metric, the final in-target assemblies (including exons and flanking sequences) of each sample were compared to a set of transcript sequences used for probe design, from only one of the design species, using BLASTN with a evalue cutoff of 1e-10. This was automated using custom scripts to produce output files of all blast hits for each sample. For each output file, any overlapping base pair coordinates for blast hits within a locus were merged. Following the merging of coordinates, the number of base pairs for all exon blast hits per locus was totaled, and was divided by the total length of the transcript sequence to calculate the sensitivity per transcript. The total number of base pairs from all exon blast hits was divided by the total number of base pairs of all the transcript sequences, producing an overall estimate of sensitivity per sample.

Sequence Capture Specificity. Specificity is a metric that measures how many base pairs of cleaned reads are aligned to target sequences, expressed as a percentage. In this experiment, the target sequences are represented in two ways: in-target assemblies (exons and their associated flanking sequences), and exons only. For each sample, bam files were converted to sam file format using the SAMTOOLS view function and the total number of base pairs aligned within the exon sequences and flanking sequences were counted by parsing the bam files. To estimate base pairs aligned with transcript exon

sequences only, the sample bam file was converted to sam format using the associated bed file containing base pair coordinates for exons only, and total aligned base pairs were calculated in the same manner. The number of cleaned read base pairs was calculated from the summing the read lengths contained within cleaned reads files.

Exon Coverage Uniformity. The uniformity of coverage across the length of exons was examined using both longer (> 200 bp) and shorter (61–100 bp) exons. Exons matching these criteria were filtered out from bed files containing exon coordinates independently for each sample. For longer exons, five bins of 10 bp increments were created for both the 5' and 3' ends, resulting in the generation of ten additional bed files per sample. For shorter exons, three bins of 10 bp increments were created for both the 5' and 3' ends, resulting in the generation of six additional bed files per sample. Each bed file was used to calculate the per base pair coverage for a specific end bin using SAMTOOLS. These per base pair coverage values were averaged within exons, and all averages of exons for a particular bin were subsequently combined across 50 randomly chosen samples. The values across bins were visualized in R to assess the median, upper and lower quartiles, and range of coverage estimates.

Effects of Phylogenetic Distance. We sought to test the relationship between phylogenetic distances and several evaluation metrics to determine if Sanger sequence data have predictive power for exon capture success. Phylogenetic distance was calculated as the average of uncorrected pairwise differences between samples and the four design species. These divergence estimates were calculated using the five positive controls (nuclear loci from Sanger sequencing). As this information would generally be available to researchers before designing such an experiment, these loci provide an a priori estimation of divergence across the focal group. The effects of phylogenetic distances on capture specificity, sensitivity, and duplication were investigated using simple linear regressions. The values for the above metrics were averaged for each genus, providing values for a total of 23 genera for comparison. Average phylogenetic distances ranged from 6.7–18.3%, representing divergences up to 103 million years old (Portik & Blackburn, in prep).

Evaluation of Exon Phylogenetic Informativeness. The resulting alignments of exon-only data and flanking region data were evaluated for taxon number, sequence length, percentage of missing data, and proportion of informative sites. These results were visualized in R, and the relationship between the number of informative sites and alignment length was investigated using a simple linear regression. The relationship between phylogenetic distance and missing data was also investigated using a simple linear regression. The percentage of missing data was calculated from the final concatenated alignment of exon-only loci that passed multiple post-processing filters, including a minimum length of 90 bp, no more than 80% missing data per sequence in alignments, and no more than 30% total missing data across an alignment. These filters were enforced using a custom alignment refinement python script for all alignments.

Availability of bioinformatics tools. All custom python scripts for sequence capture performance evaluation are available on github (https://github.com/dportik/). These

include tools for automating the calculation of coverage, duplication, sensitivity, specificity, and coverage uniformity. Additional scripts are available for evaluating and refining DNA sequence alignments.

Results

Effects of Blocking oligos

Quantitiative PCR reactions were performed for a positive control nuclear locus (KIAA2013) targeted by the hybridization probes and a negative control nuclear locus (49065) not targeted by the capture probes. All reactions were standardized for the same input amount of DNA (4ng). For the positive control, all post-capture curves show an expected leftward shift relative to the pre-capture, indicating that the concentration of copies of the KIAA2013 locus has increased significantly in the post-capture library pools. Of the three blocker types, the greatest change in enrichment is observed with the post-capture pool using xGen blocking oligos (11.9 cycle shift), rather than the universal blocking oligos or short blocking oligos (10.3 cycle shifts) (Fig. 4). For the negative control, all post-capture curves show an expected rightward shift relative to the pre-capture, indicating that the concentration of copies of the 49065 locus has decreased in the post-capture library pools. However, the universal blocking oligos and short blocking oligos show only minor differences from the pre-capture library (1.9 and 1.0 cycle shifts respectively) (Fig. 4). In contrast, the post-capture pool using xGen blocking oligos has shifted considerably (10.3 cycle shift), indicating that non-target regions have been significantly depleted from the library pool (Fig. 4). This is reflected in the post-capture library quantification, in which higher amounts of DNA were detected in the universal blocker reaction (23 ng/μL) and short blocker reaction (23.7 ng/μL), compared to the xGen reaction (14 ng/μL), indicating that more non-targeted sequences were accidentally captured during the hybridization.

Sequence Capture Data

The average number of reads sequenced across the 264 samples is 2,422,484 (range: 415,439–6,899,259 reads), and as we sequenced 100 bp paired-end reads, the average total base pair yield is 484.4 Mb (range: 83.0–1,379.8 Mb). In addition to the removal of low complexity and low quality reads, the raw reads were filtered to remove exact duplicates, adapters, and bacteria contamination. After these filtration steps, the average number of base pairs of cleaned reads was 331.9 Mb (range: 65.3–789.6 Mb); on average 68% of the raw base pairs passed the quality control filters.

The assemblies were assigned to targeted transcripts, and the resulting in-target assemblies contained a mix of exon sequences and non-coding flanking sequences (Fig. 5A). The length of the in-target assemblies were often several thousand base pairs, much longer than the original targeted transcript sequences (which were maximally 850 bp), illustrating a significant amount of non-coding flanking sequence data associated with each exon was captured (Fig. 5A). By trimming the flanking sequences, the concatenated exons closely match the original transcript sequence lengths (Fig. 5B). Across all targeted loci and samples, the median number of exons per transcript is four, but ranged from a single exon to eleven exons per transcript (Fig. 6). The average length of exons within transcripts recovered is 153 bp, but the data set revealed a wide range in sizes,

from shorter exons (< 100 bp) to longer exons (> 600 bp) that cover the entire transcript sequence used (Fig. 7).

Sequencing Depth and Duplication Levels

The sequencing depths of merged contigs showed variation between loci and across samples, but the most pronounced differences in coverage occurred between the exon and flanking regions (Fig. 5A). The average sequencing depth across all exons for all samples averaged 142.4X (n = 1,372,603 exons), whereas the flanking regions averaged 45.5X. This result is consistent with expectations for transcriptome-based exon capture, as the probe design only considered exon regions. Despite not specifically targeting these adjacent non-coding regions, this experiment clearly demonstrates non-coding regions can be captured and sequenced with sufficiently high coverage. Because the estimates of sequencing depth only consider sites that are captured, relating coverage to phylogenetic distance is not a meaningful metric. We did consider the effect of pooling size on coverage, but within a single genus that was the main focus of the experiment and for which capture results were very similar (genus *Hyperolius*). A comparison of pool sizes (1–2, 4–7) revealed no significant differences in sequencing depth across all loci based on the student's t-test (Fig. 8). Similarly, there do not appear to be to be differences in raw data yield (in total base pairs) for different pool sizes (Fig. 8), though low sample sizes in smaller pools preclude rigorous testing.

The duplication levels among reads are an indicator of the diversity of sequences captured, with high duplication implying a less diverse post-capture library relative to post-capture libraries with lower duplication levels. In general, a higher number of post-capture PCR cycles are expected to produce higher levels of duplication among samples. In this experiment, all post-capture PCR reactions used the same number of cycles; therefore, our comparison of duplication levels is an indicator of post-capture sequence diversity rather than a methodological artifact. Levels of duplication were similar between the ingroup (average: 17.2%, range: 3.4%–37.9%) and outgroups (average: 16.5%, range: 5.0%–24.8%). We tested for a relationship between duplication level and phylogenetic distance using a simple linear regression, and found the regression was not significant ($F(1, 21) = 0.79$, $p = 0.38$). Though phylogenetic relatedness may not be a predictor of duplication levels, there is a clear pattern of differences in duplication levels across pooling sizes (Fig. 8). Pools with a single individual or two individuals have much lower duplication levels (3.9% and 5.1%, respectively) than pools with four or more individuals (range of averages: 15.7%–19.0%) (Fig. 8). Small samples sizes precluded statistical testing for these differences between smaller and larger pools, but these data indicate pooling size is much more likely to affect duplication levels than other factors such as phylogenetic distance. We did not find significant differences in duplication levels between larger pools, and this strongly suggests pooling seven individuals did not negatively impact the resulting diversity of sequences captured among samples. Additional replicate captures of larger pool sizes can help determine at which point captured sequence diversity is impacted and establish limitations in pooling sizes for successful capture.

Exon Coverage Uniformity

Using 50 random ingroup samples, sequencing depth values were calculated for the edges of exons in 10 base pair bins, with 5 bins included for longer exons (> 200 bp) and 3 bins included for shorter exons (61–100 bp). At the 5' and 3' ends of longer exons, the average coverage is 117.2X and 124.0X, respectively (Fig. 9, Table 7). These values increase slightly across bins towards the center, with both the 5' and 3' 41–50 bp bins having approximately 165X coverage (Fig. 9, Table 7). The coverage values for edge bins of shorter exons were lower, but in general showed the same trend increasing towards the center (Fig. 10, Table 8). Here, the average coverage of the 5' and 3' ends is 74.9X and 80.4X, respectively, with the most central bins (21–30 bp) exhibiting 83.7X and 86.2X coverage (Fig. 10, Table 8). Together, these results indicate high uniformity in sequencing depth across the length of short exons, and demonstrate only a slight decrease in coverage towards the edges of longer exons.

Sensitivity, Specificity, and the Effects of Phylogenetic Divergence

We explored capture sensitivity, the percentage of bases of in-target assemblies that are covered by at least one read, across all samples in our experiment. In general, sensitivity varied between genera but was relatively consistent within genera (Fig. 11A). The average across all ingroup samples is 80.1% (range 52.1%–91.8%), whereas outgroup samples averaged 33.8% (range 20.7%–42.2%). A simple linear regression was calculated to predict sensitivity (%) based on phylogenetic distance. A significant regression equation was found ($F(1, 21) = 79.58$, $p < 0.001$), with an adjusted R^2 of 0.78 (Fig. 12). Sensitivity is equal to [108.19 + -4.57*(average pairwise divergence)] percent when pairwise divergence is measured as a percent; sensitivity decreased 4.57% for each percent increase of pairwise divergence.

Specificity is a metric similar to sensitivity, but it measures the percentage of base pairs of cleaned reads that can be aligned to target sequences. We investigated specificity using the in-target assemblies (exons and flanking regions) and exons only. Specificity varied among genera (Fig. 11B), and across all ingroup samples averaged 60.2% (range 32.0%–73.0%), whereas outgroup samples averaged 35.6% (range 15.0%–50.0%). As expected, specificity of the exon only data set was lower, and ingroup genera exhibited higher specificity (47.3%, range: 26.0%–65.0%) than outgroup genera (27.7%, range: 13.0%–40.0%). Using specificity results from the in-target assemblies, a simple linear regression was calculated to predict specificity (%) based on phylogenetic distance. A significant regression equation was found ($F(1, 21) = 44.1$, $p < 0.001$), with an adjusted R^2 of 0.66 (Fig. 12). Specificity is equal to 83.99 + -3.26*(average pairwise divergence) percent when pairwise divergence is measured as a percent. Specificity decreased 3.26% for each percent increase of pairwise divergence.

Sequence Alignments and Informativeness

There were a total of 1,047 exon-only alignments and 287 flanking region alignments that passed all filtering criteria. Together, the concatenated alignment of flanking and exon data totals 631,127 base pairs.

For exon-only alignments, the average number of taxa included is 250, average per locus length is 536 bp, average level of missing data is 8.9%, and the average proportion of informative sites is 38.3%. The concatenated alignment of the exon-only

loci totals 561,180 base pairs. The average proportion of missing data in the concatenated alignment for the ingroup samples is 10.7% (range 3%– 35%), and is 55.1% (range 43%– 74%) for the outgroup samples. A simple linear regression was calculated to predict missing data levels in the final exon-only alignments, based on phylogenetic distance. A significant regression equation was found ($F(1, 21) = 96.78$, $p < 0.001$), with an adjusted R^2 of 0.81 (Fig. 12). Missing data is equal to [-22.07 + 4.76*(average pairwise divergence)] percent when pairwise divergence is measured as a percent. Missing data increased 4.76 percent for each percent increase of pairwise divergence. A simple linear regression was also calculated to predict the number of informative sites in an exon-only locus based on the length of the locus. A significant regression equation was found ($F(1, 1045) = 5666$, $p < 0.001$), with an adjusted R^2 of 0.84 (Fig. 13). The number of informative sites is equal to [-1.89 + 0.38*(alignment length)]. Informative sites increased 0.38 base pairs for each base pair increase in alignment length (Fig. 13).

For flanking region alignments, the average number of taxa included is 250, the average length is 243 bp, the average level of missing data is 12.4%, and the average proportion of informative sites is substantially higher than exon-only alignments at 77.4%. The concatenated alignment of the flanking-only loci totals 69,947 base pairs. The average proportion of missing data in the concatenated alignment for the ingroup samples is 15.4% (range 6%– 40%), and is 50.6% (range 42%– 68%) for the outgroup samples. The non-coding flanking loci are generally more difficult to align, especially as phylogenetic distance increases. For the purpose of this study, we performed alignments across all samples, which is likely to have contributed to lower quality alignments and failure to pass specific missing data filters. We therefore emphasize if flanking sequence alignments are performed for the ingroup only, or even subclades of the ingroup, it should result in more and longer alignments recovered.

Discussion

We used a custom transcriptome-based exon capture, designed without the use of a reference genome, to successfully generate a large informative phylogenomic data set across divergent lineages of frogs. We accomplished this using transcriptome sequences directly for probe design, resulting in the additional recovery of a significant amount of highly variable non-coding sequence data. We generated sequence alignments for 1,047 of the 1,260 transcriptome-based loci, with an average of 250 (of 264) taxa present per alignment. The combination of exon and flanking region data resulted in a concatenated alignment of 631,127 base pairs. Based on the results of our experiment, we discuss the overall efficiency of capture, results of using transcript sequences for probe design, effects of phylogenetic distance, and recommendations for pooling size and blocking oligos.

The effects of blocking oligos are non-trivial, and have great potential to affect the capture efficiency (Fig. 4). Although the enrichment of target loci is accomplished using short blockers, universal blockers, and xGen blockers, there are critical differences in the level of depletion of non-targeted loci among blockers. The xGen blockers significantly outperformed the short blockers and universal blockers in the depletion of non-targets (Fig. 4). The higher concentration of DNA in post-capture libraries of the universal and short blockers represents a large carry-over of non-targets, which ultimately translates to significant reductions in both sensitivity and specificity and increases in PCR

duplication rates. This is particularly important to consider for organisms with larger genome sizes, such as amphibians, which are likely to suffer from reductions in sensitivity and specificity and higher duplication rates based on genome size alone. The cost of xGen blockers is significantly higher per reaction, but may ultimately reduce the amount of sequencing effort required to obtain high quality sequence data. We therefore strongly recommend the testing of blocking oligos using a qPCR assay before conducting the main capture experiment, as the specificity, sensitivity, and duplication levels can be greatly improved with appropriate blocking oligos.

A main question concerning sequence capture is simply how many individuals can be pooled in a reaction, which has important implications for reducing costs and increasing the sampling for a given project. We tested a range of pool sizes (1–7 samples) within the genus *Hyperolius* (160 samples, 28 captures) to determine the effects of pooling on raw data yield, sequencing depth, and duplication levels. We found no differences in raw data yield or sequencing depth across pools, but our results indicate duplication levels vary across pooling sizes (Fig. 8). We demonstrate duplication levels rose from 4–5% in 1–2 sample pools to an average of 15–19% in the 4–7 sample pools. These levels were acceptable for obtaining high quality sequence capture data for our experimental design. We did not detect a significant increase in duplication levels for pools of seven samples, suggesting the upper limit for sample pooling was not reached, though this topic requires additional investigation. Although pooling size does affect PCR duplication levels, we again emphasize that these effects can be strongly exacerbated through the use of less efficient blocking oligos.

Phylogenetic distance has a predictable effect on capture sensitivity, specificity, and the proportion of missing data in the final sequence alignments. As expected, more divergent species experienced drops in sensitivity and specificity, and their proportion of missing data increased (Fig. 12). Though these results are intuitive, our findings our useful in at least two ways. First, we demonstrate that for the most distant outgroup in our experiment (family Microhylidae), which shared a common ancestor with the probe design species 103 million years ago, we recovered 23% of the total exon sequence data (roughly 146,000 bp). Our experiment was focused on sequence capture within a single family, but successful sequence capture occurred for highly divergent outgroup species, albeit with a predictable reduction in efficiency. Second, the regression equations relating capture efficiency metrics to average pairwise divergence can serve as a starting point for other researchers in determining the phylogenetic breadth of their capture experiments. Our comparisons are made using nuclear sequence data generated prior to our experiment. These empirical data, though based on frogs, allow an approximation of the effects of phylogenetic distance on metrics generally used to characterize capture efficiency, and can set realistic expectations for the overall success of sequence capture based on phylogeny. This approximation requires Sanger sequencing only a small number of nuclear loci for a subset of the target group, information that should generally be acquired before beginning a large-scale capture experiment.

Our experimental design used transcriptome sequences of four species from divergent ingroup clades to design capture probes, and we recovered high quality sequence data across the ingroup genera. The use of four sets of sequences for each locus ultimately reduced the total number of loci that were included in our probe design, and the trade-off between number of loci and variability in probe design is important to

consider for exon capture design. Unfortunately we cannot assess whether probe sets from certain species were more efficient in capturing sequences, and it is unclear how using a single species would have affected the outcome of this experiment. Using a single species for probe design in our case would have allowed for the inclusion of approximately 5,000 loci, rather than 1,260. A possibility for reducing the number of design species is to reconstruct ancestral sequences for deeper nodes of the ingroup, and use these sequences for probe design. Though there are many options for probe design, our results demonstrate sampling multiple divergent ingroup species is a highly effective strategy for capture across larger phylogenetic scales.

Our experiment tested the direct use of transcriptome sequences for probe design, thereby circumventing the use of reference genomes for identifying intron-exon boundaries to filter out short exons. This approach was highly successful, and we recovered short and long exons with a high uniformity in coverage (Figs. 9, 10) as well as a large quantity of highly variable non-coding flanking regions. The average size of individual exons (153 bp) within loci is close to predictions of average exon lengths across the genome of *Xenopus laevis* (~200 bp), and we found most of the 850 bp transcriptome sequences contained four exons (Figs. 6, 7). We successfully captured large quantities of short exons (< 100 bp) (Fig. 7), a feature that may be appealing for researchers targeting short loci. The probe design spanning intron-exon boundaries did not reduce coverage towards the ends of exons (Figs. 9, 10), and returned thousands of base pairs of non-coding flanking sequences per in-target assembly. The resulting alignments of non-coding regions show high levels of variation, with an average proportion of 77% informative sites (compared to 38% of exon-only regions). These flanking regions can be incorporated into population genetics or phylogenetic analyses, similar to UCE and anchored hybrid enrichment approaches. Our pipeline allows alignments to be made with the full in-target assemblies, exon regions only, or flanking regions only, providing flexibility for decisions about sequence data analysis.

Transcriptome-based exon capture is an effective method for producing large sets of orthologous markers with predictable levels of informativeness in non-model systems. This method can be applied to population level questions by sequencing transcriptomes within a population, or applied to larger phylogenetic scales using the transcriptomes of divergent species. As this approach and other types of sequence capture gain popularity, the reporting of empirical data can enhance the ability of researchers to choose the appropriate capture approach or aid in the design of custom sequence captures. We have outlined our experimental design, including probe design from transcriptome sequences, as well as reaction-specific decisions about blockers and capture pooling schemes. For this type of transcriptome-based exon capture, information about the number of exons in transcripts, their lengths, and the recovery of flanking sequences should be discussed. Finally, efforts to relate any of the above measures to phylogenetic distance would greatly benefit researchers planning a sequence capture experiment for non-model systems.

Acknowledgments

Anticipated coauthorship: Lydia L. Smith, Ke Bi.

Lab work conducted by DMP was funded by a National Science Foundation DDIG (DEB: 1311006), the EECG Research Award (American Genetic Association), and by

D.C. Blackburn, R.C. Bell, and J.A. McGuire. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303. DMP thanks many collaborators and institutions for tissue samples processed in this study, and they will be recognized as co-authors or fully acknowledged in the resulting phylogenetic study in prep. DMP also thanks Sean Reilly and Ammon Corl for help in lab training and troubleshooting.

References

Aljanabi, S., and I. Martinez. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research, 25:4692–4693.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. Journal of Molecular Biology, 215:403–410.

Altschul S.F., T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSIBLAST: a new generation of protein database search programs. Nucleic Acids Research, 25:3389–3402.

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

Bi, K., D. Vanderpool, S. Singhal, T. Linderoth, C. Moritz, and J.M. Good. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13:403.

Bolger, A.M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30:2114–2120.

Bragg, J.G., S. Potter, and C.G. Moritz. Exon capture phylogenomics: efficacy across scales of divergence. Molecular Ecology Resources, Online Early View.

Capella-Gutierrez, S., J.M. Silla-Martinez, and T. Gabaldon. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics, 25:1972–1973.

Faircloth, B.C., J.E. McCormack, N.G. Crawford, M.G. Harvey, R.T. Brumfield, and T.C. Glenn. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic Biology 61:717–726.

Grabherr M, B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology, 15:644–652.

Gregory, T.R. 2015. Animal Genome Size Database. [http://www.genomesize.com]

Hedtke, S.M., M.J. Morgan, D.C. Cannatella, and D.M. Hillis. 2013. Targeted enrichment: maximizing orthologous gene comparisons across deep evolutionary time. PLoS ONE 8, e67908.

Huang, X., and A. Madan. 1999. CAP3: a DNA sequence assembly program. Genome Research, 9:868–877.

Katoh, K., and D.M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution, 30:722–780.

Kent, W. 2002. BLAT–the BLAST-like alignment tool. Genome Research, 12:656–64.

Langmead, B., S. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods, 9:357–359.

Lemmon, A.R., S.A. Emme, and E.M. Lemmon. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Systematic Biology, 61:727–744.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 25:1754–1760.
Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map (SAM) format and SAMtools. Bioinformatics, 25:2078–2079.
Li, W., and A. Godzik. 2006. CD-Hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics, 22:1658–1659.
Liu, B. J. Yuan, S-.M. Yiu, Z. Li, Y. Xie, Y. Chen, Y. Shi, H. Zhang, Y. Li, T-.W. Lam, and R. Luo. 2012. COPE: an accurate k-mer-based pair-end reads connection tool to facilitate genome assembly. Bioinformatics, 28:2870–2874.
Magoč, T., and S. Salzberg. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics, 27:2957–2963.
Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, 17:10–12.
McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M.A. DePristo. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research, 20:1297–1303.
McCartney-Melstad, E., G.G. Mount, and H.B. Shaffer. 2015. Exon capture optimization in large-genome amphibians. bioRxiv doi: http://dx.doi.org/10.1101/021253.
Meyer, M., and M. Kircher. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols, 2010(6): pdb.prot5448.
Miller, M.R., J.P. Dunham, A. Amores, W.A. Cresko, and E.A. Johnson. 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research, 17:240–248.
Peñalba, J.V., L.L. Smith, M.A. Tonione, C. Sass, S.M. Hykin, P.L. Skipwith, J.A. McGuire, R.C. Bowie, and C. Moritz. 2014. Sequence capture using PCR-generated probes: a cost-effective method of targeted high-throughput sequencing for nonmodel organisms. Molecular Ecology Resources, 14:1000–1010.
Rubin, B.E.R., R. Ree, and C.S. Moreau. 2012. Inferring phylogenies from RAD sequence data. PLoS ONE, 7, e33394.
Simpson, J., K. Wong, S. Jackman, J. Schein, S. Jones, and I. Birol. 2009. ABySS: a parallel assembler for short read sequence data. Genome Research, 19:1117–1123.
Singhal, S. 2013. *De novo* transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. Molecular Ecology Resources 13:403–416.
Smit, A.F.A., R. Hubley, and P. Green. 2015. *RepeatMasker Open-4.0*. [http://www.repeatmasker.org]

Figure 4. Quantitiative PCR plots for the positive control nuclear locus (KIAA2013) and the negative control nuclear locus (49065). Assessment of the relative success of the capture can be made by comparing the position of the curves of the library pool prior to capture (black) to the three curves produced by library pools after capture using different blocking oligo strategies (blue, green, red). The largest cycle shifts in both enrichment and depletion occur using the xGen blocking oligos (red), with substantially better performance occurring for the depletion of non-targeted sequences.

Figure 4

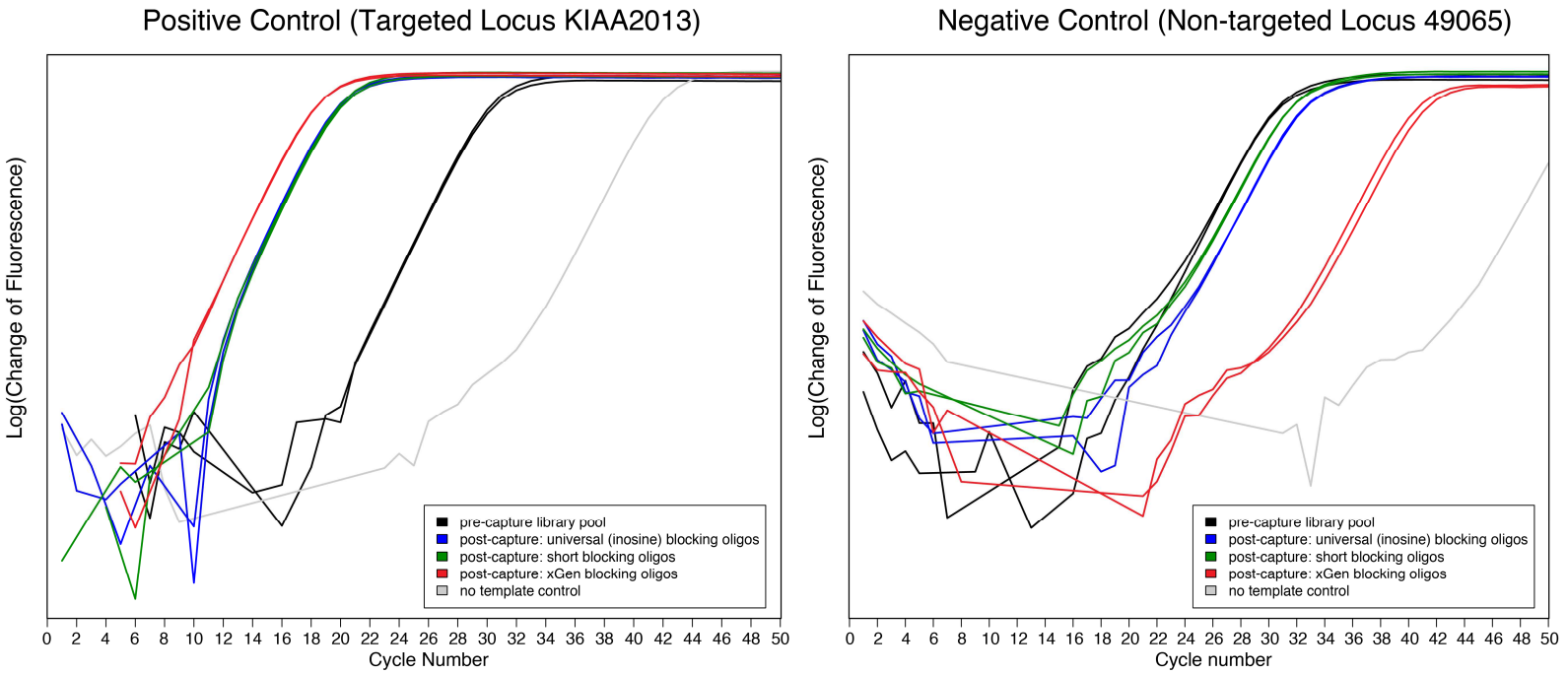

Figure 5. An example sequencing depth (coverage) plot for (A) an in-target assembly and (B) exon-only contig of the same locus from one sample. Exons matching the transcript sequence used for probe design are colored green and labeled (A–E, matching in both plots), and non-coding flanking regions are colored orange. Both 50X and 10X coverage levels are indicated by dotted lines.

Figure 5

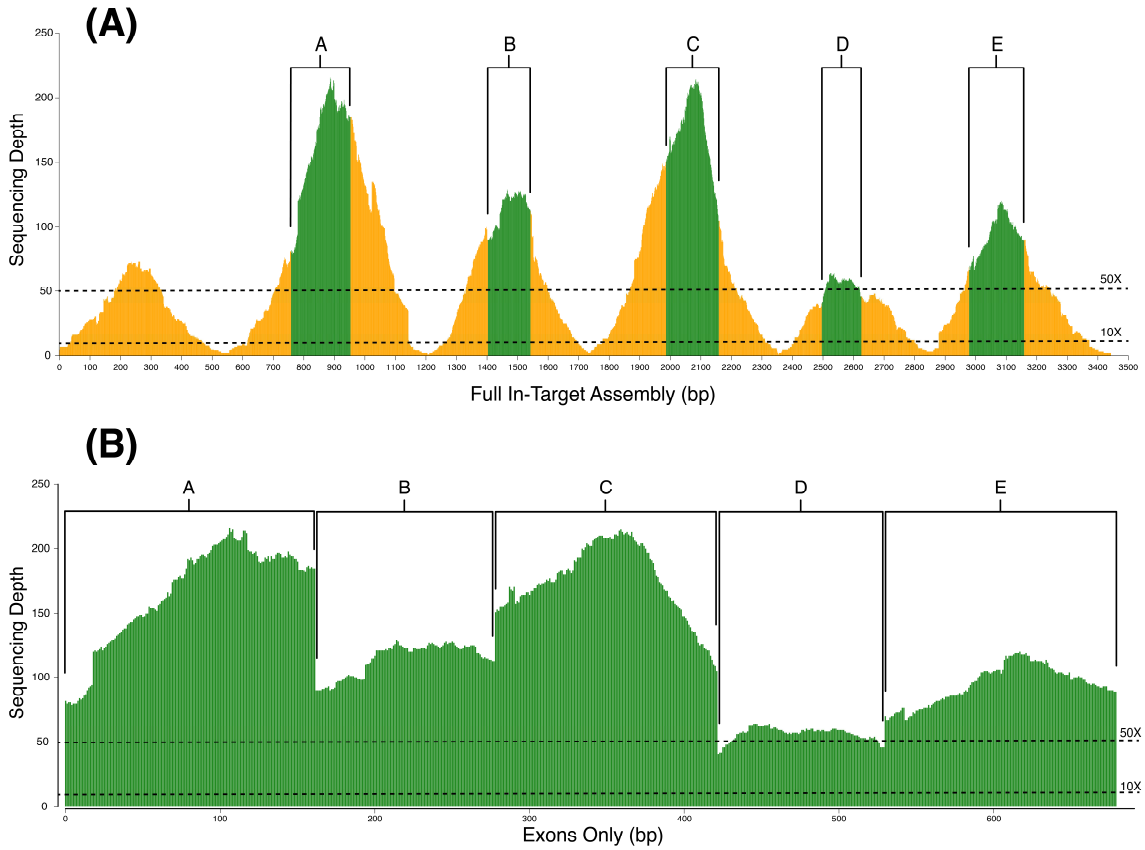

Figure 6. A frequency distribution for the number of exons detected in the fully assembled and merged contigs across all samples. The median number of exons per transcript is four.

Figure 6

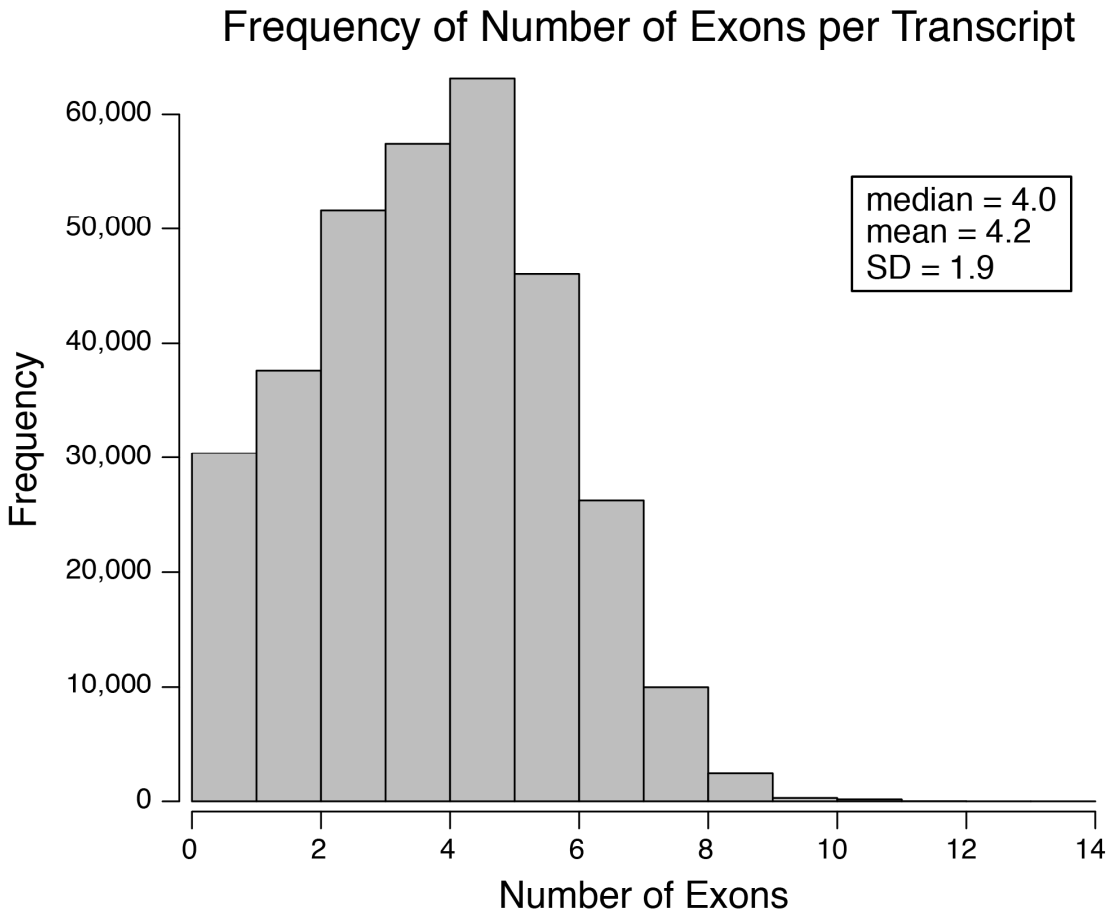

Figure 7. A frequency distribution for the length of each exon detected within a fully assembled and merged contig, across all samples. The average length is 153 bp, and the median length is 132 bp.

Figure 7

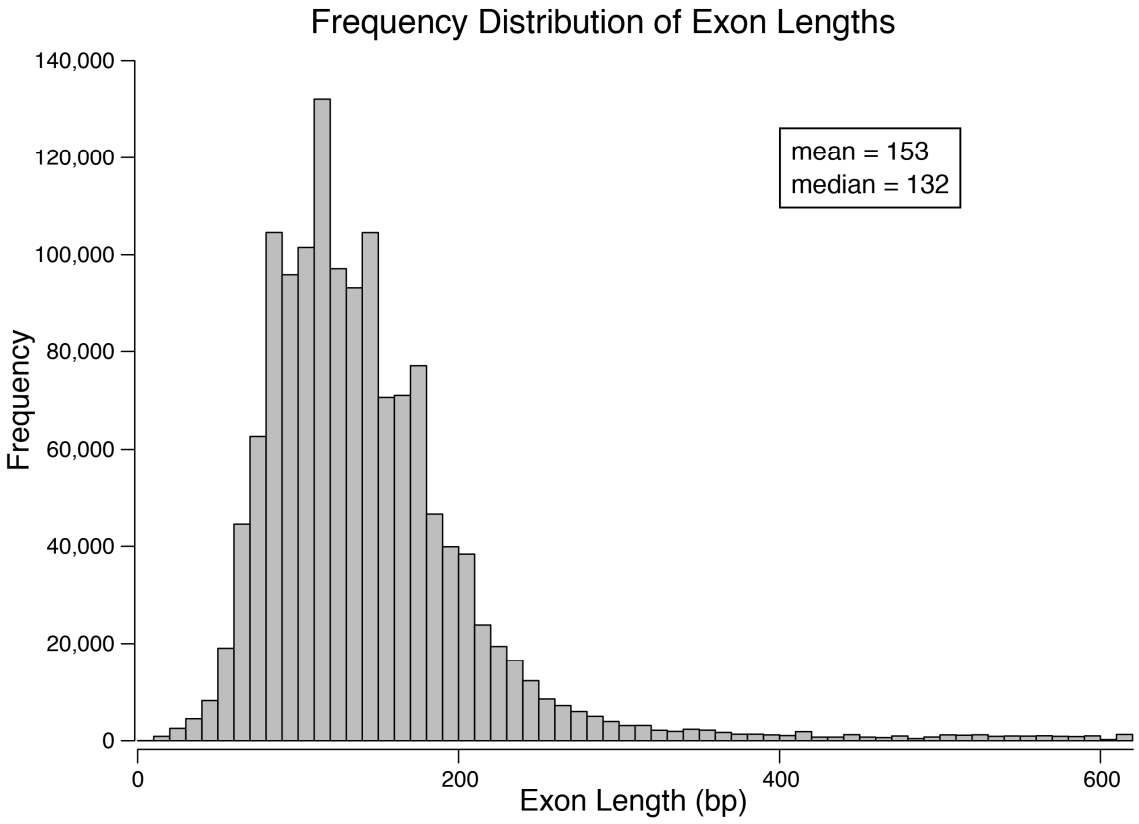

Figure 8. Boxplots of pooling sizes and (A) duplication levels, (B) raw data yield, and (C) sequencing depth. The boxplots depict the median, upper and lower quartiles, and range for each metric.

Figure 8

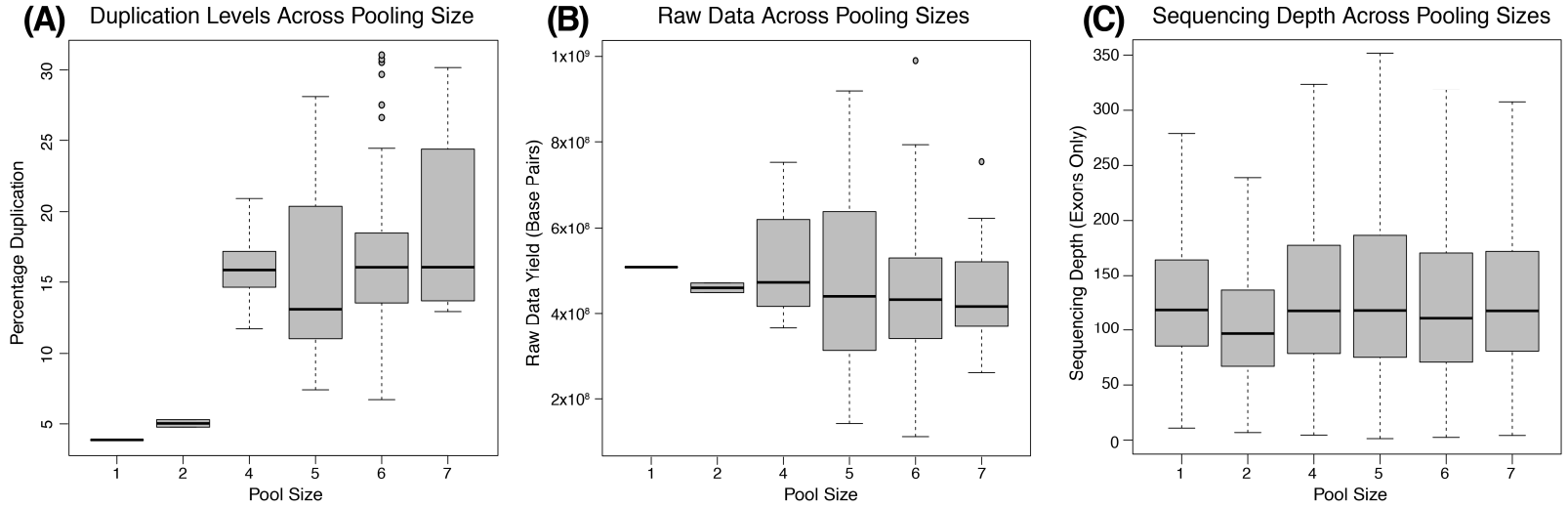

Figure 9. The average binned coverage of the edges of long exons (> 200 base pairs). Bins are in 10 base pair increments, with five bins on the 5' and 3' ends. Estimates are based on 50 randomly chosen ingroup samples.

Figure 9

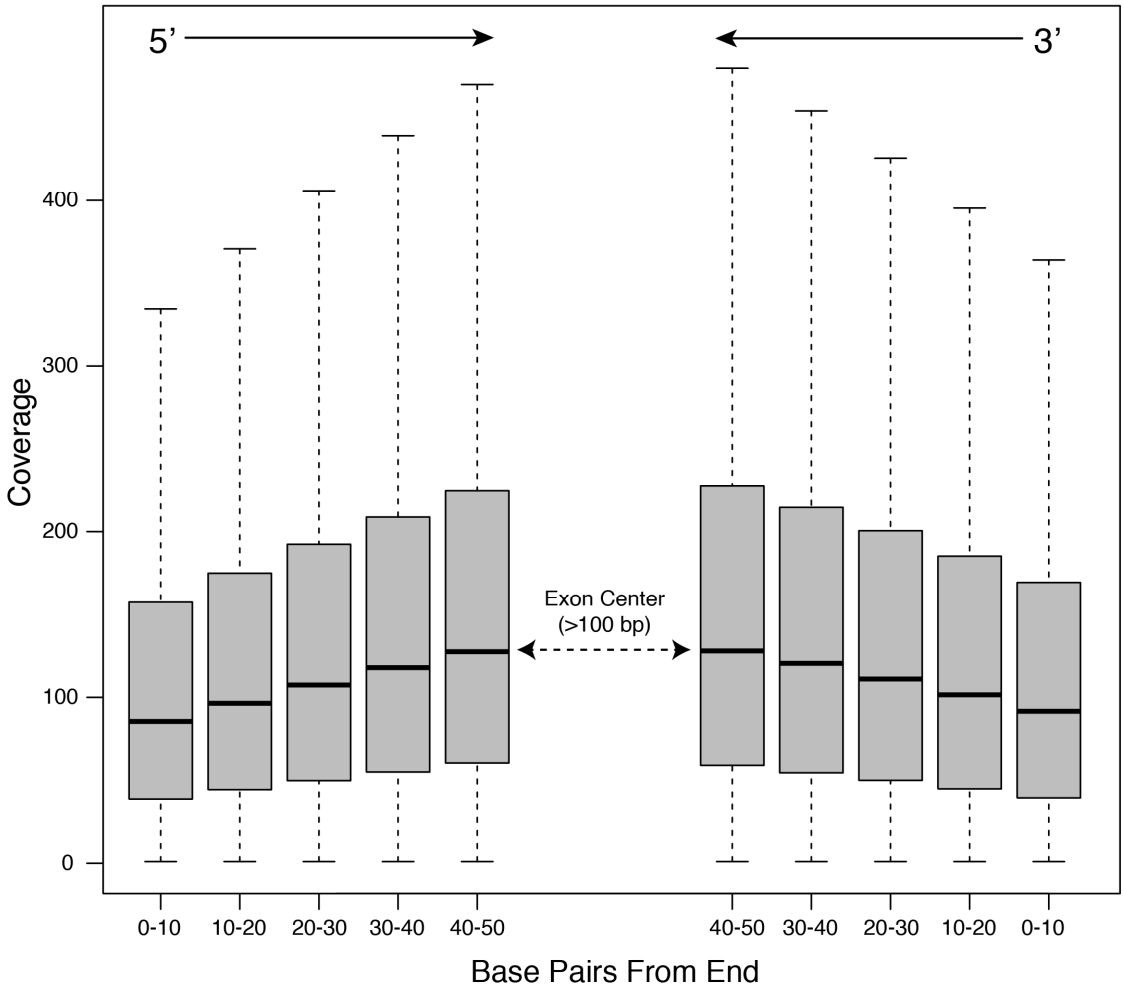

Figure 10. The average binned coverage of the edges of short exons (61–100 base pairs). Bins are in 10 base pair increments, with three bins on the 5' and 3' ends. Estimates are based on 50 randomly chosen ingroup samples.

Figure 10

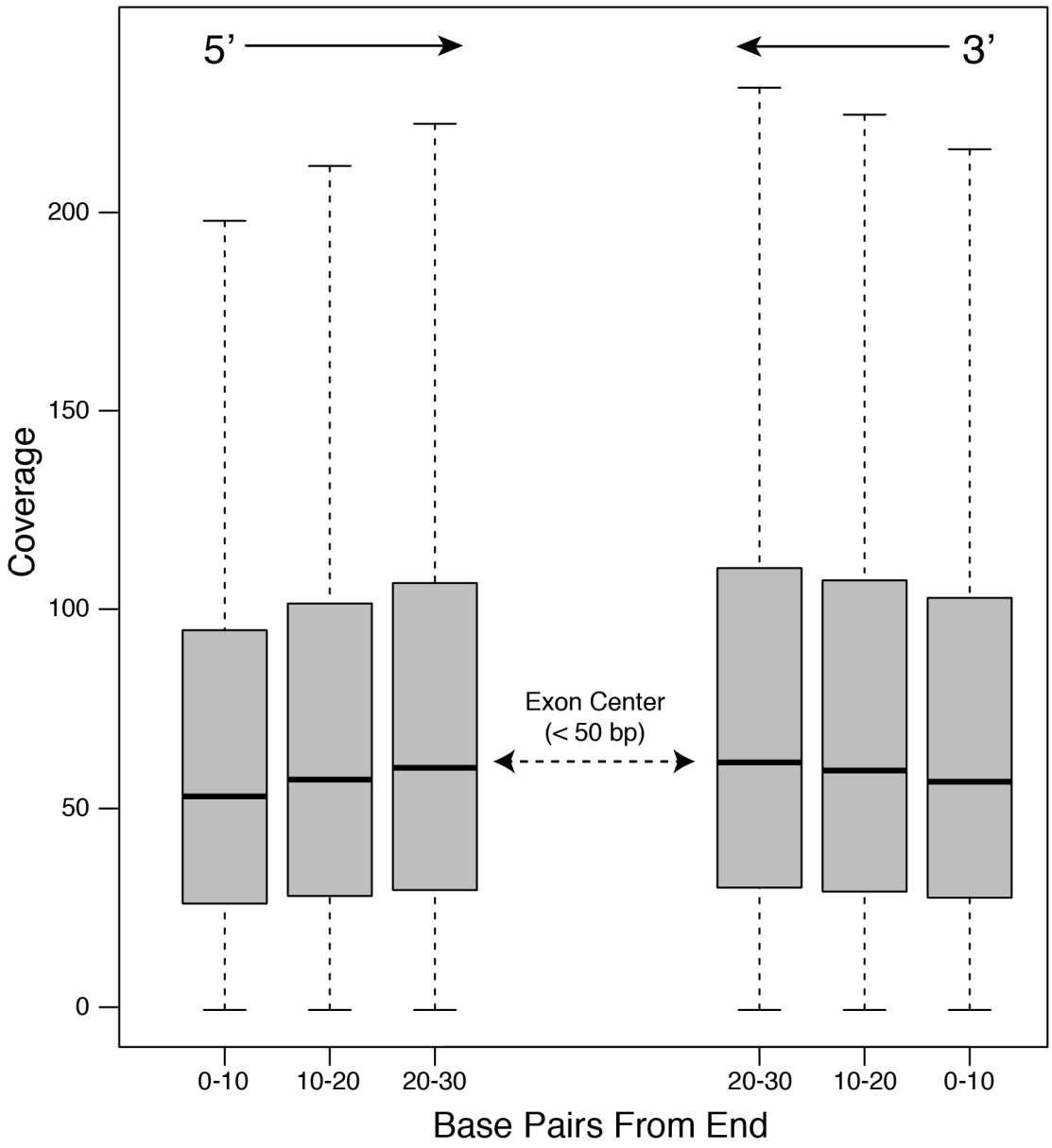

Figure 11. Barplot of (A) sensitivity and (B) specificity, across all samples. Labels A–K refer to ingroup genera denoted by green (A: *Acanthixalus*, B: *Afrixalus*, C: *Cryptothylax*, D: *Heterixalus*, E: *Hyperolius*, F: *Kassina*, G: *Morerella*, H: *Opisthothylax*, I: *Paracassina*, J: *Phlyctimantis*, K: *Tachycnemis*) and labels L–O refer to outgroups denoted by orange (L: Arthroleptidae, M: Brevicipitidae, N: Hemisotidae, O: Microhylidae). Yellow indicates the species used for transcriptome sequencing and probe design.

Figure 11

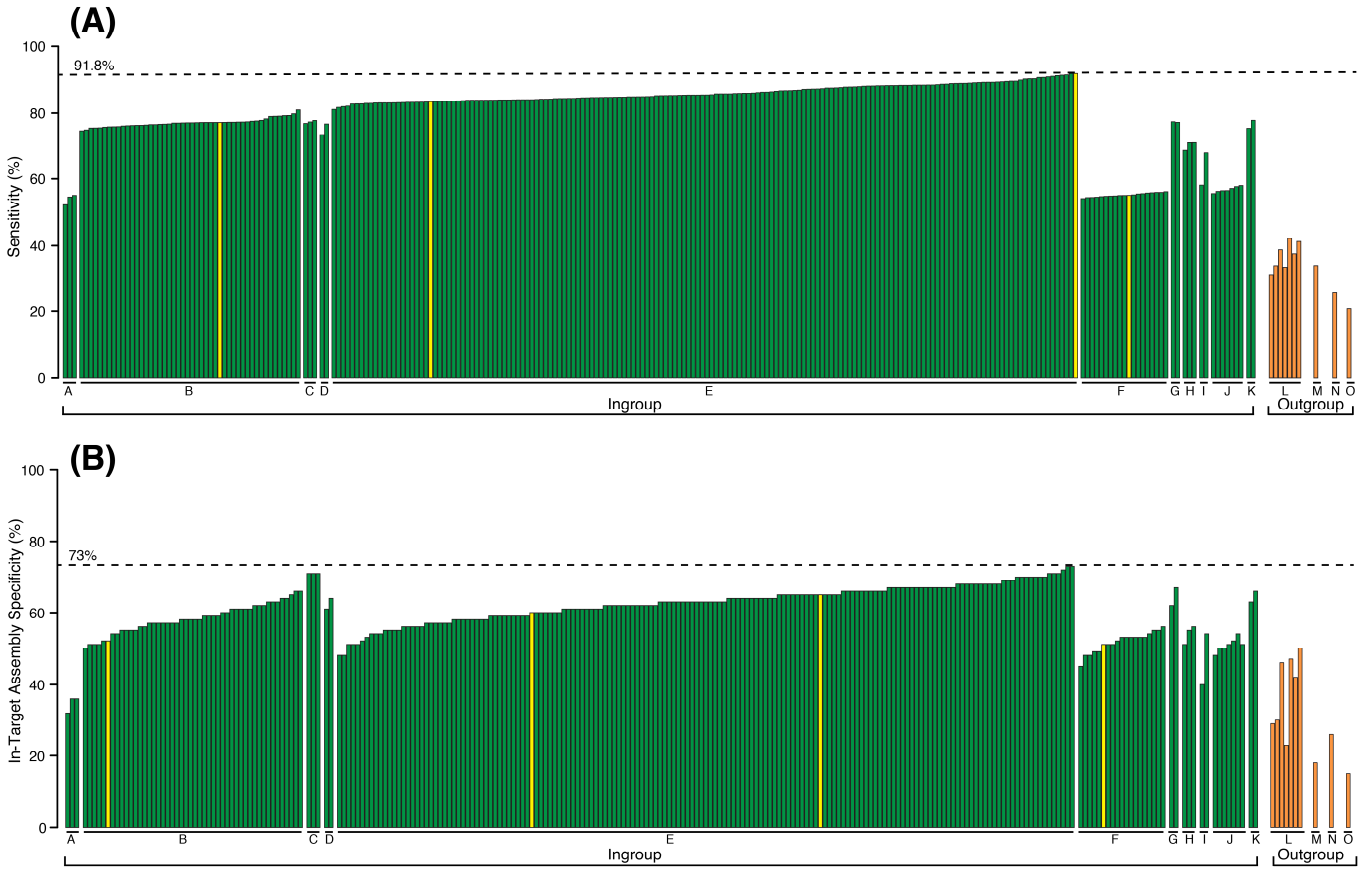

Figure 12. Plots of linear regressions of (A) missing data from the concatenated exon alignment, (B) sensitivity, and (C) specificity, using the average pairwise divergence from probe design species as the independent variable.

Figure 12

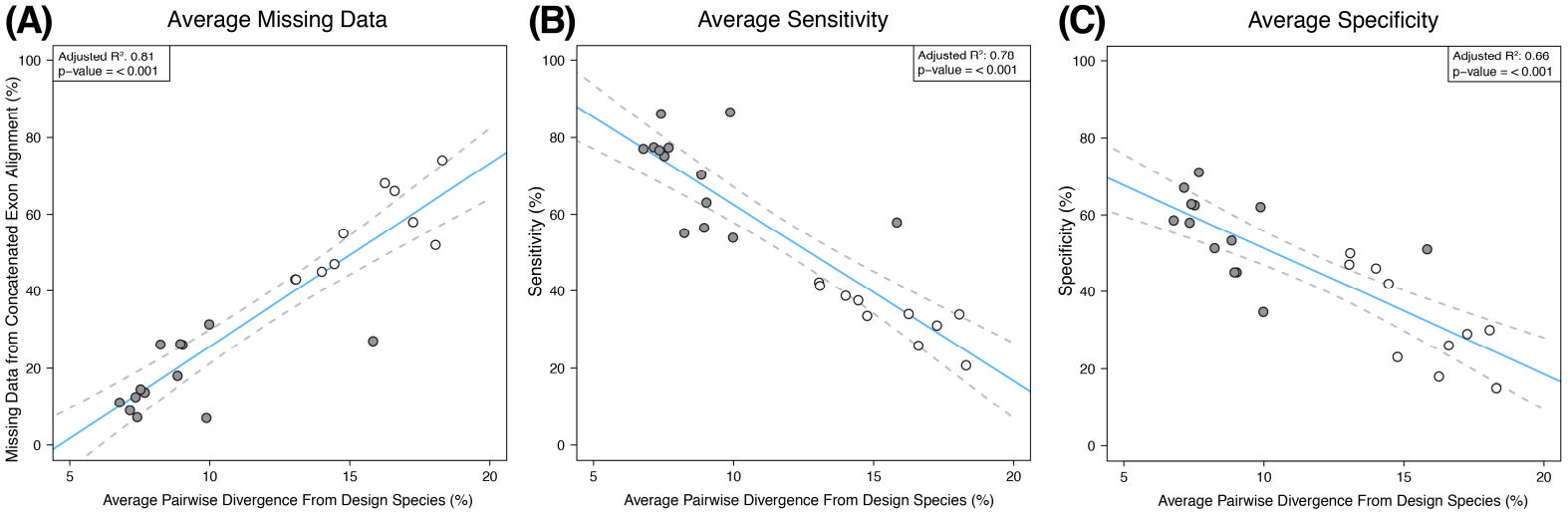

Figure 13. Linear regression of alignment length and the number of informative sites. Each dot represents a unique exon-only alignment, for a total of 1,047 loci. A significant regression equation was found ($F(1, 1045) = 5666$, $p < 0.001$), with an adjusted R^2 of 0.84.

Figure 13

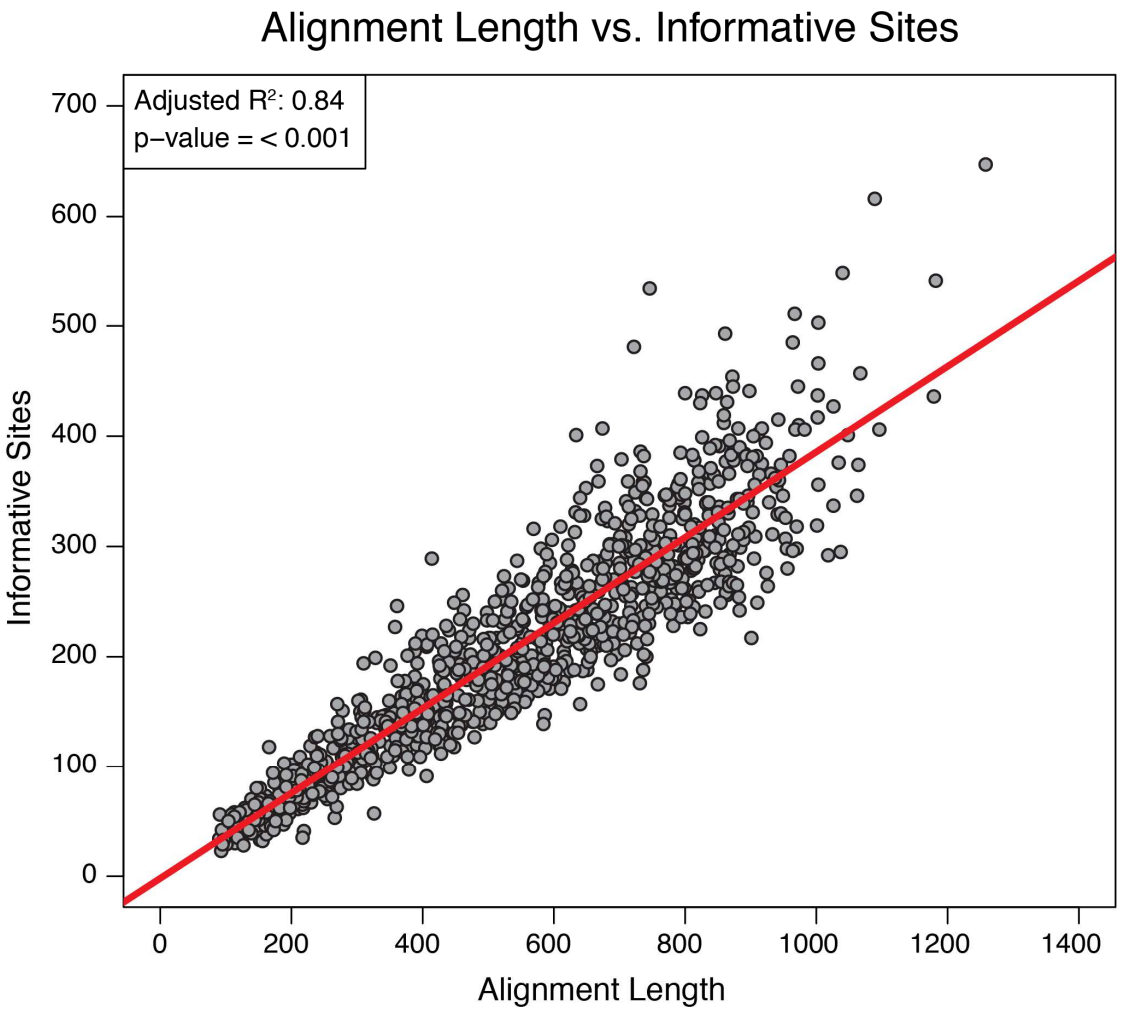

Table 7. Edge Binned Coverage of Long Exons

Table 7

End	5’	5’	5’	5’	5’	3’	3’	3’	3’	3’
Bin	0–10 bp	11–20 bp	21–30 bp	31–40 bp	41–50 bp	41–50 bp	31–40 bp	21–30 bp	11–20 bp	0–10 bp
Mean	117.2	129.8	142.2	154.0	165.2	165.7	156.4	146.1	135.2	124.0
Median	85.5	96.4	107.4	117.8	127.4	127.9	120.4	111.0	101.5	91.6

The mean and median sequencing depth values are given across 10 bp windows up to 50 bases away from the 5’ and 3’ ends of long exons (> 200 bp).

Table 8. Edge Binned Coverage of Short Exons

Table 8

End	5’	5’	5’	3’	3’	3’
Bin	0–10 bp	11–20 bp	21–30 bp	21–30 bp	11–20 bp	0–10 bp
Mean	74.9	79.9	83.7	86.2	83.8	80.4
Median	54.7	58.9	61.9	63.2	61.2	58.4

The mean and median sequencing depth values are given across 10 bp windows up to 30 bases away from the 5’ and 3’ ends of short exons (61–100 bp).

Chapter 3

Molecular systematics of African Reed Frogs (Anura: Hyperoliidae) based on 1,000 nuclear loci and a phylogenomic comparative analysis of sexual dichromatism, sexual size dimorphism, and body size evolution

Introduction

The anuran family Hyperoliidae is Africa's most diverse group of frogs, with a distribution across sub-Saharan Africa (16 genera, over 200 species), Madagascar (1 genus, 11 species), and the Seychelles Islands (1 monotypic genus). Hyperoliids are found in a variety of habitats and altitudes, but exhibit the highest species richness and co-occurring species in forested regions, including lowland rainforests and montane forests. As many as 14 hyperoliid species have been recorded in near sympatry (Lawson & Collett 2011), and species show high turnover based on both elevational and habitat gradients (Poynton & Boycott 1996, Wieczorek et al. 2000, Poynton et al. 2006, Penner et al. 2011). Hyperoliids exhibit extensive morphological diversity, ranging from semi-terrestrial forms, to walker-climbers, to largely arboreal forms, and display considerable variation in body size across species and between the sexes.

The phylogenetic relationships of species within and between genera of hyperoliid frogs are not well known, which has limited comparative evolutionary studies of this family. Phylogenetic studies have been hindered by the group's extremely complicated taxonomic history, high levels of polymorphism and geographic variation, a lack of adequate taxon sampling, and a paucity of multi-locus molecular data for phylogenetic inference. This is not to suggest that the group has received no attention from phylogeneticists. Indeed, several studies using morphological and/or molecular data have attempted to address higher-level hyperoliid relationships (Drewes 1984; Richards & Moore 1996, Wieczorek et al. 2000, Vences et al. 2003, Frost et al. 2006; Rödel et al. 2009, Veith et al. 2009, Portik & Blackburn, in prep), and others have worked to clarify relationships within species complexes (Lawson 2010, Schick et al. 2010; Channing et al. 2013; Conradie et al. 2013; Greenbaum et al. 2013; Liedtke et al. 2014; Bell et al. 2015; Bell et al. 2015; Loader et al. 2015; Bwong et al. in review), thereby making progress on these challenges. The most critical limiting factor has been specimen availability, and our coordinated international team of researchers, through extensive fieldwork, has produced sufficient taxon sampling to begin addressing fine-scale species-level phylogenetic relationships of hyperoliid frogs. Our study is premised on new high-throughput sequencing techniques that allow the collection of large molecular data sets for many samples, thereby resolving the issue of limited multi-locus data for estimating the phylogeny of Hyperoliidae.

Here, I conduct a large-scale study of hyperoliid molecular systematics using phylogenomic data and comprehensive taxon sampling resulting from my fieldwork and the combined efforts of over 30 researchers. Using more than 1,000 nuclear loci, I reconstruct the phylogenetic relationships of Hyperoliidae using concatenation and species tree approaches. I explore the evolution of several traits with the resulting

phylogeny, including body size, sexual size dimorphism, and sexual dichromatism, which are introduced below.

Sexual size dimorphism

Sexual size dimorphism (SSD) is thought to be the result of adaptation of the sexes to their differing reproductive roles, and can involve components of intrasexual selection (competition within a sex) and intersexual selection, such as female preference (Darwin 1874). Male-biased size dimorphism can be expected when larger males have a reproductive advantage in competing for mates, or as a result of female preference. In contrast, cases of female-biased size dimorphism are generally explained as selection favoring larger females because they have increased fecundity. Differences in body sizes may also be the result of ecological divergence as a result of intraspecific competition, manifested as intersexual niche-partitioning (Shine 1989). Ultimately, the degree of SSD is the result of the direction and strength of selection on each sex, as well as constraints imposed by the shared genomes of the sexes (Andersson, 1994; Fairbairn 1997; Fairbairn et al. 2007). One macroevolutionary outcome of SSD is described by Rensch's rule, which states that when males are the larger sex, SSD increases with body size, whereas when females are the larger sex, SSD decreases with body size (Rensch 1950, 1960). This allometric relationship implies that males are generally more variable in body size, and highlights selection for large male sizes (Abouheif & Fairbairn 1997; Fairbairn 1997). However, many female-biased SSD systems are inconsistent with Rensch's rule, indicating stronger selection for female body size (Fairbairn 1997).

Anurans exhibit a wide range of body sizes and degree of sexual size dimorphism, however close to 90% of species evaluated exhibit female-biased SSD (Shine 1979; Han & Fu 2013). The most common explanation for female-biased SSD in frogs is selection on female size resulting in a fecundity advantage (Shine 1988; Andersson 1994), and female fecundity increases more rapidly with body size as compared to males (Crump 1974, Trivers 1972). There is a strong correlation between increases in female size and increases in egg size or clutch size across frogs (Han & Fu 2013), though advantages may be greater for smaller species (De Lisle & Rowe 2013; Nali et al. 2014). However, selection for increased fecundity does not seem to fully explain female-biased SSD in frogs (Monroe et al. 2015). The high prevalence of female-biased anuran species has also led to the proposal of a number of hypotheses relating to selection on male body size, including energetic constraints on males resulting from reproductive behaviors (Woolbright 1983), higher predation rates on males resulting in younger and smaller individuals in the population (Wells 1977; Monnet & Cherry 2002), and the enhanced vagility of smaller male sizes (Ghiselin 1974). In contrast, male-biased SSD is often explained by intrasexual competition in the form of combat. The males of many frog species exhibit territorial behaviors that should favor larger sizes, particularly if combat occurs (Trivers 1972; Wells 1977; Shine 1979). However, a recent comparative study found no correlation between combat and male-biased SSD (Han & Fu 2013), and the males of many species with female-biased SSD also engage in combat.

Meta-analyses of SSD across frog lineages have demonstrated an isometric relationship, rather than the particular allometric relationship of Rensch's rule (De Lisle & Rowe 2013; Han & Fu 2013; Nali et al. 2014), questioning the overall strength of selection on either sex. However, clade-specific dimorphism patterns might be masked

by meta-analyses (De Lisle & Rowe 2013), which generally ignore factors unique to particular clades that potentially drive the evolution of body sizes. Beyond allometric tests for Rensch's rule, there are other methods that evaluate the evolution of body size across the sexes. Comparative phylogenetic methods can be used to estimate rates of character evolution, and to test whether the rate or direction of change differs among lineages with different selection regimes (Butler & King, 2004; Harmon et al., 2008; O'Meara & Beaulieu, 2014). This framework allows for investigation of the overall evolution of body sizes of the sexes across large clades, while also allowing examination of these patterns across regimes defined as subclades or by particular phenotypic traits.

The level of sexual size dimorphism varies substantially across hyperoliid species (pers. obs.), yet the most comprehensive studies have been quite limited, only including eight hyperoliid species (De Lisle & Rowe 2013; Han & Fu 2013). Given the large variation in body size and the large number of hyperoliid species (~230), they are an appropriate system for investigating patterns of sexual size dimorphism. There are several large subclades within the family, as well as smaller clades within the genus *Hyperolius*, that may have experienced different sets of selective pressures resulting in group-specific patterns of SSD. In this study, I quantify the variation in SSD occurring across hyperoliids, examine Rensch's rule across the family and for particular genera, and assess which evolutionary models best explain body size evolution of the sexes.

Sexual dichromatism

A striking and unique feature of many hyperoliid species is their bright coloration and patterning, which can be highly variable and driven by the presence of color polymorphism or sexual dichromatism. At sexual maturation, sex steroids can trigger a change in coloration that creates bright sexual monochromatism (both sexes become brightly colored) or sexual dichromatism, in which females undergo color change and males retain juvenile coloration (Richards 1982; Hayes & Menendez 1999). Currently 40 species of hyperoliids are estimated to display ontogenetic sexual dichromatism (Bell & Zamudio, 2012). Although prevalent in many vertebrate groups, this condition is rare among anurans and has been recorded for only 97 of the 6,500 known frog species (Bell & Zamudio, 2012). In other vertebrates, selection for dichromatism is often correlated with increasing importance of species recognition (Figuerola & Green, 2000). If sexual dichromatism enhances species recognition in hyperoliids, the condition may be more prevalent in regions with high diversity where other mating signals (such as advertisement calls) are not sufficient for identifying conspecifics. Preliminary comparative analyses suggest dichromatic species of hyperoliids have repeatedly evolved from monochromatic species (Veith et al. 2009), but improved taxon sampling and a more comprehensive phylogeny are necessary to evaluate these results.

In addition to the expected set of selective pressures driving body size evolution, the presence of sexual dichromatism is a clade-specific trait that may have played a role in the body size evolution of hyperoliid frogs. It is possible that patterns in the evolution of body size and SSD may differ between dichromatic and monochromatic species, though it is difficult to predict expectations for these relationships. In this study, I investigate the evolution of sexual dichromatism through ancestral character reconstruction, which provides an estimate of the number of transitions between monochromatism and dichromatism and the probability of either state at key nodes. Additionally, I assess

whether there are significant differences in body sizes and SSD between dichromatic and monochromatic species to see if this trait enhances or constrains body size evolution of the sexes.

Methods

Sampling

Sampling was chosen from 1,350 mtDNA barcoded individuals to best represent distinct lineages and capture diversity across species complexes (Portik, unpublished data). Specimen information for the chosen samples in provided in Appendix 3.

Transcriptome-based exon capture

The overall workflow and lab-specific protocols for obtaining transcriptome-based sequence capture data are described in Portik et al. (2015b), and summarized here. I sequenced four transcriptomes of divergent ingroup species and used them to design a MYaits-3 custom bait library (MYcroarray) that targeted 1,265 loci ranging between 500 and 850 base pairs long. I prepared a total of 264 genomic libraries (254 hyperoliid species, 10 outgroups) for multiplexed capture reactions (Meyer & Kircher 2010). The MYbaits capture reactions were performed as described in Portik et al. (2015), and resulting amplified capture products were sequenced on three lanes of an Illumina HiSeq2500 with 100 bp paired-end reads. The raw data were cleaned and assembled following workflows of Singhal et al. (2013), Bi et al. (2012), and Portik et al. (2015b), resulting in the generation of in-target exon assemblies as well as assemblies of non-coding flanking regions for all samples.

Sequence Alignments and Filtering

The in-target exon and the flanking sequences were aligned on a per locus basis using MAFFT (Katoh & Standley 2013), and subsequently trimmed using trimAl (Capella-Gutierrez et al. 2009) using the pipelines of Bi et al. (2012) and Portik et al. (2015b). I enforced multiple post-processing filters, including a minimum alignment length of 90 bp, no more than 80% missing data per sequence in alignments, and no more than 30% total missing data across an entire alignment. These filters were enforced using a custom python script for all alignments (https://github.com/dportik/Alignment_Refiner). A total of 1,047 exon-only alignments and 287 flanking region alignments passed all filtering criteria, for a combined length of 561,180 base pairs (bp) and 69,947 bp, respectively. For this study, I only used exon alignments for phylogenetic analyses.

Phylogenetic Analyses

Phylogenies were reconstructed from the genomic data sets using multiple approaches, including concatenation and species tree estimation. For concatenation, the exon alignments were combined and analyzed using RAxML v8 (Stamatakis 2014) on the Cyberinfrastructure for Phylogenetic Research (CIPRES) gateway (Miller et al. 2010). The rapid bootstrapping and best ML search option was used with 100 bootstrap replicates under the GTRCAT model, treating the data as a single partition.

Species trees were estimated using ASTRAL-II (Mirarab et al. 2014), which uses unrooted gene trees to estimate the species tree. This method uses a quartet-based approach that is consistent under the multi-species coalescent process, and therefore

appropriate for resolving gene tree discordance resulting from incomplete lineage sorting (Mirarab et al. 2014). This method also allows for missing taxa in alignments, which can be problematic for other coalescent-based summary methods such as MP-EST (Liu et al. 2010). Unrooted maximum likelihood gene trees were generated for the exon and flanking regions of each locus using RAxML v8 (Stamatakis 2014) under the GTRCAT model. The full set of gene trees based on exon data (1,047 loci) was then evaluated using ASTRAL II to estimate the species tree.

I investigated the number of loci required to achieve topological convergence using species tree estimation in ASTRAL. To achieve this, 18 bins of subsampled loci (containing 3, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, and 400 loci) were assessed. The unrooted gene trees from the exon-only data set were randomly sampled with replacement to obtain the appropriate number of loci for each bin, with 50 replicates per bin. Each subsampled replicate was analyzed using ASTRAL to obtain a subsampled species tree, for a total of 50 subsampled species trees per bin. The Robinson-Foulds (RF) or symmetric difference metric was used to measure the distance between these unrooted phylogenetic trees (Robinson and Foulds 1981). The subsampled species trees were compared to the ASTRAL species tree and the concatenated maximum likelihood tree obtained from the full exon-only data set, which were treated as optimal topologies (RF-optimal). Because the results of this comparison rely directly on the subjective decision of the "optimal" topology, a second evaluation was performed that only compared species trees within a given size bin to one another (RF-internal). The creation of randomly subsampled datasets across user-selected bin sizes, serial analysis with ASTRAL, and calculation of RF distances was streamlined using custom python scripts in the workflow package ASTRALnaut.

Morphological Data

I obtained body size measurements in the form of snout-urostyle length (SUL) for 2,757 preserved specimens representing 140 unique lineages. Measurements were made with a Mitutoyo Series 500 Digimatic Caliper (Mitutoyo U.S.A., Illinois) and recorded to the nearest 0.1 millimeters. Body size data were available for at least one male and one female for 118 of these taxa. These measurements were supplemented with literature data, which included specific measurements of a series of specimens for eight species (Schiøtz 1967; Dehling 2012; Loader et al. 2015), summary information in the form of an average taken from a series for eleven species (Nussbaum and Wu 1995; Rödel et al. 2003; Roedel et al. 2009; Schick et al. 2010; Amiet 2012; Conradie et al. 2012; Conradie et al. 2013), and the maximum size recorded for three species (Channing 2001). The literature data added 19 taxa, for a total of 137 species with at least one body size measurement per sex.

Sexual Size Dimorphism and Assessing Rensch's Rule

I calculated two sexual size dimorphism indices (SSDi) to quantify the level of sexual size dimorphism (SSD) across species. The first SSDi is noted as SSDi1, and SSDi1 = [(larger sex / smaller sex) − 1], set negative if males are the larger sex and positive if females are the larger sex (Lovich & Gibbons 1992). The SSDi1 has been widely used, is properly scaled, is symmetric around zero, and has high intuitive value because positive SSDi1 values indicate female-biased dimorphism and negative values indicate male-

biased dimorphism (Lovich & Gibbons 1992). I calculated a second SSDi (SSDi2) using a two-step procedure (Smith 1999, Nali et al. 2014), where M represents male size and F represents female size. If $M \geq F$, SSDi2 = (M / F), or else if $M < F$, SSDi2 = 2 - (F / M). The SSDi2 is symmetrical around 1, with SSDi2 values < 1 indicating female-biased dimorphism. This metric is not as intuitive and was not used in comparative analyses, and was only included to allow comparison to previous work on anuran SSD (Nali et al. 2014). I visualized the evolution of sexual size dimorphism by reconstructing ancestral character values for SSDi1 on the phylogeny using maximum likelihood with the fastAnc function in phytools (Felsenstein 1985; Schluter et al. 1997; Revell 2012, 2013).

Bivariate scatterplots were created from the log-transformed average body sizes of males and females across species, with female body size plotted on the x-axis. To assess Rensch's rule, I performed regressions to assess whether the slope (β) was significantly different from an isometric relationship ($\beta = 1$). The estimates of body sizes for both sexes (x and y) both contain error. Model I regressions (including ordinary least squares, OLS) assume the values for the x-axis are known without error, and consequently all error can be attributed to the y-axis measurements (Smith 2009). The OLS equation is not symmetrical, producing different predictions depending on which variable is x and y, and these methods also require a choice of sex to be dependent and independent variables. I therefore used reduced major axis regression (a Model II regression method), which is symmetric in producing a single regression line regardless of which sex is assigned to x or y. Reduced major axis (RMA) regression is appropriate for assessing if the slope is equal to or significantly different from a particular value (Warton et al. 2006), in this case if the relationship is isometric ($\beta = 1$) or exhibits allometry conforming to Rensch's rule ($\beta > 1$) or the converse ($\beta < 1$) (Abouheif and Fairbairn 1997). I performed RMA using the lmodel2 package (Legendre 2014) for R on the phylogenetically uncorrected data, including calculation of β and 95% confidence intervals. To account for the shared evolutionary history of these species and non-independence of data points, I also performed a phylogenetic reduced major axis regression (pRMA) (Ives et al. 2007) using the phyl.RMA function in phytools (Revell 2012). Hypothesis testing for $\beta = 1$ was calculated following Ives et al. (2007) and implemented in phytools also using the phyl.RMA function (Revell 2012). These analyses were performed for the complete data set of hyperoliids to investigate the overall relationship of body sizes between the sexes on a family level. To investigate potential genus-level patterns, these analyses were also performed for *Hyperolius* and *Afrixalus* independently.

Body Size Evolution

I visualized the evolution of body size by reconstructing ancestral character values for the body sizes of each sex on the phylogeny using maximum likelihood with the fastAnc function in phytools (Felsenstein 1985; Schluter et al. 1997; Revell 2012, 2013). To examine patterns of sex-specific size evolution, I examined body size evolution using Brownian motion and Ornstein-Uhlenbeck-based (OU) models implemented in OUwie (Beaulieu et al. 2012; O'Meara and Beaulieu 2014). This set of OU models allows the strength of selection and the rate of stochastic motion to vary across selective regimes, which are typically designated *a priori*. I tested if the body size of each sex evolves under the following models: BM: single rate Brownian motion; BMS: Brownian motion

with multiple rates; OU1: a classic OU model (Hansen 1997; Butler & King 2004), OUM: OU model with multiple optima (θ); OUMV: OU model with multiple optima and stochastic rates (σ^2); OUMA: OU model with multiple optima and selection strengths (α); and OUMVA: OU model with multiple optima, selection strengths, and stochastic rates. I checked the results of each model to ensure that the eigenvalues of the Hessian were positive, which indicates the model parameters were reliably estimated (Beaulieu et al. 2012). If negative values were found, the results for the model were not included for model comparisons. I regarded support for the BM and OU1 models as evidence that body sizes are evolving independently of selective regimes, with more complex OU models suggesting traits are evolving towards different optima corresponding to selective regimes (Butler & King 2004). I compared the fit of these models to the body size data using the Akaike information criterion corrected for sample size (AICc). I calculated delta AICc scores, and AICc values were converted to proportional Akaike weights to evaluate relative model support.

Selective regimes were investigated on two levels. The first set of analyses defined three regimes representing major clades across the family: kassinoids, *Afrixalus*, and *Hyperolius* (Fig. 15). The second set of analyses investigated patterns within the genus *Hyperolius*, using three well-defined clades revealed by the phylogenetic analyses (*Hyperolius* clade 1, clade 2, clade 3; Fig. 15).

Evolution of Sexual Dichromatism

For hyperoliid species included in the phylogeny, I scored the presence or absence of sexual dichromatism, coded as a discrete binary character. These data were compiled using a number of published sources (Schiøtz 1967, 1999; Channing 2001; Channing & Howell 2006; Rödel et al. 2009; Veith et al. 2009; Amiet 2012; Bell & Zamudio 2012; Conradie et al. 2013), my personal observations in the field, or my examination of museum specimens. A summary of the sexual dichromatism data is provided in Appendix 5. Ancestral character estimation requires an ultrametric tree with branch lengths proportional to actual or relative time. I used the optimal maximum likelihood topology inferred using RAxML for starting branch lengths, as ASTRAL returns species trees with no branch length information. These branch lengths proportional to substitutions per site were transformed using penalized likelihood (Sanderson 2002) with the chronos function in APE package in R (Paradis et al. 2004), and taxa with missing character data were pruned from the tree. The ancestral states of this trait were examined using stochastic character mapping to estimate the marginal probabilities for all nodes based on joint sampling. The equal rates model was implemented using the make.simmap function of the phytools package in R (Revell 2012), and I performed 1,000 replicates on this single best tree. The number of transitions between character states and proportion of time spent in each state across all replicates was summarized using the describe.simmap function in phytools.

To investigate the potential relationship between SSD and sexual dichromatism, I performed a phylogenetic analysis of variance (pANOVA) (Garland et al. 1993) to evaluate if values of SSD differed significantly between monochromatic and dichromatic species. This type of analysis was conducted using the aov.phylog function in the R package GEIGER (Harmon et al. 2008) and using the phylANOVA function in the R package phytools (Revell 2012). These methods calculate an F-statistic from a standard

ANOVA, but estimate significance by simulating traits and calculating F-statistic distributions using the phylogeny (Garland et al. 1993). The relationship between the body sizes of each sex and sexual dichromatism was also investigated using pANOVA.

Results

Phlyogenetic Analyses

The investigation of topological convergence revealed similar results for the best tree comparisons (ASTRAL and RAxML) and internal comparisons (Fig. 14). Drastic gains in consistency were obtained by increasing the number of gene trees/loci under consideration from 3 to 30, and variation in the species tree estimates decreased. There was additional improvement shown between 30 and 70 loci, but the improvements diminished with increasing loci (Fig. 14). Nevertheless, slight improvements in the consistency of species tree estimation were detected up to the highest bin size (350 loci), suggesting that adding additional loci even to very large data sets continues to improve species tree estimates.

The phylogenetic topologies recovered by the species tree and concatenation methods are largely concordant, but because the maximum likelihood estimate provides branch lengths and support values this topology is shown (Fig. 15). The analysis recovers a monophyletic Hyperoliidae with strong support, though placement of genera is somewhat different from previous analyses. The deepest split in the family occurs between the kassinoids + *Acanthixalus* and all other genera, and within the former group *Acanthixalus* is quite divergent from the kassinoids, which include *Semnodactylus*, *Paracassina*, *Phlyctimantis*, and *Kassina* (Fig. 15). Outside of this group, *Opisthothylax* is sister to the remaining hyperoliid genera. The genus *Afrixalus* is sister to a clade containing the Malagasy and Seychelles Island genera, *Tachycnemis* and *Heterixalus*, along with *Afrixalus enseticola*, an Ethiopian highland species (Fig. 15). The genera *Morerella* and *Cryptothylax* form a sister relationship, and together are sister to all species of *Hyperolius* (including *Chlorolius*). There are some early diverging taxa in the genus *Hyperolius*, including *Hyperolius semidiscus*, *Hyperolius parkeri*, and the *Hyperolius nasutus*/*pusillus* complex (Fig. 15). There are two major clades recovered for the remaining species. The first clade contains a well-supported split, and here the subclades are referred to as *Hyperolius* clade 1 and 2 (Fig. 15). The second major clade is referred to as *Hyperolius* clade 3, and in this group subclades are not readily defined (Fig. 15). The taxonomic implications of the phylogenetic results are beyond the scope of this study, but these relationships form a solid foundation to disentangle the complicated taxonomy of this family.

Sexual Size Dimorphism and Assessing Rensch's Rule

Across the 137 taxa, the average SSDi1 is 0.130 (range: -0.081 to 0.437) and average SSDi2 is 0.87 (range: 0.563–1.08). An SSDi1 of zero indicates lack of dimorphism, with SSDi1 greater than zero indicating female-biased dimorphism. There are 116 taxa with positive SSDi1 values and therefore female-biased dimorphism (85%), but there are 21 species (15%) with equally sized sexes or male-biased dimorphism (Fig. 16, Appendix 4). Several of the species with male-biased dimorphism are contained in the *Hyperolius marmoratus*/*parallelus* group (n = 7). The highest variation in sexual size dimorphism occurs in the genus *Hyperolius*, which spans almost the entire range of this measure (-

0.074 to 0.432; range = 0.506) for the family (Fig. 17). Comparatively, there is much less variation in sexual size dimorphism in the genus *Afrixalus* (-0.013 to 0.242; range = 0.255) and within the kassinoids (-0.05 to 0.130; range = 0.135) (Appendix 4).

The standard RMA regressions for Hyperoliidae, *Hyperolius*, and *Afrixalus* resulted in slopes that did not differ significantly from $\beta = 1$ as indicated by the 95% confidence intervals (Table 9). These results conform to a pattern of isometry, rather than adherence to Rensch's Rule or the converse of Rensch's Rule. The pRMA regressions, which account for shared evolutionary history, are mostly consistent with these results. The slopes for Hyperoliidae ($\beta = 1.08$, p = 0.12) and for *Afrixalus* ($\beta = 0.94$, p = 0.55) are close to and not significantly different from $\beta = 1$ (Table 10). When accounting for phylogenetic relationships, the slope for *Hyperolius* is much higher and consistent with Rensch's Rule ($\beta = 1.13$), but not significantly different from one (p = 0.06) (Fig. 16, Table 11). Though not statistically significant, this trend indicates a higher variation occurring in male body sizes with respect to female body sizes in the genus *Hyperolius*, a result corroborated by the high variation in size dimorphism values for this genus (Fig. 17).

Body Size Evolution

Reconstructions of body sizes across species are illustrated in Figure 18. Many species show intermediate values for size in both sexes, though a number of *Afrixalus* and *Hyperolius* lineages exhibit decreases in male sizes without a corresponding decrease in female size (Fig. 18). The largest body sizes occur in the kassinoids, particularly in the arboreal genus *Phlyctimantis*, in the species *Cryptothylax greshoffi*, and in the arboreal *Tachycnemis seychellensis*, which is restricted to the Seychelles Islands (Fig. 18).

The investigation of body size evolution of the sexes revealed the Brownian motion models were a poor fit, both across Hyperoliidae and within the genus *Hyperolius* (Table 11). Across the three-clade hyperoliid data set, the OU models were a better fit. The single optimum OU model (OU1) was the best-supported model for female body size based on AICc weight, but several other models fit similarly well (OUM, BM, BMS) (Table 11). The OU model with multiple optima (OUM) was the best-supported model for male body size evolution (Table 3), indicating multiple optima exist for males ($\theta_1 = 3.71$, $\theta_2 = 3.18$, $\theta_3 = 3.28$) across each major clade (kassinoids, *Afrixalus, Hyperolius*) (Table 12). The estimated selection strengths and evolutionary rates are higher for male body size ($\alpha = 3.88$, $\sigma^2 = 0.31$) than for female body size of the OU1 model ($\alpha = 1.02$, $\sigma^2 = 0.19$) (Table 12). Similar results were recovered for the three-regime *Hyperolius* data set, though model selection is clearer for both sexes. For the body size evolution of female *Hyperolius*, the single optimum OU model (OU1) was the best supported ($\theta = 3.39$), and for male *Hyperolius* an OU model with multiple optima ($\theta_1 = 3.10$, $\theta_2 = 3.37$, $\theta_3 = 3.22$) and multiple stochastic rates ($\sigma^2_1 = 0.27$, $\sigma^2_2 = 0.28$, $\sigma^2_3 = 0.76$) (OUMV) was the best fit (Table 12). Within *Hyperolius*, the selection strength is again higher for male body size evolution ($\alpha = 12.04$) relative to female body size evolution ($\alpha = 8.30$).

Evolution of Sexual Dichromatism

I scored 71 lineages of hyperoliids as sexually dichromatic (51%) and 67 lineages as sexually monochromatic (49%) (Appendix 5). Maximum likelihood ancestral character estimations of sexual dichromatism estimate the state of the common ancestor of

Hyperoliidae was monochromatic (posterior probability [PP] = 0.65) rather than dichromatic (PP = 0.34). Under an equal rates model, the transition rate is 1.28 with approximately 38.3 total state changes across the phylogeny. Over the 1,000 simulations, there are approximately 14.1 transitions to dichromatism from monochromatism and 24.2 transitions to monochromatism from dichromatism. The proportional time spent in dichromatism is 0.47, whereas monochromatism is 0.52.

Across the phylogeny there are two key transitions to dichromatism in hyperoliids. One of these major transitions occurred in the common ancestor of *Tachycnemis* and *Heterixalus* (PP = 0.68), which are genera restricted to Madagascar and the Seychelles Islands (Fig. 19). The second key transition occurred in the common ancestor of *Morerella, Cryptothylax,* and *Hyperolius* (PP = 0.65) (Fig. 19). Within this group, nearly all the descending nodes are reconstructed unambiguously as sexually dichromatic. All species occurring in *Hyperolius* clades 1 and 2 are sexually dichromatic, with the sole exception of *H. lamottei*. In contrast, there are at least 12 reversals to monochromatism occurring in *Hyperolius* clade 3, including some small subclades that are mostly monochromatic (Fig. 19). The *Hyperolius* clade 3 appears to be overall more variable in this trait than *Hyperolius* clades 1 and 2. The only other hyperoliid species that have dispersed to oceanic islands (*H. molleri* and *H. thomensis*, São Tomé and Principe) appear to have lost dichromatism (Fig. 19), in contrast to the gain of dichromatism in *Tachycnemis* and *Heterixalus*.

There are no significant differences in measures of sexual size dimorphism between dichromatic and monochromatic species based on the phylogenetic ANOVA (Table 5). Additionally, there were no significant differences in male or female body sizes between dichromatic and monochromatic species of hyperoliids (Table 13).

Discussion

Assessment of Rensch's Rule in hyperoliids

The degree of sexual size dimorphism in a species is the net difference between the sum of all the selective pressures affecting males and the sum of those affecting females (Ralls 1976, 1977; Andersson, 1994; Fairbairn 1997), yet this result can also be constrained by the genome shared by the sexes and the evolutionary trajectory of a clade. The macroevolutionary pattern described by Rensch tends to hold true for male-biased SSD systems, in which selection is often higher for male body sizes, and predictable allometric scaling is produced as species become larger (Rensch 1950, 1960; Abouheif & Fairbairn 1997; Fairbairn 1997). Interestingly, in systems that exhibit female-biased SSD there is typically an isometric relationship between the sexes, and less frequently, the converse of Rensch's Rule (Blanckenhorn et al. 2007; Cox et al. 2007; Webb & Freckleton 2007). Frogs are predominately a female-biased SSD group (Shine 1979; Han & Fu 2013) and have been demonstrated to show an isometric relationship between the sexes (De Lisle & Rowe 2013; Han & Fu 2013; Nali et al. 2014), with few clade-specific exceptions.

My examination of the body sizes of hyperoliid species revealed this family largely exhibits female-biased SSD, and I do not find support for Rensch's Rule in Hyperoliidae (Figs. 16, 17; Appendix 4). The average SSDi value of hyperoliid frogs (0.130, n = 137) is similar to the mean estimated for all frogs (0.163, n = 688), and well within the expected range of SSDi (hyperoliids: -0.081 to 0.437; all frogs: -0.43 to 1.12) (Han & Fu 2013). My analysis revealed patterns that are consistent with isometric

scaling between the sexes across the family and within genera (Fig. 16; Tables 9, 10), though the genus *Hyperolius* shows a trend towards Rensch's Rule when phylogeny is accounted for. This trend partially reflects the considerable variation in sexual size dimorphism for the genus *Hyperolius*. Variation in this genus is substantially greater than variation within the genus *Afrixalus* or the kassinoid frogs, and accounts for nearly the entire range of SSDi values (Fig. 17; Appendix 4). Though hyperoliids predominately exhibit female-biased SSD, in several taxa male-biased SSD does occur (Figs. 16, 17; Appendix 4). These are particularly interesting cases for exploring the selective forces operating on male body size, but they require careful examination of behavior, reproductive mode, and mating system to draw firm conclusions, and these data are not readily available for most hyperoliid species.

Evolution of body size in hyperoliids

Overall, sexual size dimorphism in anurans is a composite character, affected by mating systems, reproductive ecology and behavior, intrasexual niche-partitioning, mortality rates, growth rates, and population age structures (Shine 1979; Woolbright 1983; Shine 1989; Monnet & Cherry 2002; De Lisle & Rowe 2013; Nali et al. 2014). The most general explanation for female-biased SSD and selection on female body size is the fecundity advantage. A correlation between female size and clutch size has been demonstrated across all frogs (Han & Fu 2013), though this does not promote allometric scaling converse to Rensch's Rule (Monroe et al. 2015). The relationship between female size and clutch size has not been demonstrated in hyperoliids, though recent collections of egg masses of particular species during fieldwork (Portik, Jongsma, & Blackburn) will allow this relationship to be explicitly quantified.

Changes in SSDi can result from increases in female size or decreases in male size as well as stochastic evolution, and so a fundamental question remains: what are the overall evolutionary tendencies of male and female body sizes for this clade? Across the entire family, no clear patterns emerged (Table 11), which is not unexpected considering the tremendous differences in ecologies and reproductive modes across species. However, within the genus *Hyperolius* I found significant differences in the evolutionary trend of male and female body sizes, namely selection for a single optimal female body size and strong directional selection for males occurring in different clades. This optimal female size is reflected in the ancestral reconstructions, which show little variation in female body sizes across *Hyperolius* clades 1–3 (Fig. 18). There is considerably more variation in male body size, which is captured in the allometric plot (Fig. 16), the ancestral reconstruction of male body size (Fig. 18), and the estimates of male body size optima across clades of *Hyperolius* (Table 12). Inspection of the ancestral reconstructions of body size demonstrates males are both increasing and decreasing in size, whereas females are remaining relatively consistent in body size. Therefore, the relative changes in male body size appear to be the main driver of the variation in SSD in the genus *Hyperolius*.

The consistency in female body size is an interesting pattern, as increases in female size are expected based on sexual selection theory (Darwin 1874). Besides fecundity, there are other characteristics of reproductive mode than may help explain this evolutionary pattern. A vast majority of *Hyperolius* species utilize arboreal sites for oviposition, often on small emergent vegetation in lentic water systems (Portik &

Blackburn, in prep). This reproductive trait may constrain the maximum body size of females because the vegetation may not support the weight of larger females, particular on the ends of leaves where nests are often created. The females of other African arboreal frogs (*Leptopelis*, *Phlyctimantis*) reach considerably larger sizes than *Hyperolius*. However, these groups are using terrestrial or aquatic oviposition, rather than arboreal oviposition, relaxing this hypothetical constraint on larger body sizes. Though currently speculative, the functional relationship between female body size and the creation of arboreal egg nests may be opposing the selection for increased fecundity associated with larger female sizes. I propose this relationship can be tested by characterizing the properties of vegetation selected by females for arboreal oviposition, determining their body size and mass, and subsequently examining the weight limitations of those arboreal oviposition sites.

The disparity in *Hyperolius* male body sizes is apparent from all analyses, and a surprising result is the detection of selection for smaller males in some subclades. Sexual selection theory generally predicts increases in male sizes as a result of intersexual competition or female preference, but few explanations are given for selection of smaller males. These hypotheses are related to increased vagility and ability to locate females, energetic constraints, higher predation rates, or differences in population age structure (Ghiselin 1974; Woolbright 1983; Monnet & Cherry 2002). These theories do not fully explain whole-clade tendencies for smaller male sizes, as is the case here. I propose several plausible explanations for the patterns observed in male *Hyperolius,* including interspecific resource partitioning, division of acoustic space, and reproductive character displacement. African frog communities can contain up to 8 *Hyperolius* species occurring in direct sympatry, often with at least one pair of closely related or sister species. The large energy requirements involved with prolonged mating seasons may be driving niche-partitioning on a fine-scale level in these communities. Males are likely to be fiercely competing with conspecifics, but are also likely experiencing competition for calling sites, food resources, or shelters with similarly sized heterospecifics. The body sizes of males may therefore be the result of a combination of these selective pressures, in addition to female preference. Potential character displacement, whether reproductive or ecological, as well as fine-scale partitioning of body sizes in species-rich communities, can be investigated through targeted ecological and morphological work on these assemblages.

Protogynous sex change and the evolution of body size

The most prominent examples of male-biased SSD occur in the *Hyperolius viridiflavus-marmoratus-parallelus* supergroup (contained in *Hyperolius* clade 2), which is suspected to exhibit protogynous sex change (Grafe & Linsenmair 1989). A high density of females appears to induce this polyphenism during the mating season, and if triggered a female can complete functional sex change and become a reproductively active male (Grafe & Linsenmair 1989). In this system, protogynous sex change is most advantageous if the sexes are of similar sizes, because the acoustic signals of super-sized secondary males may not be recognizable or preferred by their potential mates. In one of the complexes of this supergroup, *H. marmoratus*, it has been demonstrated that females prefer lower frequency calls in small choruses, which are associated with larger male body sizes (Dyson & Passmore 1988; Telford et al. 1988). This empirical evidence

suggests a slightly above-average male size can produce a more attractive call (Dyson & Passmore 1988). The secondary males were found to be above the average size of primary males (Grafe & Linsenmair 1989), which would confer a calling advantage. In addition, larger males are typically more successful in frog combat (Wells 19778; Shine 1979), a behavior that has been documented for these territorial species. If protogynous sex change can be demonstrated in wild populations, this would be a novel explanation for the evolution of SSDi in anurans.

Regardless of the existence of this mating system, is selection promoting larger male sizes or smaller female sizes in this supergroup of *Hyperolius*? The analyses of body size evolution across *Hyperolius* strongly suggest there is a single optimal female body size, whereas multiple optima exist for males (Table 11). This is reflected in the estimated male optimum for *Hyperolius* clade 2, which is larger than for clades 1 and 3 ($\theta_1 = 3.10$, $\theta_2 = 3.37$, $\theta_3 = 3.22$) (Table 12). This is also corroborated by the ancestral reconstructions of body size, which show little change in female size across *Hyperolius*, but there appears to be an overall clade-level increase in body size for males in the supergroup (Fig. 18). Together, this evidence strongly suggests selection is stronger for male body size in the *Hyperolius viridiflavus-marmoratus-parallelus* supergroup. Though the combat and female preference hypotheses can explain selection for this size increase, I postulate this size increase is also likely a consequence of the evolution of a unique mating system involving protogynous sex change. Further work is required to demonstrate whether or not protogynous sex change is commonly occurring in this supergroup, and if selective forces unique to this mating system are truly balancing the body sizes of the sexes.

Evolution of sexual dichromatism

Sexual dichromatism has evolved independently in two main clades of hyperoliids: 1) the common ancestor of the Malagasy and Seychelles Islands species (*Heterixalus*, *Tachycnemis*), and 2) the common ancestor of *Hyperolius, Morerella*, and *Cryptothylax* (Fig. 19). A number of subsequent reversals to monochromatism occur in the genus *Hyperolius* and *Heterixalus*, with some cases of secondarily derived dichromatism. Overall, these patterns were somewhat captured in the study by Veith et al. (2009), and support the multiple origin of sexual dichromatism in this family. There is disparity in the distribution of this trait across the subclades of *Hyperolius*, with considerable variation occurring in *Hyperolius* clade 3, but nearly all species in *Hyperolius* clades 1 and 2 exhibit sexual dichromatism.

Ontogenetic dichromatism in hyperoliids could be driven by a combination of factors related to sexual selection and natural selection, and disentangling these patterns here is not possible. In other vertebrates, selection for dichromatism is often correlated with increasing importance of species recognition (Figuerola & Green, 2000). The role of sexual dichromatism for intraspecific signaling has not been demonstrated, though this remains a possibility. There does not appear to be any relationship between sexual dichromatism and the evolution of body size or sexual size dimorphism, suggesting dichromatism does not enhance or constrain the evolution of these traits. There may be differences in the function of the integument between the color phases within a species, and work focused on the gene expression and chemical content of the skin across phases is currently underway (Portik, Bell, and Blackburn). I have recently demonstrated that

hyperoliids do not sequester alkaloids, which suggests that the vivid coloration often present in one or both sexes does not advertise toxicity (Portik et al. 2015a). The differences in patterns between the sexes may also serve as cryptic coloration functioning in different microhabitats, if habitat partitioning is demonstrated to occur between the sexes within species. This hypothesis requires substantial ecological work, which is generally lacking across hyperoliid species. The evolution and persistence of sexual dichromatism in hyperoliids is puzzling, but the patterns uncovered by this analysis can serve as a starting point to begin generating and investigating hypotheses, and provide directions for future work.

Acknowledgments

I would like to thank Rayna Bell, David Blackburn, Gregor Jongsma, Adam Leaché, Lucinda Lawson, Václav Gvoždík, Mark-Oliver Roedel, Eli Greenbaum, Bob Drewes, Jens Vindum, Stefan Lötters, Michael Veith, Jos Kielgast, Marius Burger, Raffael Ernst, Arie van der Meijden, Susanne Müller, Alan Channing, Simon Loader, James Harvey, Werner Conradie, Bill Branch, Michele Menegon, Chris Barratt, Tim Colston, Caleb Ofori-Boateng, Zoltan Tamas Nagy, Max Dehling, Ulrich Sinsch, Daniela Rössler, Annika Hillers, Mareike Hirschfeld, and Annmarie Ohler for sampling contributions, cooperation, and willingness to participate in a large collaboration to resolve the relationships of Hyperoliidae.

I thank the Museum of Vertebrate Zoology, California Academy of Sciences, Cornell Museum of Natural History, Museum of Comparative Zoology, Burke Museum of National History, Field Museum of Chicago, South African National Biodiversity Institute, South African Institute for Aquatic Biodiversity, Port Elizabeth Museum, Museum National d'Histoire Naturelle, and North Carolina Museum of Natural Sciences for facilitating tissue and/or specimen loans. I specifically thank Carol Spencer for assistance in the Museum of Vertebrate Zoology, and Jens Vindum, Lauren Scheinberg, and Noel Graham and for assistance in the California Academy of Sciences. I thank David Blackburn, Adam Leaché, Gregor Jongsma, Marcel Kouete, Ben Evans, Caleb Ofori-Boateng, Brian Freiermuth, Lauren Scheinberg, Divine Fotibu, Jay McEntee, Matt McElroy, and Elia Mulungu for allowing my participation in or assisting with fieldwork. My lab work was funded by a National Science Foundation DDIG (DEB: 1311006), the EECG Research Award (American Genetic Association), the Museum of Vertebrate Zoology, and by David Blackburn, Rayna Bell, and Jim McGuire. I thank Lydia Smith and Ke Bi for help in the planning and execution of the transcriptome-based exon capture experiment, and Sean Reilly and Ammon Corl for help in lab training and troubleshooting. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303. This work benefitted greatly from my discussions with the McGuire lab, including Jim McGuire, Philip Skipwith, and Sean Reilly, and also David Blackburn, Rayna Bell, Don Miles, David Cannatella, David Wake, Ammon Corl, and Gregor Jongsma.

References

Abouheif, E., and D.J. Fairbairn. 1997. A comparative analysis of allometry for sexual size dimorphism: assessing Rensch's Rule. The American Naturalist 149:540–562.

Amiet, J.-L. 2012. Les Rainettes du Cameroun (Amphibiens Anoures). Saint-Nazaire, France: La Nef des Livres.

Andersson, M. 1994. Sexual selection. Princeton University Press, Princeton, NJ.

Beaulieu J.M., D.C. Jhwueng, C. Boettiger, and B.C. O'Meara. 2012. Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. Evolution 66:2369–2383.

Bell, R.C., and K.R. Zamudio. 2012. Sexual dichromatism in frogs: natural selection, sexual selection and unexpected diversity. Proceedings of the Royal Society B 279:4687–4693.

Bell, R.C., R.C. Drewes, and K.R. Zamudio. 2015a. Reed frog diversification in the Gulf of Guinea: Overseas dispersal, the progression rule, and in situ speciation. Evolution 69:904–915.

Bell, R.C., R.C. Drewes, A. Channing, V. Gvoždík, J. Kielgast, S. Lötters, B.L. Stuart, and K.R. Zamudio. 2015b. Overseas dispersal of *Hyperolius* reed frogs from Central Africa to the oceanic islands of São Tomé and Príncipe. Journal of Biogeography 42:65–75.

Bi, K., D. Vanderpool, S. Singhal, T. Linderoth, C. Moritz, and J.M. Good. 2012. Transcriptome-based exon capture enables highly cost-effective comparative genomic data collection at moderate evolutionary scales. BMC Genomics 13:403.

Blanckenhorn, W.U., R. Meier, and T. Teder. 2007. Rensch's rule in insects: patterns within and among species. In: Fairbairn, D.J., W.U. Blanckenhorn, T. Szekely, eds. Sex, size and gender roles: evolutionary studies of sexual size dimorphism. Oxford: Oxford University Press, 60–70.

Butler, M.A., and A.A. King. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. American Naturalist 164:683–695.

Bwong, B.A., L.P. Lawson, J.N. Nyamache, M. Menegon, C.D. Barratt, D.M. Portik, P. Malonza, P. Nagel, and S.P. Loader. Phylogenetic, ecological and morphological variation in the congeners *Hyperolius mitchelli* and *Hyperolius rubrovermiculatus* from East Africa. Zootaxa, *Submitted.*

Capella-Gutierrez, S., J.M. Silla-Martinez, and T. Gabaldon. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics, 25:1972–1973.

Channing, A. 2001. Amphibians of Central and Southern Africa. Cornell University Press, Ithaca, New York, USA.

Channing, A., and K.M. Howell. 2006. Amphibians of East Africa. Edition Chimaira, Frankfurt, Germany.

Channing, A., A. Hillers, S. Lötters, M.-O. Rödel, S. Schick, W. Conradie, D. Rödder, V. Mercurio, P. Wagner, J.M. Dehling, L.H. du Preez, J. Kielgast, and M. Burger. 2013. Taxonomy of the super-cryptic *Hyperolius nasutus* group of long reed frogs of Africa (Anura: Hyperoliidae), with descriptions of six new species. Zootaxa 3620:301–350.

Conradie, W., W.R. Branch, G.J. Measey, and K.A. Tolley. 2012. A new species of *Hyperolius* Rapp, 1842 (Anura: Hyperoliidae) from the Serra da Chela mountains, south-western Angola. Zootaxa 3269:1–17.
Conradie, W., W.R. Branch, and K.A. Tolley. 2013. Fifty shades of grey: giving colour to the poorly known Angolan Ashy reed frog (Hyperoliidae: *Hyperolius cinereus*), with the description of a new species. Zootaxa 3635:201–223.
Cox RM, Butler MA, John-Alder HB. 2007. The evolution of sexual size dimorphism in reptiles. In: Fairbairn, D.J., W.U. Blanckenhorn, T. Szekely, eds. Sex, size and gender roles: evolutionary studies of sexual size dimorphism. Oxford: Oxford University Press, 38–49.
Crump, M.L. 1974. Reproductive strategies in a tropical community. University of Kansas Museum of Natural History Miscellaneous Publication 61:1–68.
Darwin, C.R. 1874. The descent of man, and selection in relation to sex. 2nd ed, Appleton, New York.
De Lisle, S.P., and L. Rowe. 2013. Correlated evolution of allometry and sexual dimorphism across higher taxa. The American Naturalist 183:630–639.
Dehling, J.M. 2012. An African glass frog: A new *Hyperolius* species (Anura: Hyperoliidae) from Nyungwe National Park, southern Rwanda. Zootaxa 3391:52–64.
Drewes, R.C. 1984. A phylogenetic analysis of the Hyperollidae (Anura): treefrogs of Africa, Madagascar, and the Seychelles Islands. Occasional Papers of the California Academy of Sciences 139:1–70.
Dyson, M.L., and N.I. Passmore. 1988. Two-choice phonotaxis in *Hyperolius marmoratus* (Anura: Hyperoliidae): the effect of temporal variation in presented stimuli. Animal Behavior 36:648–652.
Fairbairn, D.J. 1997. Allometry for sexual size dimorphism: pattern and process in the coevolution of body size in males and females. Annual Review of Ecology and Systematics 28:659–687.
Fairbairn, D.J., W.U. Blanckenhorn, and T. Szekeley. 2007. Sex, size and gender roles: evolutionary studies of sexual size dimorphism. Oxford: Oxford University Press.
Felsenstein, J. 1985. Phylogenies and the comparative method. American Naturalist 125:1–15.
Figuerola, J. and A.J. Green. 2000. The evolution of sexual dimorphism in relation to mating patterns, cavity nesting, insularity, and sympatry in the Anseriformes. Functional Ecology 14:701–710.
Frost, D.R., T. Grant, J. Faivovich, R.H. Bain, A. Haas, C.F.B. Haddad, R.O. de Sa, A. Channing, M. Wilkinson, S.C. Donnellan, C.J. Raxworthy, J.A. Campbell, B.L. Blotto, P. Moler, R.C. Drewes, R.A. Nussbaum, J.D. Lynch, D.M. Green, and W.C. Wheeler. 2006. The amphibian tree of life. Bulletin of the American Museum of Natural History 297:1–370.
Garland, T., A.W. Dickerman, C.M. Janis, and J.A. Jones. 1993. Phylogenetic analysis of covariance by computer simulation. Systematic Biology 42:265–292.
Ghiselin, M.T. 1974. The economy of nature and the evolution of sex. University of California Press, Berkeley.
Grafe, T.U., and K.E. Linsenmair. 1989. Protogynous sex change in the reed frog. Copeia 1989:1024–1029.

Greenbaum, E., U. Sinsch, E. Lehr, F. Valdez, and C. Kusamba. 2013. Phylogeography of the reed frog *Hyperolius castaneus* (Anura: Hyperoliidae) from the Albertine Rift of Central Africa: Implications for taxonomy, biogeography, and conservation. Zootaxa 3731:473–494.

Han, X., and J. Fu. 2013. Does life history shape sexual size dimorphism in anurans? A comparative analysis. BMC Evolutionary Biology 13:27.

Hansen, T.F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

Harmon, L. J., J.T. Weir, C.D. Brock, R.E. Glor, and W. Challenger. 2008. GEIGER: investigating evolutionary radiations. Bioinformatics 24:129–131.

Hayes, T., and K. Menendez. 1999. The effect of sex steroids on primary and secondary sex differentiation in the sexually dichromatic reedfrog (*Hyperolius argus*: Hyperolidae) from the Arabuko Sokoke Forest of Kenya. General and Comparative Endocrinology 115:188–199.

Ives, A.R., P.E. Midford, and T. Garland Jr. 2007. Within-species variation and measurement error in phylogenetic comparative methods. Systematic Biology 56:252–270.

Katoh, K., and D.M. Standley. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution, 30:722–780.

Lawson, L. P. 2010. The discordance of diversification: evolution in the tropical-montane frogs of the Eastern Arc Mountains of Tanzania. Molecular Ecology 19:4046–4060.

Lawson, L.P., and L. Collett. 2011. Herpetofauna of Montane Areas of Tanzania. 1. Results from Two Amphibian Surveys of Malundwe Mountain, Mikumi National Park. Fieldiana Life and Earth Sciences 4:74–80.

Legendre, P. 2014. Model II regression user's guide, R edition. Available at http://cran.r-project.org/web/packages/lmodel2/lmodel2.pdf. Accessed February 24, 2014.

Liedtke, H.C., D. Hügli, J.M. Dehling, F. Pupin, M. Menegon, A.J. Plumptre, D. Kujirakwinja, and S.P. Loader. 2014. One or two species? On the case of *Hyperolius discodactylus* Ahl, 1931 and *H. alticola* Ahl, 1931 (Anura: Hyperoliidae). Zootaxa 3768:253–290.

Liu, L., L. Yu, and S.V. Edwards. 2010. A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. BMC Evolutionary Biology 10:302.

Loader, S.P., L.P. Lawson, D.M. Portik, and M. Menegon. 2015. Three new species of spiny throated reed frogs (Anura: Hyperoliidae) from evergreen forests of Tanzania. BMC Research Notes 8:167.

Lovich, J.E., and J.W. Gibbons. 1992. A review of techniques for quantifying sexual size dimorphism. Growth, Development and Aging 56:269–281.

Meyer, M., and M. Kircher. 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor Protocols 2010(6): pdb.prot5448.

Miller, M.A., W. Pfeiffer, and T. Schwartz. 2010. "Creating the CIPRES Science Gateway for inference of large phylogenetic trees" in Proceedings of the Gateway

Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA pp 1–8.
Mirarab, S., R. Reaz, Md.S. Bayzid, T. Zimmermann, M.S. Swenson, and T. Warnow. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics 30:i541–i548.
Monnet, J.-M., and M.I. Cherry. 2002. Sexual size dimorphism in anurans. Proceedings of the Royal Society B: Biological Sciences 269:2301–2307.
Monroe, M.J., S.H. South, and S.H. Alonzo. 2015. The evolution of fecundity is associated with female body size but not female-biased sexual size dimorphism among frogs. Journal of Evolutionary Biology 28:1793–1803.
Nali, R.C, K.R. Zamudio, C.F.B. Haddad, and C.P.A. Prado. 2014. Size-dependent selective mechanisms on males and females and the evolution of sexual size dimorphism in frogs. The American Naturalist 184:727–740.
Nussbaum, R.A., and S.H. Wu. 1995. Distribution, variation, and systematics of the Seychelles treefrog, *Tachycnemis seychellensis* (Amphiba: Anura: Hyperoliidae). Journal of Zoology, London 236:383–406.
O'Meara, B.C., and J.M. Beaulieu. 2014. Modelling stabilizing selection: the attraction of Ornstein-Uhlenbeck models. In: Modern Comparative Methods and Their Application in Evolutionary Biology (L.Z. Garamszegi, ed), pp. 381–393. Springer, New York, NY.
Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analysis of phylogenetics and evolution in R language. Bioinformatics 20:289–290.
Penner, J., M. Wegmann, A. Hillers, M. Schmidt, and M.-O. Rödel. 2011. A hotspot revisited - a biogeographical analysis of West African amphibians. Molecular Phylogenetics and Evolution 176:1077–1088.
Portik, D.M., L. Scheinberg, D.C. Blackburn, and R.A. Saporito. 2015a. Lack of defensive alkaloids in the integumentary tissue of four brilliantly colored African Reed Frog species (Hyperoliidae: *Hyperolius*). Herpetological Conservation and Biology, *in press*.
Portik, D.M., L.L. Smith, and K. Bi. 2015b. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence. bioRxiv doi: http://dx.doi.org/10.1101/031468.
Poynton, J., and R. Boycott. 1996. Species turnover between Afromontane and eastern African lowland faunas: Patterns shown by amphibians. Journal of Biogeography 23:669–680.
Poynton, J. C., Loader, S. P., Sherratt, E., & Clarke, B. T. 2006. Amphibian Diversity in East African Biodiversity Hotspots: Altitudinal and latitudinal Patterns. Biodiversity and Conservation 16:1103–1118.
Ralls, K. 1976. Mammals in which females are larger than males. The Quarterly Review of Biology 51:245–276.
Ralls, K. 1977. Sexual dimorphism in mammals: avian models and unanswered questions. The American Naturalist 111:917–938.
Rensch, B. 1950. Die Abhängigkeit der relative sexualdifferenz von der Körpergröβe. Bonner Zoologische Beiträge 1:58–69.
Rensch, B. 1960. Evolution above the species level. Columbia University Press, New York.

Revell, L.J. 2010. Phylogenetic signal and linear regression on species data. Methods in Ecology and Evolution 1:319–329.
Revell, L.J. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution 3:217–223.
Revell, L.J. 2013. Two new graphical methods for mapping trait evolution on phylogenies. Methods in Ecology and Evolution 4:754–759.
Richards, C. M. 1982. The alteration of chromatophore expression by sex hormones in the Kenyan reed frog, *Hyperolius viridiflavus*. General and comparative endocrinology 46:59–67.
Richards, C. M., and W.S. Moore. 1996. A phylogeny for the African treefrog family Hyperoliidae based on mitochondrial rDNA. Molecular Phylogenetics and Evolution 5:522–532.
Robinson, D.R., and L.R. Foulds. 1981. Comparison of phylogenetic trees. Mathematical Biosciences 53: 131–147.
Rödel, M.-O., J. Kosuch, M. Veith, and R. Ernst. 2003. First record of the genus *Acanthicalus* Laurent, 1944 from the Upper Guinean Forest, West Africa, with the description of a new species. Journal of Herpetology 37:43–52.
Rödel, M.-O., J. Kosuch, T.U. Grafe, R. Boistel, N.E. Assemian, N.G. Kouamé, B. Tohé, G. Gourène, J.-L. Perret, K. Henle, P. Tafforeau, N. Pollet, and M. Veith. 2009. A new tree-frog genus and species from Ivory Coast, West Africa (Amphibia: Anura: Hyperoliidae). Zootaxa 2044:23–45.
Sanderson, M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* **19**:101–109.
Schick, S., J. Kielgast, D. Rödder, V. Muchai, M. Burger, and S. Lötters. 2010. New species of reed frog from the Cong basin with discussion of paraphyly in Cinnamon-belly reed frogs. Zootaxa 2501:23–36.
Schiøtz, A. 1967. The treefrogs (Rhacophoridae) of West Africa. Spolia Zoologica Musei Hauniensis 25:1–346.
Schiøtz, A. 1999. Treefrogs of Africa. Edition Chimaira, Frankfurt, Germany.
Schluter, D., T. Price, A.Ø. Mooers, and D. Ludwig. 1997. Likelihood of ancestor states in adaptive radiation. Evolution 51:1699–1711.
Shine, R. 1979. Sexual selection and sexual dimorphism in the Amphibia. Copeia 1979:297–306.
Shine, R. 1989. Ecological causes for the evolution of sexual dimorphism: a review of the evidence. The Quarterly Review of Biology 64:419–461.
Singhal, S. 2013. *De novo* transcriptomic analyses for non-model organisms: an evaluation of methods across a multi-species data set. Molecular Ecology Resources 13:403–416.
Smith, R.J. 1999. Statistics of sexual size dimorphism. Journal of Human Evolution 36:423–459.
Smith, R.J. 2009. Use and misuse of the reduced major axis for line-fitting. American Journal of Anthropology 140:476–486.
Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Trivers, R.L. 1972. Parental investment and sexual selection. Pp 136–179, in B. Campbell ed. Sexual selection and the descent of man, 1871–1971. Aldine, Chicago.
Veith, M., J. Kosuch, M.-O. Rödel, A. Hillers, A. Schmitz, M. Burger, and S. Lötters. 2009. Multiple evolution of sexual dichromatism in African reed frogs. Molecular Phylogenetics and Evolution 51:388–393.
Vences, M., J. Kosuch, F. Glaw, W. Böhme, and M. Veith. 2003. Molecular phylogeny of hyperoliid treefrogs: biogeographic origin of Malagasy and Seychellean taxa and re-analysis of familial paraphyly. Canadian Journal of Zoology 41:205–215.
Warton, D.I., I.J. Wright, D.S. Falster, and M. Westoby. 2006. Bivariate line-fitting methods for allometry. Biological Reviews 81:259–291.
Webb T.J., and R.P. Freckleton. 2007. Only half right: species with female-biased sexual size dimorphism consistently break Rensch's rule. PLoS ONE 2: e897.
Wells, K.D. 1977. The social behavior of anuran amphibians. Animal Behavior 25:666–693.
Wieczorek, A., R. Drewes, and A. Channing. 2000. Biogeography and evolutionary history of *Hyperolius* species: application of molecular phylogeny. Journal of Biogeography 27:1231–1243.
Woolbright, L.L. 1983. Sexual selection and size dimorphism in anuran Amphibia. The American Naturalist 121:110–119.

Figure 14. Topological convergence using species tree estimation in ASTRAL across bin sizes, compared to A) full data set ASTRAL species tree, B) optimal topology RAxML tree, and C) other ASTRAL species trees reconstructed in the same bin.

Figure 14

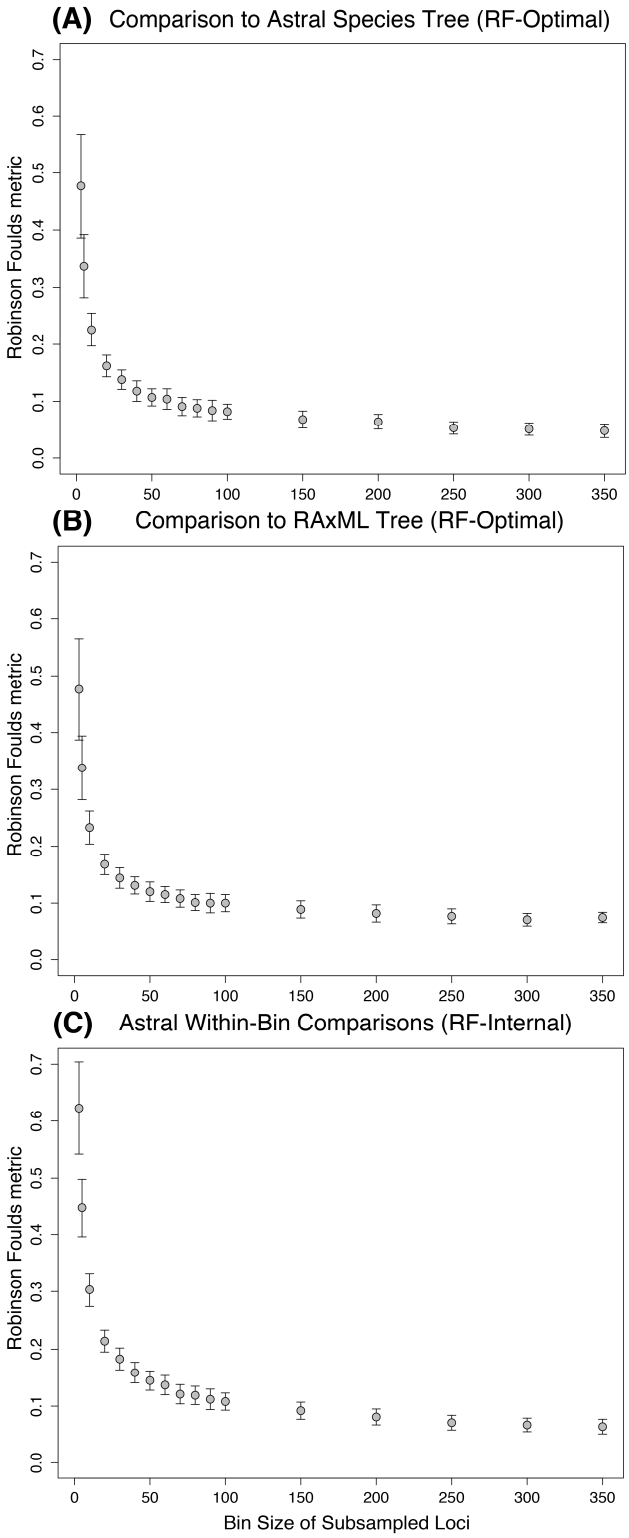

Figure 15. The optimal ML topology of the full 1,047 locus data set returned by RAxML. Shaded nodes have > 90% bootstrap support, white nodes < 90% bootstrap support. Clades used as regimes in evolutionary model testing are labeled.

Figure 15

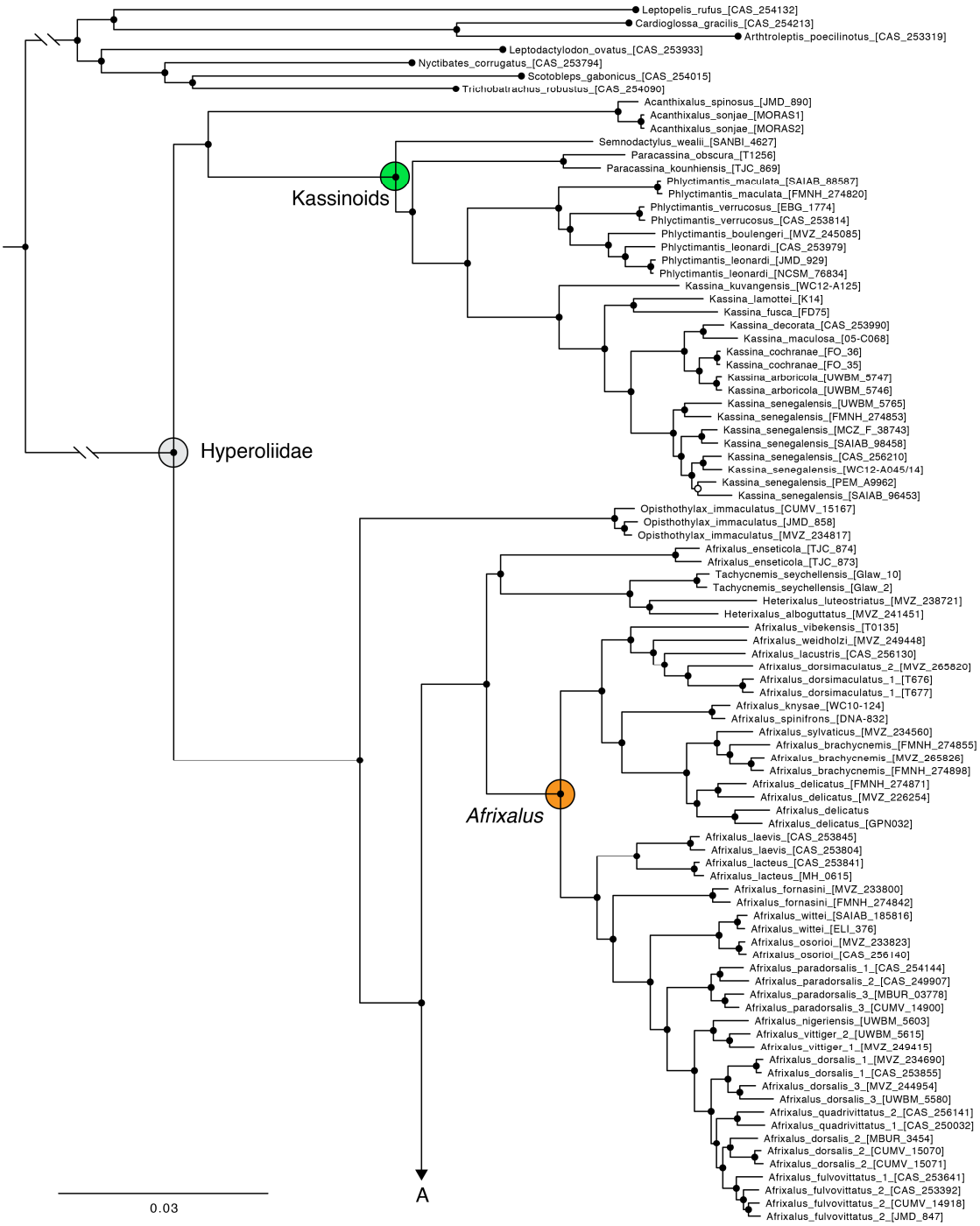

Figure 15 (continued).

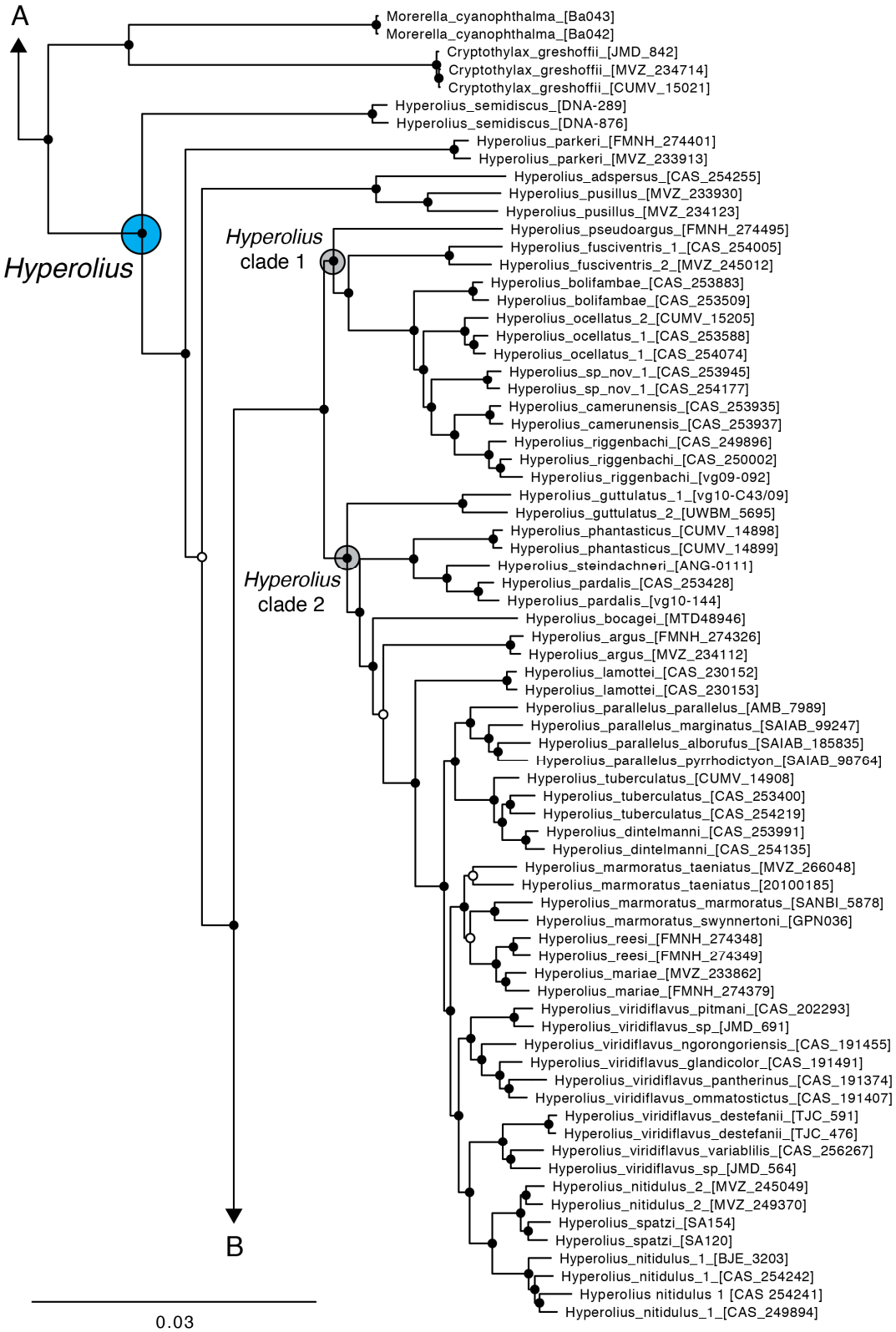

Figure 15 (continued).

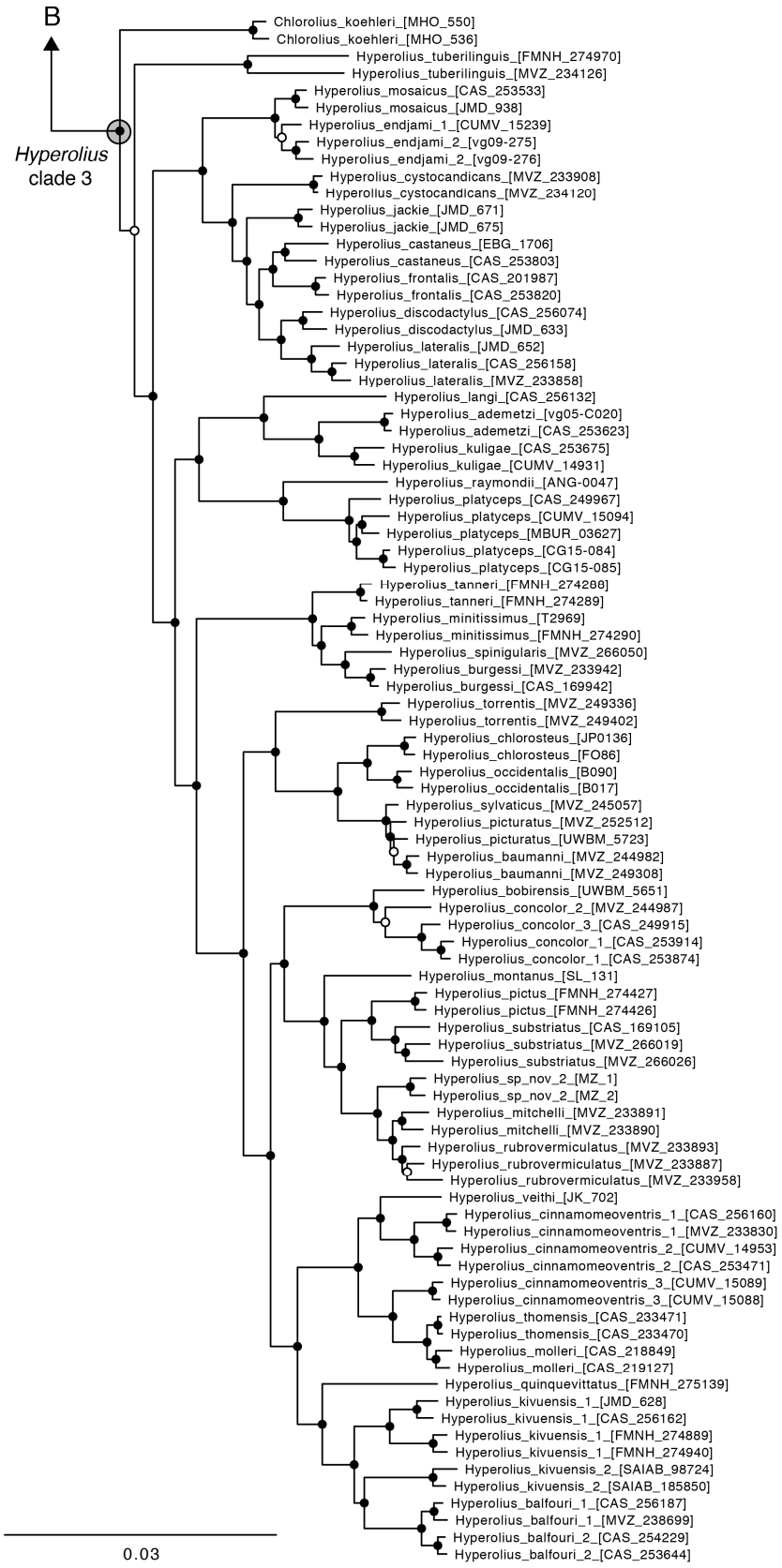

Figure 16. Test of Rensch's rule across Hyperoliidae. Dotted line indicates isometry (β = 1). Solid lines result from phylogenetic reduced major axis regressions: black, all hyperoliids; orange, *Afrixalus*; and light blue, *Hyperolius*.

Figure 16

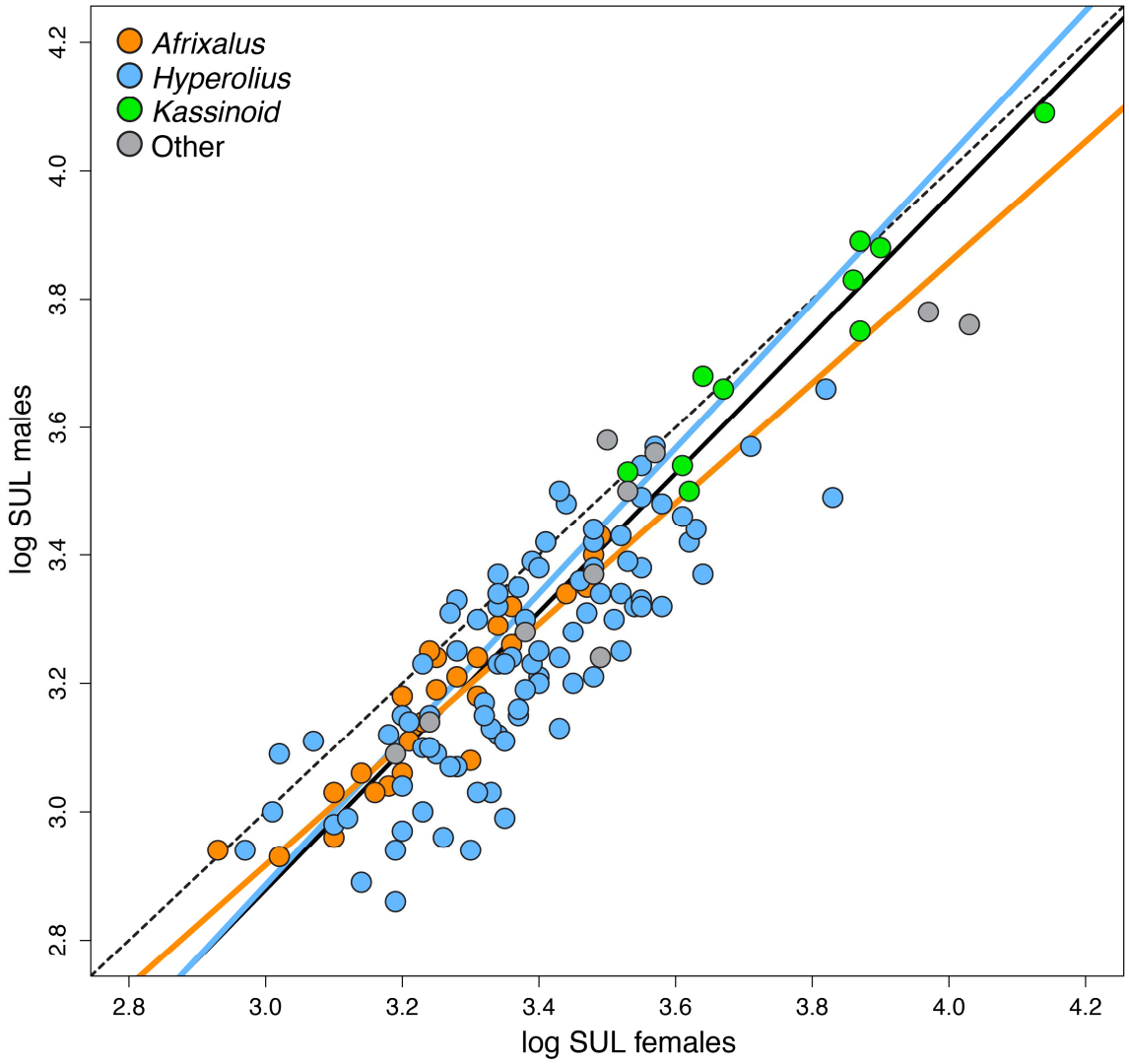

Figure 17. Maximum likelihood reconstruction of sexual size dimorphism index for hyperoliid frogs. Warmer colors represent male-biased SSD or similar sizes of the sexes and cooler colors indicate female-biased SSD.

Figure 17

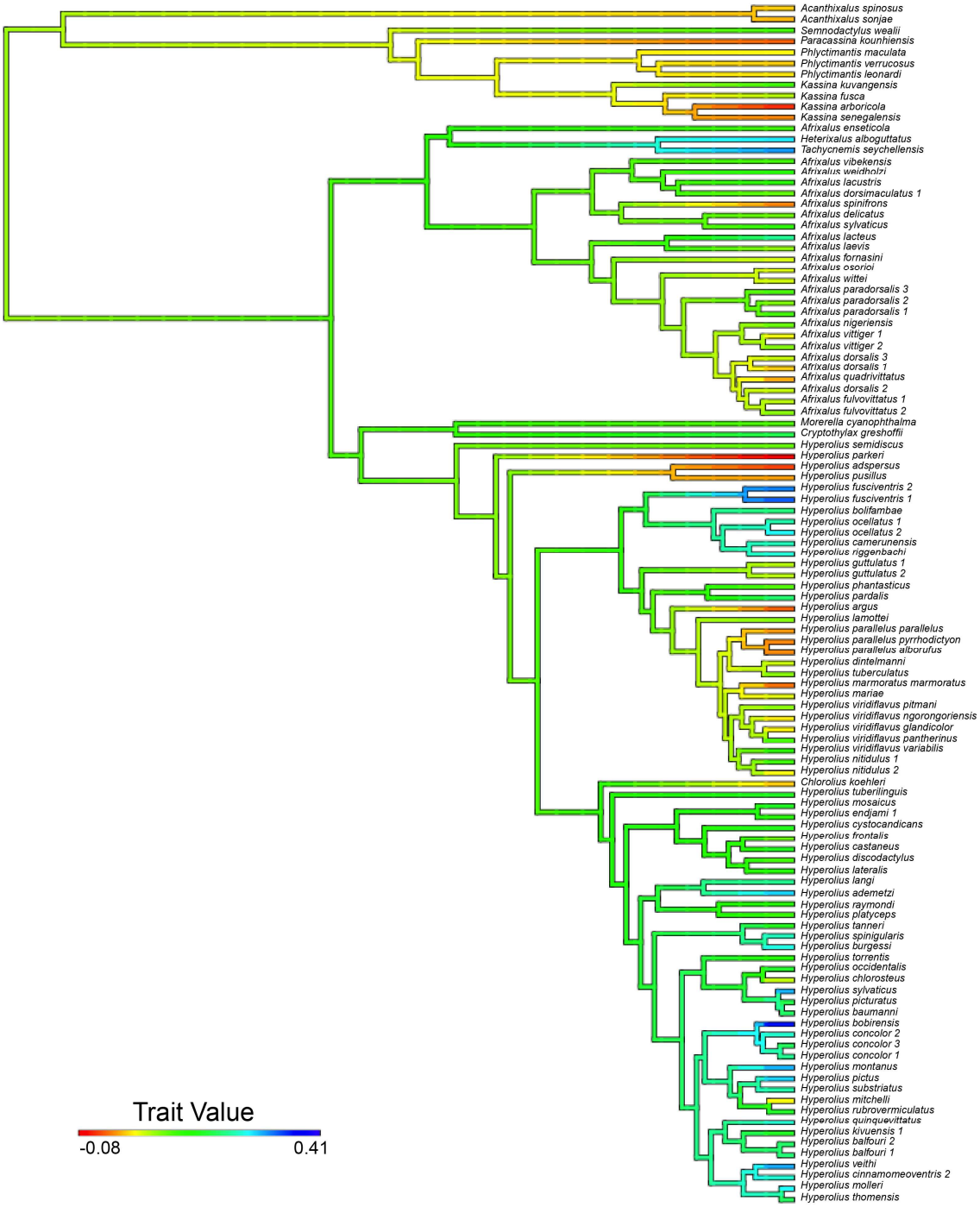

Figure 18. Reconstruction of male and female body sizes for species of hyperoliid frogs. Warmer colors represent smaller body sizes, with cooler colors indicating larger body sizes.

Figure 18

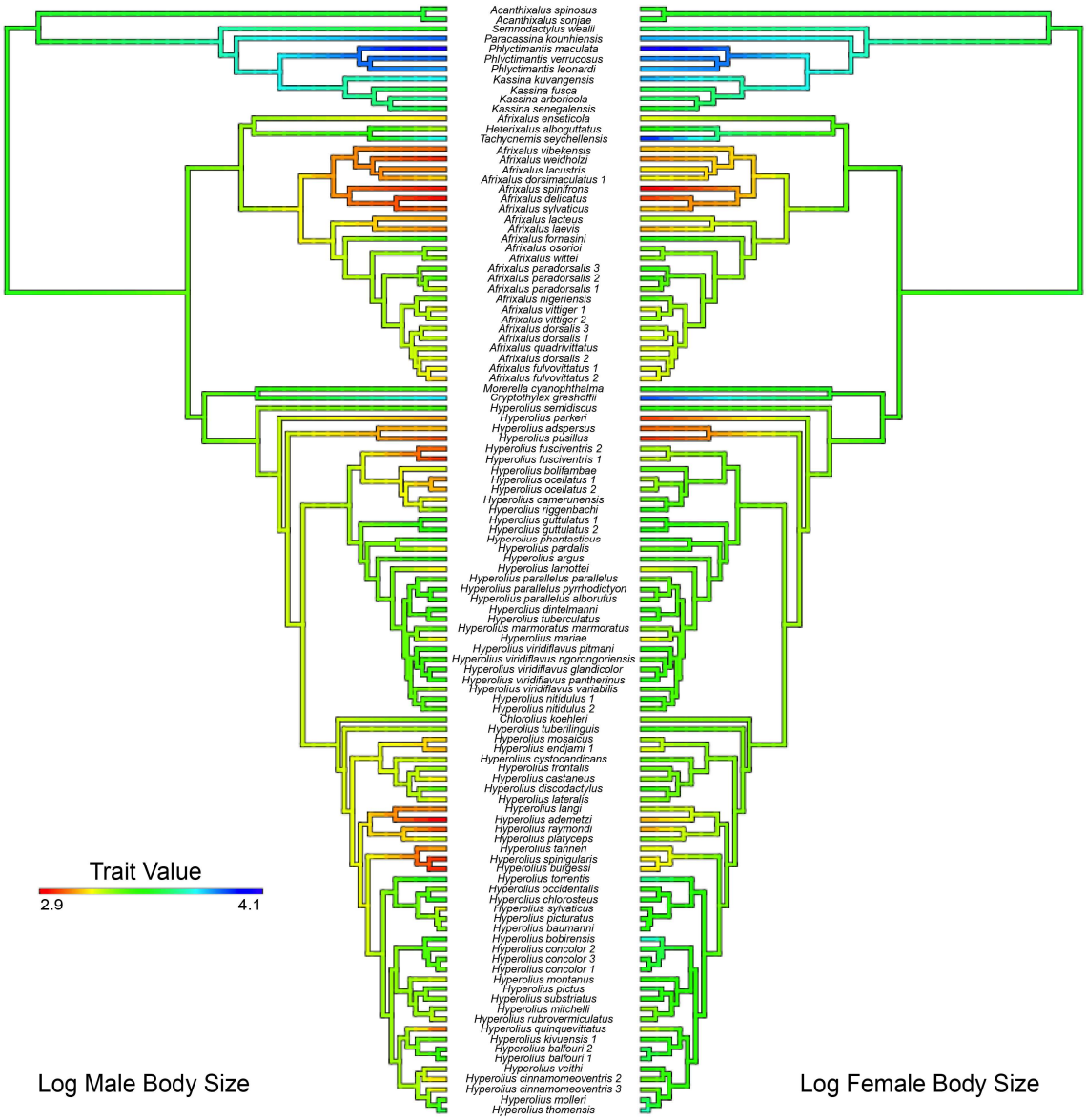

Figure 19. Mapping of sexual dichromatism onto the phylogeny of Hyperoliidae. Boxes represent characters states, with pie charts at nodes representing posterior probabilities of character states. Grey specifies monochromatism and black represents sexual dichromatism. The arrow indicates where the phylogeny continues.

Figure 19

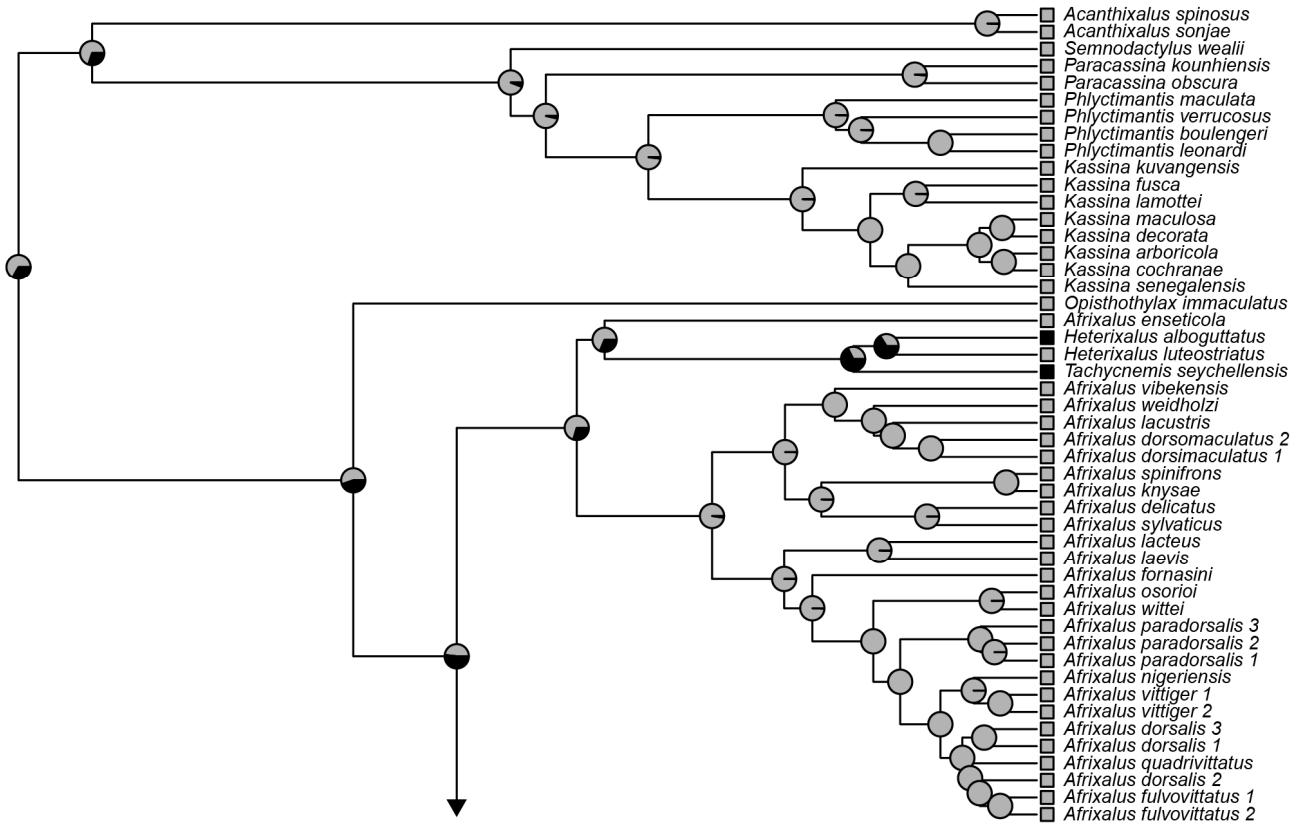

Figure 19 (continued).

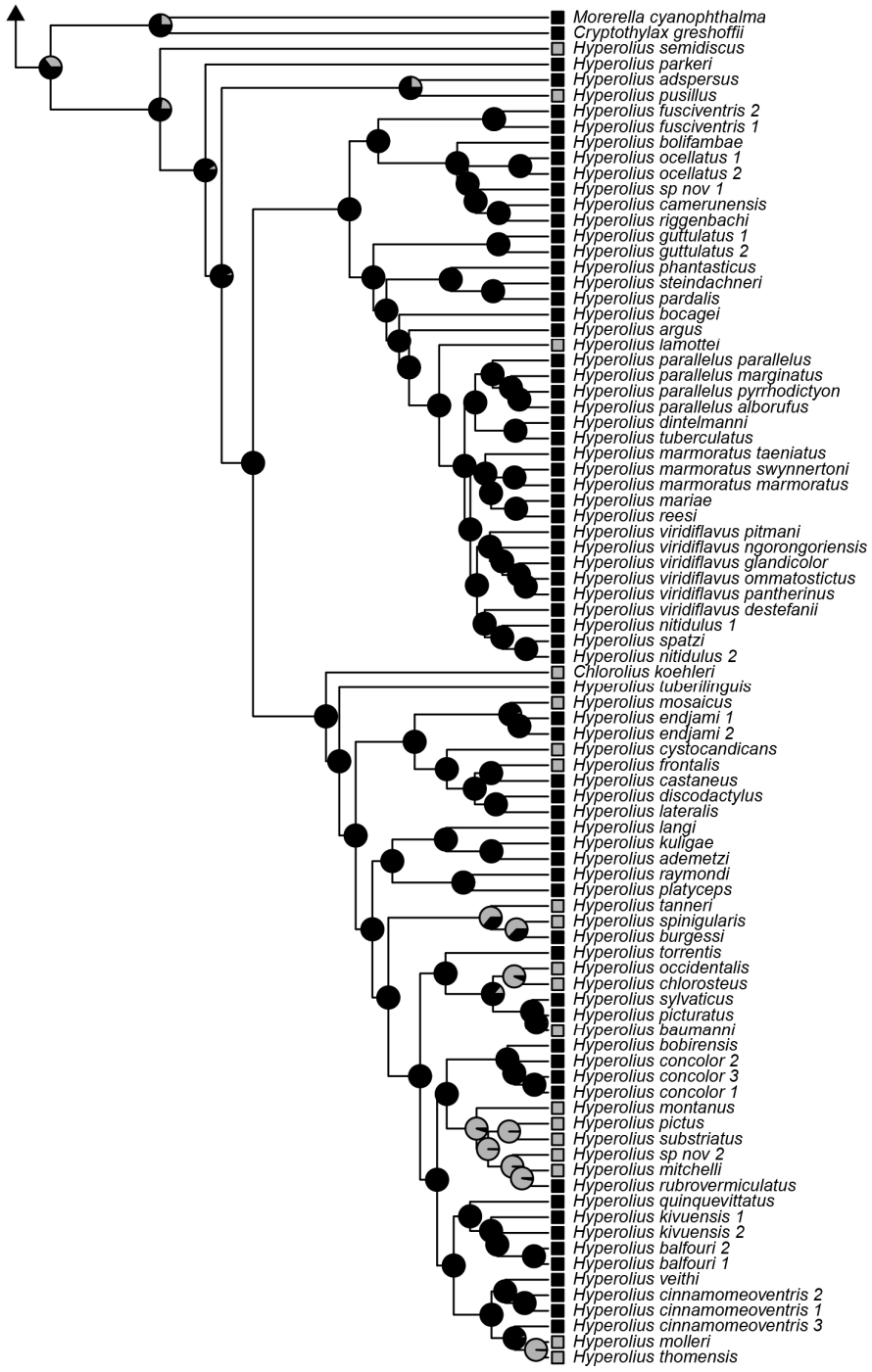

Table 9. Results of non-phylogenetically corrected RMA regressions of body sizes.

Table 9

Group	β	95% CI	Intercept	95% CI	r2
Hyperoliidae	1.03	0.95–1.11	-0.23	-0.52–0.03	0.83
Hyperolius	1.01	0.88–1.17	-0.21	-0.71–0.23	0.67
Afrixalus	1.03	0.87–1.21	-0.18	-0.77–0.32	0.86

Table 10. Results of phylogenetically corrected RMA regressions of body sizes.

Table 10

Group	β	Intercept	r2	T	df	p-value
Hyperoliidae	1.08	-0.36	0.72	1.56	81.2	0.12
Hyperolius	1.13	-0.52	0.69	1.88	52.5	0.06
Afrixalus	0.94	0.09	0.77	0.60	17.8	0.55

Table 11. Fit of evolutionary models to body size evolution of different regimes.

Table 11

Group	Model	AICc	ΔAICc	Weight
Hyperoliidae: Males	**OUM**	**-79.30**	**0.00**	**0.5019**
	OUMV	-77.20	2.10	0.1756
	OUMA	-76.85	2.45	0.1475
	OUMVA	-75.43	3.87	0.0723
	BMS	-74.95	4.36	0.0568
	OU1	-74.26	5.04	0.0403
	BM1	-70.33	8.97	0.0057
Hyperoliidae: Females	OU1	-88.23	0.00	0.3081
	BM1	-87.69	0.54	0.2357
	BMS	-87.46	0.77	0.2098
	OUM	-86.99	1.24	0.1659
	OUMV	-84.22	4.00	0.0416
	OUMA	-83.70	4.53	0.0320
	OUMVA	-80.63	7.60	0.0069
Hyperolius: Males	**OUMV**	**-59.49**	**0.00**	**0.6619**
	OUM	-57.35	2.14	0.2270
	OUMVA	-54.41	5.09	0.0520
	OU1	-54.27	5.23	0.0485
	BMS	-50.99	8.50	0.0094
	BM1	-46.65	12.85	0.0011
Hyperolius: Females	**OU1**	**-72.82**	**0.00**	**0.7081**
	OUMV	-69.18	3.64	0.1148
	OUM	-69.07	3.75	0.1085
	OUMVA	-65.64	7.17	0.0196
	BM1	-65.50	7.32	0.0182
	BMS	-65.38	7.44	0.0172
	OUMA	-64.92	7.90	0.0136

Best models are in bold, and were chosen based on AICc weights.

Table 12. Parameter estimates of best models for the evolution of body size of each sex.

Table 12

Group	Sex	Model	α	σ^2_1	σ^2_2	σ^2_3	θ_1	θ_2	θ_3
Hyperoliidae	Males	OUM	3.88	0.31	—	—	3.71	3.18	3.28
	Females	OU1	1.02	0.19	—	—	3.55	—	—
Hyperolius	Males	OUMV	12.04	0.27	0.28	0.76	3.10	3.37	3.22
	Females	OU1	8.30	0.36	—	—	3.39	—	—

Models were chosen based on AICc weights. Empty cells indicate the model involved a single estimate for the σ^2 or θ parameter across regimes, otherwise values for each selective regime (1–3) are provided.

Table 13. Results of comparison of sexual size dimorphism index and sexual dichromatism using phylogenetic analysis of variance (pANOVA).

Table 13

	phylANOVA		aov.phylog		
	F value	p-value	*F* value	Uncorrected p-value	Corrected p-value
SSDi1	10.52	0.42	10.5	< 0.01	0.42
Male body size	0.55	0.86	0.54	0.46	0.86
Female body size	0.25	0.90	0.25	0.61	0.90

Appendix 3**.** Specimen information for samples included in phylogenomic work.

Lineage	**Tissue ID**	**Country/Island**	**Locality**
Acanthixalus sonjae	MORAS1		
Acanthixalus sonjae	MORAS2		
Acanthixalus spinosus	JMD 890	Gabon	Parc National de l'Ivindo
Afrixalus "brachycnemis"	FMNH 274898	Malawi	-16.01366, 35.65175
Afrixalus "brachycnemis"	FMNH 274855	Malawi	-12.11386, 33.72457
Afrixalus "brachycnemis"	MVZ 265826	Mozambique	-15.384275, 37.071958
Afrixalus "delicatus"	MVZ 226254	Kenya	-3.33333, 39.86667
Afrixalus "delicatus"	GPN032		
Afrixalus "delicatus"	FMNH 274871	Malawi	-16.04831, 35.71076
Afrixalus "delicatus"			
Afrixalus dorsalis 1	CAS 253855	Cameroon	4.95539, 9.86786
Afrixalus dorsalis 1	MVZ 234690	Ghana	6.27555, -1.0081
Afrixalus dorsalis 2	CUMV 15070	Gabon	01.80976°S, 09.38770°E
Afrixalus dorsalis 2	CUMV 15071	Gabon	01.80976°S, 09.38770°E
Afrixalus dorsalis 2	MBUR 3454	Republic of Congo	
Afrixalus dorsalis 3	MVZ 244954	Ghana	5.29004, -2.63957
Afrixalus dorsalis 3	UWBM 5580	Ghana	6.72604 N, 01.52750 W
Afrixalus dorsimaculatus 1	T676	Tanzania	West Usambaras
Afrixalus dorsimaculatus 1	T677	Tanzania	West Usambaras
Afrixalus dorsimaculatus 2	MVZ 265820	Mozambique	-12.86737656, 35.17335137
Afrixalus enseticola	TJC 873	Ethiopia	Oromia, Bale
Afrixalus enseticola	TJC 874	Ethiopia	Oromia, Bale
Afrixalus fornasini	FMNH 274842	Malawi	-11.98449, 34.04609
Afrixalus fornasini	MVZ 233800	Kenya	-3.333, 39.867
Afrixalus fulvovittatus 1	CAS 253641	Cameroon	05□56'07", 10□07'41"

Afrixalus fulvovittatus 2	CUMV 14918	Gabon	00.51121°N, 12.80278°E
Afrixalus fulvovittatus 2	CAS 253392	Cameroon	2.59461, 14.02827
Afrixalus fulvovittatus 2	JMD 847	Gabon	Parc National de l'Ivindo
Afrixalus knysae	WC10-124	South Africa	33.95022222, 23.6015
Afrixalus lacteus	CAS 253841	Cameroon	4º 57' 33.98" N, 9º 51' 54.22" E
Afrixalus lacteus	MH 0615	Cameroon	
Afrixalus lacustris	CAS 256130	Uganda	0.4508 N, 32.9493 E
Afrixalus laevis	CAS 253804	Cameroon	4.32228, 9.06635
Afrixalus laevis	CAS 253845	Cameroon	4º 57' 33.98" N, 9º 51' 54.22" E
Afrixalus nigeriensis	UWBM 5603	Ghana	6.24246 N, 0.55710 W
Afrixalus osorioi	CAS 256140	Uganda	0.43226 N, 32.96107 E
Afrixalus osorioi	MVZ 233823	Kenya	0.2663333, 34.80675
Afrixalus paradorsalis 1	CAS 254144	Cameroon	4□57.636, 9□39.222
Afrixalus paradorsalis 2	CAS 249907	Cameroon	5°43.493' N, 9°20.963' E
Afrixalus paradorsalis 3	CUMV 14900	Gabon	00.45358°N, 10.27809°E
Afrixalus paradorsalis 3	MBUR 03778	Republic of Congo	02 25 14.9 S, 12 52 20.1 E
Afrixalus quadrivittatus 1	CAS 250032	Cameroon	6°10.517' N, 10°04.474' E
Afrixalus quadrivittatus 2	CAS 256141	Uganda	0.43226 N, 32.96107 E
Afrixalus spinifrons	DNA-832	South Africa	31.65386, 29.50683
Afrixalus sylvaticus	MVZ 234560	Kenya	-4.2664167, 39.3405
Afrixalus vibekensis	T01.35		
Afrixalus vittiger 1	MVZ 249415	Ghana	8.42942, 0.56544
Afrixalus vittiger 2	UWBM 5615	Ghana	10.42018 N, 2.07402 W
Afrixalus weidholzi	MVZ 249448	Ghana	8.42942, 0.56544
Afrixalus wittei	ELI 376	DRC	S7.3382, E28.0088
Afrixalus wittei	SAIAB 185816	Zambia	-12.25386111, 25.34419444
Arthtroleptis poecilinotus	CAS 253319	Cameroon	2.90072, 13.90334
Breviceps adspersus	MVZ 265908	Mozambique	-21.996, 35.323

Cardioglossa gracilis	CAS 254213	Cameroon	5.15301, 10.53249
Chlorolius koehleri	MHO 536	Cameroon	
Chlorolius koehleri	MHO 550	Cameroon	
Cryptothylax greshoffii	JMD 842	Gabon	Parc National de l'Ivindo
Cryptothylax greshoffii	MVZ 234714	Cameroon	2.96625, 12.0816
Cryptothylax greshoffii	CUMV 15021	Gabon	00.49986°N, 12.80182°E
Hemisus marmoratus	MVZ 249304	Ghana	8.28022, 0.53053
Heterixalus alboguttatus	MVZ 241451	Madagascar	-21.25, 47.45
Heterixalus luteostriatus	MVZ 238721	Madagascar	-21.89067, 46.76333
Hyperolius ademetzi	CAS 253623	Cameroon	05□51.344', 10□09.512'
Hyperolius ademetzi	vg05-C020	Cameroon	6.2038, 10.4596
Hyperolius adspersus	CAS 254255	Cameroon	5.10863, 10.61671
Hyperolius argus	FMNH 274326	Tanzania	S 04.82645°, E 038.78822°
Hyperolius argus	MVZ 234112	Kenya	-3.1999167, 40.0077
Hyperolius balfouri 1	CAS 256187	Uganda	0.44423 N, 32.97309 E
Hyperolius balfouri 1	MVZ 238699	Kenya	0.042222222, 34.74694444
Hyperolius balfouri 2	CAS 253644	Cameroon	05□56'07", 10□07'41"
Hyperolius balfouri 2	CAS 254229	Cameroon	5.15301, 10.53249
Hyperolius baumanni	MVZ 244982	Ghana	8.3226, 0.55622
Hyperolius baumanni	MVZ 249308	Ghana	8.40584, 0.59387
Hyperolius bobirensis	UWBM 5651	Ghana	6.24246 N, 0.55710 W
Hyperolius bocagei	MTD48946	Angola	
Hyperolius bolifambae	CAS 253509	Cameroon	3.17376, 12.52714
Hyperolius bolifambae	CAS 253883	Cameroon	4.84975, 9.77183
Hyperolius burgessi	CAS 169942	Tanzania	5 5 60 S, 38 37 60 W
Hyperolius burgessi	MVZ 233942	Tanzania	-5.1000333, 38.6266333
Hyperolius camerunensis	CAS 253935	Cameroon	4.84975, 9.77183
Hyperolius camerunensis	CAS 253937	Cameroon	4.84975, 9.77183

Hyperolius castaneus	EBG 1706	DRC	S3.36255, E28.87713
Hyperolius castaneus	CAS 253803	Uganda	1.05315 S, 29.76138 E
Hyperolius chlorosteus	FO86		
Hyperolius chlorosteus	JP0136		
Hyperolius cinnamomeoventris 1	MVZ 233830	Kenya	0.2663333, 34.80675
Hyperolius cinnamomeoventris 1	CAS 256160	Uganda	0.43226 N, 32.96107 E
Hyperolius cinnamomeoventris 2	CUMV 14953	Gabon	00.51121°N, 12.80278°E
Hyperolius cinnamomeoventris 2	CAS 253471	Cameroon	2.93994, 11.97633
Hyperolius cinnamomeoventris 3	CUMV 15088	Gabon	01.80976°S, 09.38770°E
Hyperolius cinnamomeoventris 3	CUMV 15089	Gabon	01.80976°S, 09.38770°E
Hyperolius concolor 1	CAS 253874	Cameroon	4.95539, 9.86786
Hyperolius concolor 1	CAS 253914	Cameroon	4.84975, 9.77183
Hyperolius concolor 2	MVZ 244987	Ghana	8.3226, 0.55621
Hyperolius concolor 3	CAS 249915	Cameroon	5°43.493' N, 9°20.963' E
Hyperolius cystocandicans	MVZ 233908	Kenya	-1.1241667, 36.67375
Hyperolius cystocandicans	MVZ 234120	Kenya	-1.2446833, 36.65705
Hyperolius dintelmanni	CAS 254135	Cameroon	4□57.636, 9□39.222
Hyperolius dintelmanni	CAS 253991	Cameroon	4.84975, 9.77183
Hyperolius discodactylus	JMD 633	Rwanda	Nyungwe
Hyperolius discodactylus	CAS 256074	Uganda	1º 4' 26.4" S, 29º 45' 57.6" E
Hyperolius endjami 1	CUMV 15239	Bioko Island	3.4673166667, 8.6411
Hyperolius endjami 2	vg09-275	Cameroon	4.96344, 9.67492
Hyperolius endjami 2	vg09-276	Cameroon	4.96344, 9.67492
Hyperolius frontalis	CAS 201987	Uganda	0 58 42 S, 29 41 42 E
Hyperolius frontalis	CAS 253820	Uganda	1.04903 S, 29.79025 E
Hyperolius fusciventris 1	CAS 254005	Cameroon	4.84975, 9.77183
Hyperolius fusciventris 2	MVZ 245012	Ghana	5.28434, -2.64861
Hyperolius guttulatus 1	vg10-C43/09	Cameroon	4.96413333333333, 8.90711666666667

Hyperolius guttulatus 2	UWBM 5695	Ghana	6.83673 N, 1.72202 W
Hyperolius jackie	JMD 671	Rwanda	Nyungwe
Hyperolius jackie	JMD 675	Rwanda	Nyungwe
Hyperolius kivuensis 1	FMNH 274889	Malawi	-14.24392, 35.39216
Hyperolius kivuensis 1	FMNH 274940	Malawi	-14.24392, 35.39216
Hyperolius kivuensis 1	JMD 628	Rwanda	Rukarara
Hyperolius kivuensis 1	CAS 256162	Uganda	0.43226 N, 32.96107 E
Hyperolius kivuensis 2	SAIAB 185850	Zambia	-12.25386111, 25.34419444
Hyperolius kivuensis 2	SAIAB 98724	Zambia	-9.362555556, 30.12366667
Hyperolius kuligae	CUMV 14931	Gabon	00.51121°N, 12.80278°E
Hyperolius kuligae	CAS 253675	Cameroon	05□55'34", 10□07'32"
Hyperolius lamottei	CAS 230152	Sierra Leone	8 5 24 N, 12 26 51.6 W
Hyperolius lamottei	CAS 230153	Sierra Leone	08 05.40 N, 12 26.86 W
Hyperolius langi	CAS 256132	Uganda	0.4508 N, 32.9493 E
Hyperolius lateralis	JMD 652	Rwanda	Butare
Hyperolius lateralis	MVZ 233858	Kenya	0.2663333, 34.80675
Hyperolius lateralis	CAS 256158	Uganda	0.43226 N, 32.96107 E
Hyperolius mariae	FMNH 274379	Tanzania	S 04.82645°, E 038.78822°
Hyperolius mariae	MVZ 233862	Tanzania	-6.6067583, 38.8202167
Hyperolius marmoratus marmoratus	SANBI 5878	South Africa	-29.9396667, 30.9935837
Hyperolius marmoratus swynnertoni	GPN036		
Hyperolius marmoratus taeniatus	2010.0185	Mozambique	
Hyperolius marmoratus taeniatus	MVZ 266048	Mozambique	-15.384275, 37.071958
Hyperolius minitissimus	T2969	Tanzania	Kihanga
Hyperolius minitissimus	FMNH 274290	Tanzania	S 08.3295°, E 035.939983°
Hyperolius mitchelli	MVZ 233890	Tanzania	-6.6067583, 38.8202167
Hyperolius mitchelli	MVZ 233891	Tanzania	-6.6067583, 38.8202167
Hyperolius molleri	CAS 218849	Sao Tome and Principe	0 17 52.8 N, 6 43 49.3 E

Hyperolius molleri	CAS 219127	Sao Tome and Principe	1 39 7.7 N, 7 24 57.8 E
Hyperolius montanus	SL 131	Kenya	
Hyperolius mosaicus	CAS 253533	Cameroon	3.17916, 12.53352
Hyperolius mosaicus	JMD 938	Gabon	Parc National des Monts de Cristal
Hyperolius nitidulus 1	BJE 3203	Cameroon	6°32.917' N, 10°45.599' E
Hyperolius nitidulus 1	CAS 249894	Cameroon	5°58.063' N, 10°02.382' E
Hyperolius nitidulus 1	CAS 254241	Cameroon	5.15301, 10.53249
Hyperolius nitidulus 1	CAS 254242	Cameroon	5.15301, 10.53249
Hyperolius nitidulus 2	MVZ 245049	Ghana	5.88026, 0.03782
Hyperolius nitidulus 2	MVZ 249370	Ghana	9.25988, -1.86001
Hyperolius occidentalis	B017	Sierra Leone	
Hyperolius occidentalis	B090	Sierra Leone	
Hyperolius ocellatus 1	CAS 253588	Cameroon	3.17334, 12.52902
Hyperolius ocellatus 1	CAS 254074	Cameroon	4º 50' 59.1" N, 9º 46' 18.5874" E
Hyperolius ocellatus 2	CUMV 15205	Bioko Island	03 28.039, 08 38.466
Hyperolius parallelus alborufus	SAIAB 185835	Zambia	-12.18691667, 25.21452778
Hyperolius parallelus marginatus	SAIAB 99247	Zambia	-12.28622222, 30.6215
Hyperolius parallelus parallelus	AMB 7989	Namibia	18 02 04.7S, 20 51 41.0 E
Hyperolius parallelus pyrrhodictyon	SAIAB 98764	Zambia	-15.60325, 28.26433333
Hyperolius pardalis	CAS 253428	Cameroon	2.59461, 14.02827
Hyperolius pardalis	vg10-144	Cameroon	2.44217, 15.43078
Hyperolius parkeri	MVZ 233913	Kenya	04 15.042 S, 39 20.654 E
Hyperolius parkeri	FMNH 274401	Tanzania	S 05.1133° E, E 038.75257°
Hyperolius phantasticus	CUMV 14898	Gabon	00.45358°N, 10.27809°E
Hyperolius phantasticus	CUMV 14899	Gabon	00.45358°N, 10.27809°E
Hyperolius picturatus	UWBM 5723	Ghana	6.24246 N, 0.55710 W
Hyperolius picturatus	MVZ 252512	Ghana	6.27555, -1.0081
Hyperolius pictus	FMNH 274426	Tanzania	S 08.09955°, E 035.89848°

Hyperolius pictus	FMNH 274427	Tanzania	S 08.09955°, E 035.89848°
Hyperolius platyceps	CG15-084	DRC	-3.2923, 16.1623
Hyperolius platyceps	CG15-085	DRC	-3.2923, 16.1623
Hyperolius platyceps	CUMV 15094	Gabon	01.77104°S, 09.36941°E
Hyperolius platyceps	MBUR 03627	Republic of Congo	02 12 38.3 S, 12 49 56.6 E
Hyperolius platyceps	CAS 249967	Cameroon	4°36.693' N, 12°13.525' E
Hyperolius pseudoargus	FMNH 274495	Tanzania	S 08.51464°, E 035.4540583°
Hyperolius pusillus	MVZ 233930	Tanzania	Dar es Salaam Region
Hyperolius pusillus	MVZ 234123	Kenya	-3.15845, 39.07105
Hyperolius quinquevittatus	FMNH 275139	Malawi	-11.66565, 33.7613
Hyperolius raymondii	ANG-0047	Angola	7.744222222, 19.95466667
Hyperolius reesi	FMNH 274348	Tanzania	S 07.8421° , E 034.87814°
Hyperolius reesi	FMNH 274349	Tanzania	S 07.8421° , E 034.87814°
Hyperolius riggenbachi	CAS 250002	Cameroon	6°32.917' N, 10°45.599' E
Hyperolius riggenbachi	CAS 249896	Cameroon	6°00.855' N, 10°16.216' E
Hyperolius riggenbachi	vg09-092	Cameroon	7.74049, 12.72305
Hyperolius rubrovermiculatus	MVZ 233887	Tanzania	05 6.002 S, 38 37.598 E
Hyperolius rubrovermiculatus	MVZ 233893	Kenya	Coast Province, Arabuko Sokoke Forest
Hyperolius rubrovermiculatus	MVZ 233958	Kenya	-4.2664167, 39.3405
Hyperolius semidiscus	DNA-289	South Africa	32.28392, 28.84885
Hyperolius semidiscus	DNA-876	South Africa	31.6501, 29.50513
Hyperolius sp. nov. 1	CAS 253945	Cameroon	4.84975, 9.77183
Hyperolius sp. nov. 1	CAS 254177	Cameroon	5.17626, 10.34873
Hyperolius sp. nov. 2	MZ_1	Mozambique	Cabo Delgabo Province, Palama
Hyperolius sp. nov. 2	MZ_2	Mozambique	Cabo Delgabo Province, Palama
Hyperolius spatzi	SA120		
Hyperolius spatzi	SA154		
Hyperolius spinigularis	MVZ 266050	Mozambique	-15.384275, 37.071958

Hyperolius steindachneri	ANG-0111	Angola	8.338472222, 20.2425
Hyperolius substriatus	CAS 169105	Tanzania	4 47 60 S, 38 30 0 W
Hyperolius substriatus	MVZ 266019	Mozambique	-12.84388577, 35.17817779
Hyperolius substriatus	MVZ 266026	Mozambique	-15.384275, 37.071958
Hyperolius sylvaticus	MVZ 245057	Ghana	5.28434, -2.64861
Hyperolius tanneri	FMNH 274288	Tanzania	S 04.829116° , E 038.51273°
Hyperolius tanneri	FMNH 274289	Tanzania	S 04.74783°, E 038.29743°
Hyperolius thomensis	CAS 233470	Sao Tome and Principe	0 16 34 N, 6 36 20 E
Hyperolius thomensis	CAS 233471	Sao Tome and Principe	0 16 34 N, 6 36 20 E
Hyperolius torrentis	MVZ 249336	Ghana	8.33383, 0.59449
Hyperolius torrentis	MVZ 249402	Ghana	8.43424, 0.60552
Hyperolius tuberculatus	CUMV 14908	Gabon	00.62120°N, 10.40763°E
Hyperolius tuberculatus	CAS 253400	Cameroon	2.59461, 14.02827
Hyperolius tuberculatus	CAS 254219	Cameroon	5.15301, 10.53249
Hyperolius tuberilinguis	MVZ 234126	Kenya	-3.1699667, 39.8628333
Hyperolius tuberilinguis	FMNH 274970	Malawi	-16.04831, 35.71076
Hyperolius veithi	JK 702		
Hyperolius viridiflavus destefanii	TJC 476	Ethiopia	7.09845, 38.4164
Hyperolius viridiflavus destefanii	TJC 591	Ethiopia	6.58604, 38.38871
Hyperolius viridiflavus glandicolor	CAS 191491	Kenya	3 30 41.4 S, 38 15 7.8 E
Hyperolius viridiflavus ngorongoriensis	CAS 191455	Tanzania	3 22 24 S, 35 51 18 E
Hyperolius viridiflavus ommatostictus	CAS 191407	Tanzania	3 21 33 S, 36 50 14.4 E
Hyperolius viridiflavus pantherinus	CAS 191374	Kenya	0 9 18 S, 37 0 36 E
Hyperolius viridiflavus pitmani	CAS 202293	Uganda	0 59 12.9 S, 29 38 3.5 E
Hyperolius viridiflavus sp.	JMD 564	Rwanda	Lac Muhazi
Hyperolius viridiflavus sp.	JMD 691	Rwanda	Marais Rugesi
Hyperolius viridiflavus variablilis	CAS 256267	Uganda	0.4293 N, 33.1958 E
Kassina arboricola	UWBM 5746	Ghana	6.23899 N, 0.55824 W

Kassina arboricola	UWBM 5747	Ghana	6.23899 N, 0.55824 W
Kassina cochranae	FO 35	Guinea	08 31.861' N, 08 54.361' W
Kassina cochranae	FO 36	Guinea	08 31.861' N, 08 54.361' W
Kassina decorata	CAS 253990	Cameroon	4.84975, 9.77183
Kassina fusca	FD75	Guinea	11 46.073 N, 12 40.844' W
Kassina kuvangensis	WC12-A125	Angola	-13.71616667, 17.09661111
Kassina lamottei	K14	Cote d'Ivoire	Tai NP
Kassina maculosa	05-C068	Cameroon	6.13263, 10.17813
Kassina senegalensis	UWBM 5765	Ghana	8.29326 N, 2.28347 W
Kassina senegalensis	FMNH 274853	Malawi	-16.04831, 35.71076
Kassina senegalensis	MCZ F 38743	South Africa	22° 42' 23" S, 29° 49' 44" E
Kassina senegalensis	PEM A9962	South Africa	31.82296 S, 29.30773 E
Kassina senegalensis	SAIAB 96453	Malawi	-16.02456, 35.49902
Kassina senegalensis	SAIAB 98458	Zambia	-11.83111111, 31.42580556
Kassina senegalensis	WC12-A045/14	Angola	-14.67458333, 17.73544444
Kassina senegalensis	CAS 256210	Uganda	0.44423 N, 32.97309 E
Leptodactylodon ovatus	CAS 253933	Cameroon	4.84975, 9.77183
Leptopelis rufus	CAS 254132	Cameroon	4.9606, 9.6686
Morerella cyanophthalma	Ba04.2	Cote d'Ivoire	
Morerella cyanophthalma	Ba04.3	Cote d'Ivoire	
Nyctibates corrugatus	CAS 253794	Cameroon	4.32676, 9.06609
Opisthothylax immaculatus	CUMV 15167	Bioko Island	03 27.181, 08 30.410
Opisthothylax immaculatus	JMD 858	Gabon	Parc National de l'Ivindo
Opisthothylax immaculatus	MVZ 234817	Cameroon	Sud Prov., Sangmelima
Paracassina kounhiensis	TJC 869	Ethiopia	Oromia, Bale
Paracassina obscura	T1256		
Phlyctimantis boulengeri	MVZ 245085	Ghana	5.29004, -2.63957
Phlyctimantis leonardi	CAS 253979	Cameroon	4.84975, 9.77183

Phlyctimantis leonardi	JMD 929	Gabon	Parc National des Monts de Cristal
Phlyctimantis leonardi	NCSM 76834	Gabon	0.45358 N, 10.27809 E
Phlyctimantis maculata	FMNH 274820	Malawi	-16.01366, 35.65175
Phlyctimantis maculata	SAIAB 88587	Mozambique	-15.05994444, 40.67088889
Phlyctimantis verrucosus	CAS 253814	Uganda	1.04903 S, 29.79025 E
Phlyctimantis verrucosus	EBG 1774	DRC	N00.68335, E29.67143
Phrynomantis microps	MVZ 249480	Ghana	9.259, -1.85416
Scotobleps gabonicus	CAS 254015	Cameroon	4.84975, 9.77183
Semnodactylus wealii	SANBI 4627	South Africa	
Tachycnemis seychellensis	Glaw 10	Seychelles Islands	Mahe
Tachycnemis seychellensis	Glaw 2	Seychelles Islands	Praslin
Trichobatrachus robustus	CAS 254090	Cameroon	4.94781, 9.69563

Appendix 4**.** Summary of body size measurements and sexual size dimorphism indices for all species examined.

Genus	Species	Male SUL	Female SUL	log Male SUL	log Female SUL	SSDi 1	SSDi 2
Acanthixalus	*sonjae*	35.10	35.50	3.56	3.57	0.01	0.99
Acanthixalus	*spinosus*	33.32	34.08	3.50	3.53	0.02	0.98
Afrixalus	*delicatus*	18.65	20.47	2.93	3.02	0.10	0.90
Afrixalus	*dorsalis 1*	25.64	25.93	3.24	3.25	0.01	0.99
Afrixalus	*dorsalis 2*	24.84	26.41	3.21	3.28	0.06	0.94
Afrixalus	*dorsalis 3*	25.54	27.27	3.24	3.31	0.07	0.93
Afrixalus	*dorsimaculatus*	23.07	25.45	3.14	3.23	0.10	0.90
Afrixalus	*enseticola*	23.00	25.82	3.14	3.24	0.12	0.88
Afrixalus	*fornasini*	31.00	32.64	3.43	3.49	0.05	0.95
Afrixalus	*fulvovittatus 1*	24.41	25.91	3.19	3.25	0.06	0.94
Afrixalus	*fulvovittatus 2*	22.41	24.70	3.11	3.21	0.10	0.90
Afrixalus	*lacteus*	21.86	27.16	3.08	3.30	0.24	0.76
Afrixalus	*lacustris*	21.37	24.46	3.06	3.20	0.15	0.86
Afrixalus	*laevis*	21.38	23.06	3.06	3.14	0.08	0.92
Afrixalus	*nigeriensis*	28.19	31.29	3.34	3.44	0.11	0.89
Afrixalus	*orophilus*	20.72	22.08	3.03	3.10	0.07	0.93
Afrixalus	*osorioi*	27.58	28.96	3.32	3.36	0.05	0.95
Afrixalus	*paradorsalis 1*	24.16	27.40	3.18	3.31	0.13	0.87
Afrixalus	*paradorsalis 2*	29.95	32.35	3.40	3.48	0.08	0.92
Afrixalus	*paradorsalis 3*	28.52	32.12	3.35	3.47	0.13	0.87
Afrixalus	*quadrivittatus*	25.78	25.66	3.25	3.24	-0.01	1.01
Afrixalus	*spinifrons*	19.04	18.80	2.94	2.93	-0.01	1.01
Afrixalus	*sylvaticus*	21.00	24.00	3.04	3.18	0.14	0.86
Afrixalus	*vibekensis*	20.90	23.50	3.03	3.16	0.12	0.88
Afrixalus	*vittiger 1*	24.06	24.62	3.18	3.20	0.02	0.98
Afrixalus	*vittiger 2*	26.12	28.91	3.26	3.36	0.11	0.89
Afrixalus	*weidholzi*	19.30	22.30	2.96	3.10	0.16	0.85
Afrixalus	*wittei*	26.77	28.30	3.29	3.34	0.06	0.94
Callixalus	*pictus*	35.71	33.02	3.58	3.50	-0.08	1.08
Chlorolius	*koehleri*	27.70	28.10	3.32	3.34	0.01	0.99
Chrysobatrachus	*cupreonitens*	19.83	28.49	2.99	3.35	0.44	0.56
Cryptothylax	*greshoffii*	44.27	53.15	3.78	3.97	0.20	0.80
Heterixalus	*alboguttatus*	25.64	32.81	3.24	3.49	0.28	0.72
Heterixalus	*betsileo*	22.02	24.24	3.09	3.19	0.10	0.90
Heterixalus	*boettgeri*	26.66	29.40	3.28	3.38	0.10	0.90
Hyperolius	*ademetzi*	17.98	23.28	2.89	3.14	0.30	0.71

Hyperolius	*adspersus*	22.39	21.54	3.11	3.07	-0.04	1.04
Hyperolius	*argus*	32.38	31.37	3.48	3.44	-0.03	1.03
Hyperolius	*balfouri 1*	30.64	37.42	3.42	3.62	0.22	0.78
Hyperolius	*balfouri 2*	31.31	37.82	3.44	3.63	0.21	0.79
Hyperolius	*baumanni*	28.21	33.88	3.34	3.52	0.20	0.80
Hyperolius	*benguellensis*	19.68	22.19	2.98	3.10	0.13	0.87
Hyperolius	*bobirensis*	32.81	46.20	3.49	3.83	0.41	0.59
Hyperolius	*bolifambae*	24.70	30.14	3.21	3.40	0.22	0.78
Hyperolius	*burgessi*	18.99	24.26	2.94	3.19	0.28	0.72
Hyperolius	*camerunensis*	23.41	29.19	3.15	3.37	0.25	0.75
Hyperolius	*castaneus*	23.33	27.65	3.15	3.32	0.19	0.82
Hyperolius	*chelaensis*	24.60	31.50	3.20	3.45	0.28	0.72
Hyperolius	*chlorosteus*	35.30	35.40	3.57	3.57	0.00	1.00
Hyperolius	*chrysogaster*	20.68	27.98	3.03	3.33	0.35	0.65
Hyperolius	*cinereus*	21.00	24.60	3.04	3.20	0.17	0.83
Hyperolius	*cinnamomeoventris 1*	22.69	28.38	3.12	3.34	0.25	0.75
Hyperolius	*cinnamomeoventris 2*	27.83	34.36	3.32	3.54	0.24	0.77
Hyperolius	*concolor 1*	28.07	35.01	3.33	3.55	0.25	0.75
Hyperolius	*concolor 2*	29.29	34.84	3.38	3.55	0.19	0.81
Hyperolius	*concolor 3*	25.40	29.79	3.23	3.39	0.17	0.83
Hyperolius	*cystocandicans*	18.82	26.95	2.94	3.30	0.43	0.57
Hyperolius	*davenporti*	31.03	32.65	3.44	3.48	0.05	0.95
Hyperolius	*dintelmanni*	29.05	32.48	3.37	3.48	0.12	0.88
Hyperolius	*discodactylus*	22.10	25.20	3.10	3.23	0.14	0.86
Hyperolius	*endjami*	27.14	29.48	3.30	3.38	0.09	0.91
Hyperolius	*frontalis*	19.34	25.96	2.96	3.26	0.34	0.66
Hyperolius	*fusciventris 1*	20.85	27.54	3.03	3.31	0.32	0.68
Hyperolius	*fusciventris 2*	32.74	34.83	3.49	3.55	0.06	0.94
Hyperolius	*guttulatus 1*	21.98	25.75	3.09	3.25	0.17	0.83
Hyperolius	*guttulatus 3*	30.78	32.79	3.42	3.48	0.07	0.94
Hyperolius	*horstockii*	34.66	34.91	3.54	3.55	0.01	0.99
Hyperolius	*kachalolae*	25.43	28.27	3.23	3.34	0.11	0.89
Hyperolius	*kivuensis*	28.17	32.79	3.34	3.49	0.16	0.84
Hyperolius	*kuligae*	17.54	24.24	2.86	3.19	0.38	0.62
Hyperolius	*lamottei*	23.35	25.47	3.15	3.24	0.09	0.91
Hyperolius	*langi*	21.56	26.51	3.07	3.28	0.23	0.77
Hyperolius	*lateralis*	23.75	27.59	3.17	3.32	0.16	0.84
Hyperolius	*leucotaenius*	22.92	27.82	3.13	3.33	0.21	0.79
Hyperolius	*mariae*	23.30	24.55	3.15	3.20	0.05	0.95
Hyperolius	*mariae*	25.32	25.25	3.23	3.23	0.00	1.00
Hyperolius	*marmoratus argentovittis*	33.10	30.81	3.50	3.43	-0.07	1.07

Hyperolius	*marmoratus marmoratus*	29.17	28.28	3.37	3.34	-0.03	1.03
Hyperolius	*marmoratus verrucosus*	27.82	26.48	3.33	3.28	-0.05	1.05
Hyperolius	*mitchelli*	27.43	26.23	3.31	3.27	-0.05	1.05
Hyperolius	*molleri*	27.76	36.09	3.32	3.58	0.30	0.70
Hyperolius	*montanus*	24.78	32.61	3.21	3.48	0.32	0.68
Hyperolius	*mosaicus*	23.32	27.66	3.15	3.32	0.19	0.81
Hyperolius	*nasutus*	19.06	19.61	2.94	2.97	0.03	0.97
Hyperolius	*nitidulus 1*	29.58	33.98	3.39	3.53	0.15	0.85
Hyperolius	*nitidulus 2*	28.65	29.23	3.35	3.37	0.02	0.98
Hyperolius	*occidentalis*	25.53	31.04	3.24	3.43	0.22	0.78
Hyperolius	*ocellatus 1*	23.13	24.80	3.14	3.21	0.07	0.93
Hyperolius	*ocellatus 2*	22.30	28.71	3.11	3.35	0.29	0.71
Hyperolius	*ocellatus 3*	21.48	26.38	3.07	3.27	0.23	0.77
Hyperolius	*parallelus alborufus*	30.45	30.22	3.42	3.41	-0.01	1.01
Hyperolius	*parallelus parallelus*	28.43	28.32	3.34	3.34	0.00	1.00
Hyperolius	*parallelus pyrrhodictyon*	30.48	30.16	3.42	3.41	-0.01	1.01
Hyperolius	*pardalis*	23.71	29.10	3.16	3.37	0.23	0.77
Hyperolius	*parkeri*	21.99	20.47	3.09	3.02	-0.07	1.07
Hyperolius	*phantasticus*	32.36	35.85	3.48	3.58	0.11	0.89
Hyperolius	*picturatus*	27.36	32.21	3.31	3.47	0.18	0.82
Hyperolius	*pictus*	29.00	38.00	3.37	3.64	0.31	0.69
Hyperolius	*platyceps*	25.50	28.69	3.24	3.36	0.13	0.88
Hyperolius	*pusillus*	20.17	20.25	3.00	3.01	0.00	1.00
Hyperolius	*quinquevittatus*	20.12	25.38	3.00	3.23	0.26	0.74
Hyperolius	*raymondi*	19.80	22.70	2.99	3.12	0.15	0.85
Hyperolius	*riggenbachi*	27.68	34.96	3.32	3.55	0.26	0.74
Hyperolius	*rubrovermiculatus*	24.47	30.18	3.20	3.40	0.23	0.77
Hyperolius	*semidiscus*	29.42	32.50	3.38	3.48	0.11	0.90
Hyperolius	*spinigularis*	19.59	24.50	2.97	3.20	0.25	0.75
Hyperolius	*substriatus*	27.21	33.63	3.30	3.51	0.24	0.76
Hyperolius	*sylvaticus 1*	22.87	30.90	3.13	3.43	0.35	0.65
Hyperolius	*sylvaticus 2*	24.44	29.54	3.19	3.38	0.21	0.79
Hyperolius	*sylvaticus 3*	25.16	28.48	3.23	3.35	0.13	0.87
Hyperolius	*tanneri*	22.22	25.66	3.10	3.24	0.16	0.85
Hyperolius	*thomensis*	38.78	45.58	3.66	3.82	0.18	0.83
Hyperolius	*torrentis*	35.66	40.82	3.57	3.71	0.15	0.86
Hyperolius	*tuberculatus*	28.86	31.79	3.36	3.46	0.10	0.90
Hyperolius	*tuberilinguis*	26.70	31.43	3.28	3.45	0.18	0.82
Hyperolius	*veithi*	25.70	33.90	3.25	3.52	0.32	0.68
Hyperolius	*viridiflavus ferniquei*	31.13	32.57	3.44	3.48	0.05	0.95
Hyperolius	*viridiflavus glandicolor*	29.74	29.79	3.39	3.39	0.00	1.00

Hyperolius	*viridiflavus goetzei*	25.78	26.59	3.25	3.28	0.03	0.97
Hyperolius	*viridiflavus ngorongoriensis*	29.55	30.11	3.38	3.40	0.02	0.98
Hyperolius	*viridiflavus pantherinus*	31.88	37.14	3.46	3.61	0.17	0.84
Hyperolius	*viridiflavus pitmani*	31.01	33.85	3.43	3.52	0.09	0.91
Hyperolius	*viridiflavus rubripes*	22.71	24.00	3.12	3.18	0.06	0.94
Hyperolius	*viridiflavus variabilis*	25.83	29.97	3.25	3.40	0.16	0.84
Hyperolius	*viridiflavus viridiflavus*	27.03	27.47	3.30	3.31	0.02	0.98
Kassina	*arboricola*	39.85	37.94	3.68	3.64	-0.05	1.05
Kassina	*fusca*	34.36	36.88	3.54	3.61	0.07	0.93
Kassina	*kuvangensis*	42.58	48.10	3.75	3.87	0.13	0.87
Kassina	*maculifer*	38.71	39.18	3.66	3.67	0.01	0.99
Kassina	*senegalensis*	34.24	34.01	3.53	3.53	-0.01	1.01
Morerella	*cyanophthalma*	29.00	32.30	3.37	3.48	0.11	0.89
Paracassina	*kounhiensis*	49.00	47.90	3.89	3.87	-0.02	1.02
Phlyctimantis	*leonardi*	46.16	47.64	3.83	3.86	0.03	0.97
Phlyctimantis	*maculata*	60.22	62.58	4.09	4.14	0.04	0.96
Phlyctimantis	*verrucosus*	48.54	49.59	3.88	3.90	0.02	0.98
Semnodactylus	*wealii*	33.16	37.26	3.50	3.62	0.12	0.88
Tachycnemis	*seychellensis*	43.18	56.66	3.76	4.03	0.31	0.69

Appendix 5**.** Summary of sexual dichromatism character data for lineages included in the phylogenetic analysis.

Genus	Species	Sexual Dichromatism
Acanthixalus	*sonjae*	0
Acanthixalus	*spinosus*	0
Afrixalus	*delicatus*	0
Afrixalus	*dorsalis 1*	0
Afrixalus	*dorsalis 2*	0
Afrixalus	*dorsalis 3*	0
Afrixalus	*dorsimaculatus 1*	0
Afrixalus	*dorsomaculatus 2*	0
Afrixalus	*enseticola*	0
Afrixalus	*fornasini*	0
Afrixalus	*fulvovittatus 1*	0
Afrixalus	*fulvovittatus 2*	0
Afrixalus	*knysae*	0
Afrixalus	*lacteus*	0
Afrixalus	*lacustris*	0
Afrixalus	*laevis*	0
Afrixalus	*nigeriensis*	0
Afrixalus	*osorioi*	0
Afrixalus	*paradorsalis 1*	0
Afrixalus	*paradorsalis 2*	0
Afrixalus	*paradorsalis 3*	0
Afrixalus	*pygmaeus*	0
Afrixalus	*quadrivittatus*	0
Afrixalus	*spinifrons*	0
Afrixalus	*sylvaticus*	0
Afrixalus	*vibekensis*	0
Afrixalus	*vittiger 1*	0
Afrixalus	*vittiger 2*	0
Afrixalus	*weidholzi*	0
Afrixalus	*wittei*	0
Callixalus	*pictus*	1
Chlorolius	*koehleri*	0
Cryptothylax	*greshoffii*	1
Heterixalus	*alboguttatus*	1
Heterixalus	*luteostriatus*	0
Hyperolius	*ademetzi*	1
Hyperolius	*adspersus*	1

Hyperolius	*argus*	1
Hyperolius	*balfouri 1*	1
Hyperolius	*balfouri 2*	1
Hyperolius	*baumanni*	0
Hyperolius	*bobirensis*	1
Hyperolius	*bocagei*	1
Hyperolius	*bolifambae*	1
Hyperolius	*burgessi*	1
Hyperolius	*camerunensis*	1
Hyperolius	*castaneus*	1
Hyperolius	*chlorosteus*	0
Hyperolius	*cinnamomeoventris 1*	1
Hyperolius	*cinnamomeoventris 2*	1
Hyperolius	*cinnamomeoventris 3*	1
Hyperolius	*concolor 1*	1
Hyperolius	*concolor 2*	1
Hyperolius	*concolor 3*	1
Hyperolius	*cystocandicans*	0
Hyperolius	*dintelmanni*	1
Hyperolius	*discodactylus*	1
Hyperolius	*endjami 1*	1
Hyperolius	*endjami 2*	1
Hyperolius	*frontalis*	0
Hyperolius	*fusciventris 1*	1
Hyperolius	*fusciventris 2*	1
Hyperolius	*guttulatus 1*	1
Hyperolius	*guttulatus 2*	1
Hyperolius	*kivuensis 1*	1
Hyperolius	*kivuensis 2*	1
Hyperolius	*kuligae*	1
Hyperolius	*lamottei*	0
Hyperolius	*langi*	1
Hyperolius	*lateralis*	1
Hyperolius	*mariae*	1
Hyperolius	*marmoratus marmoratus*	1
Hyperolius	*marmoratus swynnertoni*	1
Hyperolius	*marmoratus taeniatus*	1
Hyperolius	*mitchelli*	0
Hyperolius	*molleri*	0
Hyperolius	*montanus*	0
Hyperolius	*mosaicus*	0

Hyperolius	*nitidulus 1*	1
Hyperolius	*nitidulus 2*	1
Hyperolius	*occidentalis*	0
Hyperolius	*ocellatus 1*	1
Hyperolius	*ocellatus 2*	1
Hyperolius	*parallelus alborufus*	1
Hyperolius	*parallelus marginatus*	1
Hyperolius	*parallelus parallelus*	1
Hyperolius	*parallelus pyrrhodictyon*	1
Hyperolius	*pardalis*	1
Hyperolius	*parkeri*	1
Hyperolius	*phantasticus*	1
Hyperolius	*picturatus*	1
Hyperolius	*pictus*	0
Hyperolius	*platyceps*	1
Hyperolius	*pusillus*	0
Hyperolius	*quinquevittatus*	1
Hyperolius	*raymondi*	1
Hyperolius	*reesi*	1
Hyperolius	*riggenbachi*	1
Hyperolius	*rubrovermiculatus*	1
Hyperolius	*semidiscus*	0
Hyperolius	*sp 1*	1
Hyperolius	*sp 2*	0
Hyperolius	*spatzi*	1
Hyperolius	*spinigularis*	0
Hyperolius	*steindachneri*	1
Hyperolius	*substriatus*	0
Hyperolius	*sylvaticus*	1
Hyperolius	*tanneri*	0
Hyperolius	*thomensis*	0
Hyperolius	*torrentis*	1
Hyperolius	*tuberculatus*	1
Hyperolius	*tuberilinguis*	1
Hyperolius	*veithi*	1
Hyperolius	*viridiflavus destefanii*	1
Hyperolius	*viridiflavus glandicolor*	1
Hyperolius	*viridiflavus ngorongoriensis*	1
Hyperolius	*viridiflavus ommatostictus*	1
Hyperolius	*viridiflavus pantherinus*	1
Hyperolius	*viridiflavus pitmani*	1

Hyperolius	*viridiflavus variablilis*	1
Kassina	*arboricola*	0
Kassina	*cochranae*	0
Kassina	*decorata*	0
Kassina	*fusca*	0
Kassina	*kuvangensis*	0
Kassina	*lamottei*	0
Kassina	*maculosa*	0
Kassina	*senegalensis*	0
Morerella	*cyanophthalma*	0
Opisthothylax	*immaculatus*	0
Paracassina	*kounhiensis*	0
Paracassina	*obscura*	0
Phlyctimantis	*boulengeri*	0
Phlyctimantis	*leonardi*	0
Phlyctimantis	*maculata*	0
Phlyctimantis	*verrucosus*	0
Semnodactylus	*wealii*	0
Tachycnemis	*seychellensis*	1

Character scoring: 1 indicates presence of sexual dichromatism, 0 indicates absence.

Chapter 4

Monitors cross the Red Sea: The biogeographic history of *Varanus yemenensis*

This chapter was published prior to the filing of my dissertation as follows:
Portik, D.M., and T.J. Papenfuss. 2012. Monitors cross the Red Sea: The biogeographic history of *Varanus yemenensis*. Molecular Phylogenetics and Evolution 62:561–565.

Abstract
The Red Sea has had a profound biogeographic effect on organisms with Afro-Asian distributions, resulting in complex patterns of admixture on the Arabian Peninsula. We investigate the phylogenetic affinities of a monitor lizard (*Varanus yemenensis*) restricted to the southwestern Arabian Peninsula by sequencing all African monitor species and several Asian monitor species for the mitochondrial gene *ND2* and the nuclear marker *RAG-1*. We find evidence that *V. yemenensis* is of African origin, being most closely related to the white-throat monitor, *V. albigularis*, an African species complex distributed from the Horn of Africa to southern Africa. Using divergence-dating analyses, we investigate several biogeographic hypotheses to infer the likely mechanism of colonization of the Arabian Peninsula by this species. Our results reveal that both dispersal across a southern landbridge and overwater dispersal are potential explanations. The patterns observed in *V. yemenensis* are contrasted with other taxa having similar Afro-Arabian disjunct distributions to better understand the complex biogeographic history of this region.

Introduction
Faunal exchanges between Africa and Asia, driven by climatic oscillations and the availability of land bridge connections, have resulted in patterns of admixture across Northern Africa, the Arabian Peninsula, and throughout much of the Middle East (Amer and Kumazawa 2005; Fujita and Papenfuss 2011; Macey et al. 2008; Newman et al. 2004; Papenfuss et al. 2009; Pook et al. 2009; Vargas-Ramìrez et al. 2010; Winney et al. 2004; Wong et al. 2010; Zinner et al. 2009). The Arabian Peninsula and the Red Sea have a complex geologic history (Bohannon 1986; Bosworth et al. 2005), and varying levels of connectivity with Africa have allowed for scattered vicariance events and uneven periods of potential dispersal opportunities (Fernandes et al. 2006). While many birds, mammals, and reptiles maintain large ranges throughout Northern Africa, much of the Arabian Peninsula, and Asian deserts, other species that are found widely across Africa occur only in the southwestern portion of the Arabian Peninsula. Several mammalian and reptilian taxa and species complexes with ranges extending around the Horn of Africa display a disjunct distribution between Yemen and southwestern Saudi Arabia (Newman et al. 2004; Pook et al. 2009; Vargas-Ramìrez et al. 2010; Winney et al. 2004; Wong et al. 2010; Zinner et al. 2009). These congruent patterns point to the importance of the Red Sea in sculpting biogeographic patterns, however this region is largely understudied in this context.

The geology of the Red Sea is well understood, and its formation is comprised of several major events (Fig. 20). Initial rifting and formation of the southern Red Sea

occurred in the early Miocene (~27 mya), and partial embayment occurred as a result of the connection of the Eritrean Red Sea to the Gulf of Aden through the Strait of Bab el Mandeb (Bohannon 1986; Bosworth et al. 2005). Subsequently, northward extension of the Red Sea was rapid, and largely complete by ~23 mya, with the connection between the northern Red Sea and the Neotethys Sea (now Mediterranean Sea) established. Approximately 14 mya, coinciding with the collision between the Arabian plate and Eurasia, a reduction in marine flow occurred in the northern Red Sea, however marine deposits indicate flow between the southern Red Sea and the Gulf and Aden was largely maintained (Bosworth et al. 2005). Massive halite deposition across the entire Red Sea occurred 10 mya, indicating a significant drop in embayment levels and the potential closing of the Strait of Bab el Mandeb. This pattern was maintained until around 5.3 mya, when Pliocene marine sediment deposits formed, indicating a marine connection was reestablished through Bab el Mandeb (Bosworth et al. 2005). When sea level was lowest (10–5.3 mya) the Strait of Bab el Mandeb was approximately 5 km wide, whereas currently it is approximately 30 km wide. While some authors have claimed the existence of temporary land bridges throughout the Pleistocene related to glacial cycles (Rohling et al. 1998), a recent study suggests the existence of these land bridges is unlikely (Fernandes et al. 2006).

The dynamic geology of the formation of the Red Sea leads to several hypotheses that can explain the disjunct southwest Arabian distribution of many African taxa: 1) vicariance coinciding with the initial formation of the Red Sea (27–24 mya), 2) dispersal across a northern land bridge followed by subsequent range reduction or extinction events (14 mya to present), 3) dispersal across a southern land bridge (10–5.3 mya) followed by vicariance caused by the reopening of Strait of Bab el Mandeb (5.3 mya), and 4) overwater sweepstakes dispersal events (5.3 mya and onwards). Alternatively, 5) human-mediated dispersal is a plausible non-natural explanation for a disjunct Afro-Arabian distribution pattern. To deduce which scenario is most likely for a particular species group or taxon, the divergence dates of the Arabian and African populations must be investigated in a phylogenetic framework. Additionally, the ecology of the organism can also inform the plausibility of these hypotheses, as vagility, physiology, habitat preference, and other life history characteristics affect both dispersal ability and niche breadth.

While few phylogenetic or phylogeographic studies have been conducted on taxa with Afro-Arabian distributions, the results of these studies suggest that each of the five proposed biogeographic hypotheses are supported in at least one group or taxon (Table 14). There appear to be no consistent patterns across groups, owing to a combination of stochastic natural processes and differences in life history characteristics. However, the most recent documented natural invasion of the Arabian Peninsula occurred during the Pleistocene in the hamadryas baboon, a species with high vagility and dispersal ability.

The monitor lizards (family Varanidae) are an assemblage of small to large-sized anguimorph lizards with an Old World distribution. There are currently 58 recognized varanid species, six of which are distributed across Africa and the Arabian Peninsula (*Varanus albigularis*, *V. exanthematicus*, *V. griseus*, *V. niloticus*, *V. ornatus*, and *V. yemenensis*) (Böhme 2003). The African monitors are all large bodied, conspicuous, highly vagile animals that occupy a variety of habitat types. While *V. albigularis*, *V. exanthematicus*, *V. niloticus*, and *V. ornatus* each have distributions on continental

Africa, *V. griseus* is distributed throughout Northern Africa across the Sahara Desert as well as the Arabian Peninsula, southwestern Asia, and northwestern India (Fig. 21) (Bayless 1997, 2002; Böhme 2003). The recently described *V. yemenensis* is found only on the Arabian Peninsula in Yemen and southwest Saudi Arabia (Fig. 21) (Böhme et al. 1987; Böhme et al. 1989). Based on hemipeneal morphology, *V. yemenensis* is thought to be most closely related to *V. albigularis*, a species distributed from the Horn of Africa to southern Africa (Böhme et al. 1989). Although *V. griseus* was not included in the morphological study, thereby preventing investigation of a possible Asian origin for *V. yemenensis*, these two species are morphologically dissimilar and occur in sympatry and are therefore unlikely to be sister taxa.

Given the putative sister relationship of *V. yemenensis* to *V. albigularis*, it is likely that the Red Sea has been the primary factor for diversification between these species. However, it is unclear which biogeographic scenario best explains the arrival of *V. yemenensis* to the Arabian Peninsula. In this study, we use both nuclear and mitochondrial data to reconstruct phylogenies and estimate divergence timings. Using the results of these analyses, we investigate which biogeographic hypotheses are most plausible for *V. yemenensis*, and compare our results to previously examined taxa in an attempt to better understand the complex history of this region.

Methods

Sampling

Our sampling includes thirteen individuals from eight recognized varanid species (Table 15). These include all five varanid species known to occur in Africa as well as representatives from the Arabian Peninsula, Asia-Minor, and Southeast Asia. We also included the following outgroup anguimorph sequences from GenBank: *Lanthanotus borneensis* (AY662537, AY662609), *Shinisaurus crocodilurus* (AF085604, AY662610), *Elgaria multicarinata* (AF085620, GU457977), *Heloderma suspectum* (AF407540, AY662606), and *Xenosaurus grandis* (XGU71333, AY662607).

Molecular Data

Whole genomic DNA was extracted from liver or scale tissue using a Qiagen DNeasy™ tissue kit. PCR amplifications were conducted in 25 µL volume reactions initiated at 95°C for 2 min followed by 31 cycles of 95°C for 35s, 56°C (nucDNA) or 50°C (mtDNA) for 35s, and 72°C for 1 min 35s (with extension increasing 4s per cycle). PCR products were visualized on 1% agarose gels stained with ethidium bromide and purified using ExoSAPIT (USB Corp.). Products were forward and reverse sequenced using BigDye v.3.1 chemistry (Applied Biosystems, Inc.) on an ABI3730 Genetic Analyzer (Applied Biosystems, Inc.).

We obtained partial sequence from the mitochondrial (mtDNA) gene *ND2* and partial exonic sequence from nuclear (nucDNA) gene *RAG-1* using primers designed for this study (Table 16). Sequences were analyzed using Geneious Pro version 5.3.4 (Drummond et al. 2009) and aligned by eye with no ambiguities. Sequences were translated to ensure conservation of the correct reading frame. All sequences are deposited in Genbank (Accession Nos. JN673342–JN673367).

Phylogenetic Analyses

Gene trees were constructed using both maximum likelihood and Bayesian analyses. Genes were analyzed independently and partitioned by codon position, and we also analyzed a concatenated data set partitioned by gene and codon. The Akaike Information Criterion (AIC) in jModelTest v0.1.1 (Posada 2008) was used to find the appropriate models of evolution for each gene and the appropriate model for codon partitions (Posada and Buckley 2004). Maximum likelihood analyses for each data set were performed with three replicates in GARLI-Part v0.97 (Zwickl 2006) using specific models selected in jModelTest. For each analysis, 500 nonparametric bootstrap replicates were performed and a 50% majority-rule consensus tree was produced. Bayesian analyses were conducted on each data set using MrBayes v3.1.2 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). Parallel runs were performed with random starting trees and allowed to run for $2x10^6$ generations, with Markov chains sampled every 1,000 generations. All analyses were examined with Tracer v1.5.0 (Rambaut and Drummond 2009) to ensure parameters and statistics had reached stationarity and sufficient ESS values (>150). The first 25% of the total number of generations were discarded as burn-in, leaving 15,000 trees for each parallel run. The resulting post-burn-in trees from the two parallel runs were combined and a 50% majority rule consensus tree was calculated from a total of 30,000 trees.

Dating Analyses

To infer timing of lineage divergences, we utilized a *RAG-1* data set incorporating representatives of all anguimorph families (Helodermatidae, Anguidae, Xenosauridae, Shinisauridae, Lanthanotidae, Varanidae) and three snake families (Viperidae, Uropeltidae, Colubridae). Dating analyses were carried out in BEAST v1.6.1 (Drummond and Rambaut 2007). We calibrated analyses using four calibration points, including the time to most recent common ancestry (tmrca) of Helodermatidae and Xenosauridae = 99 Mya (Nydam 2000), the tmrca of Viperidae, Uropeltidae, and Colubridae = 93.5 Mya (Zaher and Rieppel 2002), the tmrca of Viperidae and Colubridae = 33.7 Mya (Rage et al. 1992), and a minimal constraint of 37 mya for the genus *Varanus* based on the earliest fossil series of vertebrae attributed to the genus (Holmes et al. 2010). We included only one exemplar per species and therefore used the Yule model of speciation as the tree prior. An uncorrelated lognormal clock model was used to account for lineage-specific rate heterogeneity and analyses were run $5x10^7$ generations sampled every 1000 generations. Analyses were examined with Tracer v1.5.0 (Rambaut and Drummond 2009) to ensure parameters and statistics had reached stationarity and sufficient ESS values (>150). The tmrca was investigated for (1) the genus *Varanus* and (2) *Varanus yemenensis* and its sister taxon.

Molecular Diversity

To further assess levels of divergence within African varanids, uncorrected pairwise differences were estimated using PAUP* (Swofford 2002).

Results

The final alignment of *ND2* contains 591 bp (247 variable ingroup sites, 198 parsimony informative) and the final alignment of *RAG-1* contains 758 bp (70 variable ingroup sites, 57 parsimony informative) (TreeBASE ID 11910). Both mitochondrial and nuclear data recover *Varanus* as a monophyletic genus with strong support (Fig. 22; Bayesian posterior probability [PP] = 1.0, ML bootstrap [ML] =100%). The *Varanus* gene tree based on nuclear data recovers an African clade and an Asian clade, and the African clade contains all the African-distributed varanid species as well as *V. griseus* and *V. yemenensis*. *Varanus albigularis microstictus* is recovered as the closest relative of *V. yemenensis* with high support (PP = 0.98, ML = 72%). The tree topology of the concatenated data set is identical to that of the nuclear data set, and all nodes are strongly supported (not shown). The mitochondrial gene tree recovers a polytomy of three clades: an Asian clade (PP = 0.99, ML = 99%); an African clade containing *V. albigularis*, *V. exanthematicus* and *V. yemenensis* (PP = 1.0, ML = 100%); and a poorly supported African clade containing *V. griseus*, *V. niloticus* and *V. ornatus* (PP = 0.55, ML = 50%). In this phylogeny, *V. yemenensis* is also recovered as sister to *V. a. microstictus* with high support (PP = 1.0, ML = 99%). The uncorrected pairwise distances reveal *V. yemenensis* and *V. a. microstictus* are approximately 5.8% divergent based on mitochondrial data and 1.2% divergent based on nuclear data (Table 17). The sequence divergences between the subspecies *V. a. albigularis* and *V. a. microstictus* are 9.6% and 1.8% based on mtDNA and nucDNA, respectively. Values of uncorrected pairwise sequence divergences between other varanid species range from 16.2 – 24.3% and 2.6 – 5.3% based on mtDNA and nucDNA, respectively, with greatest sequence differences occurring between African and Asian taxa (Table 17).

The BEAST analysis based on *RAG-1* returned a topology identical to that recovered with Bayesian and maximum likelihood analyses, with both an Asian and African clade being recovered with high support (Fig. 22, TreeBASE ID 11910). In this topology *V. yemenensis* and *V. a. microstictus* are again recovered as sister taxa (PP = 1.0). The tmrca for the genus *Varanus* is 40.1 mya, with a 95% highest posterior density region (HPD) of 27.6–53.8 mya. The estimate for the divergence between *V. yemenensis* and *V. a. microstictus* is 6.9 mya (95% HPD = 2.1–12.5 mya).

Discussion

Both mitochondrial and nuclear data unequivocally demonstrate *V. yemenensis* is most closely related to *V. a. microstictus*, a subspecies of the *V. albigularis* complex. This subspecies is distributed from the Horn of Africa to Mozambique, and is in close geographic proximity to the Arabian Peninsula and the known range of *V. yemenensis*. We find no evidence for an Arabian or Asian origination of *V. yemenensis*, which would necessitate a sister taxon relationship with the sympatrically distributed *V. griseus* or Asian monitors. In the phylogenies recovered, *V. griseus* is either the most basal member of the African clade or weakly supported as the sister clade to the *Varanus niloticus*/*ornatus* complex. Therefore, based on our molecular evidence, we can refute an Arabian or Asian origin of *V. yemenensis* in favor of an African origin, a finding that is also supported by hemipeneal and external morphology (Böhme et al. 1989).

The *V. albigularis* complex is currently rendered paraphyletic based on the phylogenetic position of *V. yemenensis*. Additionally, the level of genetic divergence

between *V. a. microstictus* and *V. a. albigularis* exceeds that of *V. a. microstictus* and *V. yemenensis*, indicating that the *V. albigularis* complex warrants further investigation. As our only sample of *V. a. albigularis* is from Botswana (approximately 3800 km from our sample of *V. a. microstictus*), we cannot determine if the levels of genetic divergence are a result of geographic distance or actual genetic isolation between the proposed subspecies. Additional sampling of *V. a. albigularis* and *V. a. microstictus* from their inferred contact zone in Mozambique could help clarify which scenario best explains our results.

Given the origination of *V. yemenensis* from the *V. albigularis* complex, it is clear that this species attained its current range either through vicariance or via dispersal from the Horn of Africa. The divergence estimate of 6.9 mya (HPD = 2.1–12.5 mya) suggests that neither the primary vicariance (initial rifting of the Red Sea) nor human-mediated dispersal hypotheses are likely scenarios for this group. However, our timing estimate and the associated error measure allows consideration of the northern landbridge, southern landbridge, and overwater dispersal scenarios. While a northern dispersal route was already well established by 9.5–5.0 mya, the current absence of the *V. albigularis* complex north of the Horn of Africa makes this hypothesis unlikely, as extinction of this group across both northern Africa and the northern Arabian Peninsula must be invoked to explain current species distributions. The southern landbridge hypothesis posits the ancestor to *V. yemenensis* dispersed to the Arabian Peninsula during a southern landbridge connection and subsequently experienced vicariance with the reopening of the Strait of Bab el Mandeb after 5.3 mya. The overwater dispersal hypothesis suggests the ancestor of *V. yemenensis* dispersed sometime after 5.3 mya across the open ocean of the Strait of Bab el Mandeb to reach the Arabian Peninsula. Recent work on divergence dating analyses suggests that estimating divergence dates using a single gene tree will inherently produce older divergence estimates than dates obtained using species trees based on multilocus data in a coalescent framework (McCormack et al. 2011). However, given the subtle difference in timing between these two hypotheses, the closeness in timing of both hypotheses to our mean divergence date estimate, and the relative error associated with our dating analyses (2.1–12.5 mya), we cannot confidently distinguish between these two scenarios.

While overwater dispersal events may seem unlikely for many groups, varanids are excellent swimmers and tolerant to marine conditions. Varanids were the first terrestrial vertebrates to recolonize the Krakatau island group after the fauna was erased by volcanic eruption, a clear case of successful overwater marine dispersal (Rawlinson et al. 1990). It is therefore conceivable that this species could have crossed the Strait of Bab el Mandeb in a similar manner. This scenario is even more plausible because the Strait of Bab el Mandeb was historically much narrower (minimally 5 km) during lowered sea levels that resulted from high latitude glaciation cycles (Fernandes et al. 2006; Siddall et al. 2003). Based on divergence times obtained from previous studies, overwater dispersal is also a possible explanation for historical Afro-Arabian movements in the common puff adder (Pook et al. 2009), African helmeted terrapins (Wong et al. 2010), and Hamadryas baboons (Newman et al. 2004; Winney et al. 2004; Zinner et al. 2009).

The Red Sea appears to currently be preventing gene flow in several reptile species or species complexes which exhibit Oligocene or Miocene Afro-Arabian divergences, including agamids, chameleons, elapids, gekkonids, and vipers (Amer and

Kumazawa 2005; Fujita and Papenfuss 2011; Macey et al. 2008; Papenfuss et al. 2009; Pook et al. 2009). While it is clear that gene flow has ceased in more recently disjunct Afro-Arabian puff adders, varanids, and helmeted terrapins, the story is not as clear for Hamadryas baboons, which may have experienced multiple colonization events or introgression of mitochondrial haplotypes (Zinner et al. 2009). Therefore, although the Red Sea currently appears to be a strong biogeographic barrier for several reptile groups, it may only act as a partial barrier to highly vagile groups such as primates. As more phylogeographic and phylogenetic studies continue to be conducted on a variety of taxa distributed across the Horn of Africa and southwestern Arabian Peninsula, we can begin to better understand the complex historical and current biogeographic effects of the Red Sea.

Acknowledgments
We would like to thank Amoud University (Borama, Somaliland), the California Academy of Sciences (CAS), the South African National Biodiversity Institute (SANBI), and the Yemeni Institute for Yemeni Studies for tissues, funding, or access to field sites. Additional tissues were provided by Aaron Bauer and Adam Leache.

References

Amer, S.A.M., and Y. Kumazawa. 2005. Mitochondrial DNA sequences of the Afro-Arabian spiny-tailed lizards (genus *Uromastyx*; family Agamidae): phylogenetic analyses and evolution of gene arrangements. Biological Journal of the Linnean Society 85:247–260.

Bayless, M.K. 1997. The distribution of African monitor lizards (Sauria: Varanidae). African Journal of Ecololgy 35:374–377.

Bayless, M.K. 2002. Monitor lizards: A pan-African check-list of their zoogeography (Sauria: Varanidae: *Polydaedalus*). Journal of Biogeography 29:1643–1701.

Benton, M.J., and P.C.J. Donoghue. 2007. Paleontological evidence to date the tree of life. Molecualar Biology and Evolution 24:26–53.

Bohannon, R.G. 1986. Tectonic configuration of the Western Arabian Continental Margin, southern Red Sea. Tectonics 5:477–499.

Böhme, W. 2003. Checklist of the living monitor lizards of the world (family Varanidae). Zoologische Verhandelingen 341:1–43.

Böhme, W., J.P. Fritz, and F. Schutte. 1987. Neuentdeckung einer Grobechse (Sauria: *Varanus*) aus der Arabischen Republik Jemen. Herpetofauna 9:13–20.

Böhme, W., U. Joger, and B. Schatti. 1989. A new monitor lizard (Reptilia: Varanidae) from Yemen, with notes on ecology, phylogeny and zoogeography. Fauna Saudi Arabia 10:433–448.

Bosworth, W., P. Huchon, and K. McClay. 2005. The Red Sea and Gulf of Aden Basins. Journal of African Earth Sciences 43:334–378.

Drummond, A., B. Ashton, M. Cheung, J. Heled, M. Kearse, R. Moir, S. Stones-Havas, T. Thierer, and A. Wilson. 2009. Geneious v5.3. Available from http://www.geneious.com/.

Drummond, A.J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214–222.

Fernandes, C.A., E.J. Rohling, M. Siddall. 2006. Absence of post-Miocene Red Sea land bridges: biogeographic implications. Journal of Biogeography 33:961–966.

Fujita, M.K., and T.J. Papenfuss. 2011. Molecular systematics of *Stenodactylus* (Gekkonidae), an Afro-Arabian gecko species complex. Molecular Phylogenetics and Evolution 58:71–75.

Holmes, R.B., A.M. Murray, Y.S. Attia, E.L. Simons, and P. Chatrath. 2010. Oldest known *Varanus* (Squamata: Varanidae) from the upper Eocene and lower Oligocene of Egypt: support for an African origin of the genus. Palaeontology 53:1099–1110.

Huelsenbeck, J.P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Hugall, A.F., R. Foster, and M.S.Y. Lee. 2007. Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. Systematic Biology 56:543–563.

Lavin, B.R., and T.J. Papenfuss. 2012. The phylogenetic position of *Chalcides ocellatus* (Squamata: Scincidae) from Yemen and Somalia. Zootaxa 3221:26–36.

Macey, J.R., J.V. Kuehl, A. Larson, M.D. Robinson, I.H. Ugurtas, N.B. Ananjeva, H. Rahman, H.I. Javed, R.M. Osmani, A. Doumma, and T.J. Papenfuss. 2008.

Socotra Island the forgotten fragment of Gondwana: Unmasking chameleon lizard history with complete mitochondrial genomic data. Molecular Phylogenetics and Evolution 49:1015–1018.
McCormack, J.E., J. Heled, K.S. Delaney, A.T. Peterson, and L.L. Knowles. 2011. Calibrating divergence times on species trees versus gene trees: implications for speciation history of *Aphelocoma* jays. Evolution 65:184–202.
Muller, J., and R.R. Reisz. 2005. Four well-constrained calibration points from the vertebrate fossil record for molecular clock estimates. BioEssays 27:1069–1075.
Near, T.J., P.A. Meylan, and H.B. Shaffer. 2005. Assessing concordance of fossil calibration points in molecular clock studies: An example using turtles. American Naturalist 165:137–146.
Newman, T.K., C.J. Jolly, and J. Rogers. 2004 Mitochondrial phylogeny and systematics of baboons (*Papio*). American Journal of Physical Anthropology 124:17–27.
Nydam, R. L. 2000. A new taxon of helodermatid-like lizard from the Albian-Cenomanian of Utah. Journal of Vertebrate Paleontology 20:285–294.
Papenfuss, T.J., T. Jackman, A. Bauer, B.L. Stuart, M.D. Robinson, and J.F. Parham. 2009. Phylogenetic relationships among species in the sphaerodactylid lizard genus *Pristurus*. Proceedings of the California Academy of Sciences 60:675–681.
Pook, C.E., U. Joger, N. Stümpel, and W. Wüster. 2009. When continents collide: Phylogeny, historical biogeography and systematics of the medically important viper genus *Echis* (Squamata: Serpentes: Viperidae). Molecular Phylogenetics and Evolution 53:792–807.
Posada, D. 2008. jModelTest: Phylogenetic Model Averaging. Molecular Biology and Evolution 25:1253–1256.
Posada, D., and T.R. Buckley. 2004. Model selection and model averaging in phylogenetics: advantages of the AIC and Bayesian approaches over likelihood ratio tests. Systematic Biology 53:793–808.
Rage, J.C., E. Buffetaut, H. Buffetaut-Tong, Y. Chaimanee, S. Ducrocq, J.J. Jaeger, and V. Suteethorn. 1992. A colubrid snake in the late Eocene of Thailand: the oldest known Colubridae (Reptilia, Serpentes). Comptes Rendus de l'Académie des Sciences 314:1085–1089.
Rambaut, A., and A.J. Drummond. 2009. Tracer v1.5.0. Avaliable from http://beast.bio.ed.ac.uk/.
Rawlinson, P.A., A.H.T. Widjoya, M.N. Hutchinson, and G.W. Brown. 1990. The terrestrial vertebrate fauna of the Krakatau Islands Sunda Strait Indonesia 1883–1986. Philosophical Transactions of the Royal Society B 328:3–28.
Reisz, R.R., and J. Muller. 2004. Molecular timescales and the fossil record: a paleontological perspective. Trends in Genetics 20:237–241.
Rohling, E.J., M. Fenton, F.J. Jorissen, P. Bertrand, G. Ganssen, and J.P. Caulet. 1998. Magnitudes of sea-level lowstands of the past 500,000 years. Nature 394:162–165.
Ronquist, F., and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.
Siddall, M., E.J. Rohling, A. Almogi-Labin, C. Hemleben, D. Meischner, I. Schmelzer, and D.A. Smeed. 2003. Sea-level fluctuations during the last glacial cycle. Nature 423:853–858.

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

Vargas-Ramìrez, M., M. Vences, W.R. Branch, S.R. Daniels, F. Glaw, M.D. Hofmeyr, G. Kuchling, J. Maran, T.J. Papenfuss, P. Sirok, D.R. Vieites, and U. Fritz. 2010. Deep genealogical lineages in the widely distributed African helmeted terrapin: Evidence from mitochondrial and nuclear DNA (Testudines: Pelomedusidae: *Pelomedusa subrufa*). Molecular Phylogenetics and Evolution 56:428–440.

Winney, B.J., R.L. Hammond, W. Macasero, B. Flores, A. Boug, V. Biquand, S. Biquand, and M.W. Bruford. 2004. Crossing the Red Sea: phylogeography of the hamadryas baboon, *Papio hamadryas hamadryas*. Molecular Ecology 13:2819–2827.

Wong, R.A., J.J. Fong, and T.J. Papenfuss. 2010. Phylogeography of the African helmeted terrapin, *Pelomedusa subrufa*: genetic structure, dispersal, and human introduction. Proceedings of the California Academy of Sciences 61:575–585.

Zaher, H., and O. Rieppel. 2002. On the phylogenetic relationships of the Cretaceous snakes with legs, with special reference to *Pachyrhachis problematicus* (Squamata, Serpentes). Journal of Vertebrate Paleontology 22:104–109.

Zinner, D., L.F. Groeneveld, C. Keller, and C. Roos. 2009. Mitochondrial phylogeography of baboons (*Papio* spp.) - indication for introgressive hybridization? BMC Evolutionary Biology 9:1–15.

Zwickl, D.J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. The University of Texas at Austin.

Figure 20. The geography of the Red Sea, Arabian Peninsula, and Horn of Africa.

Figure 20

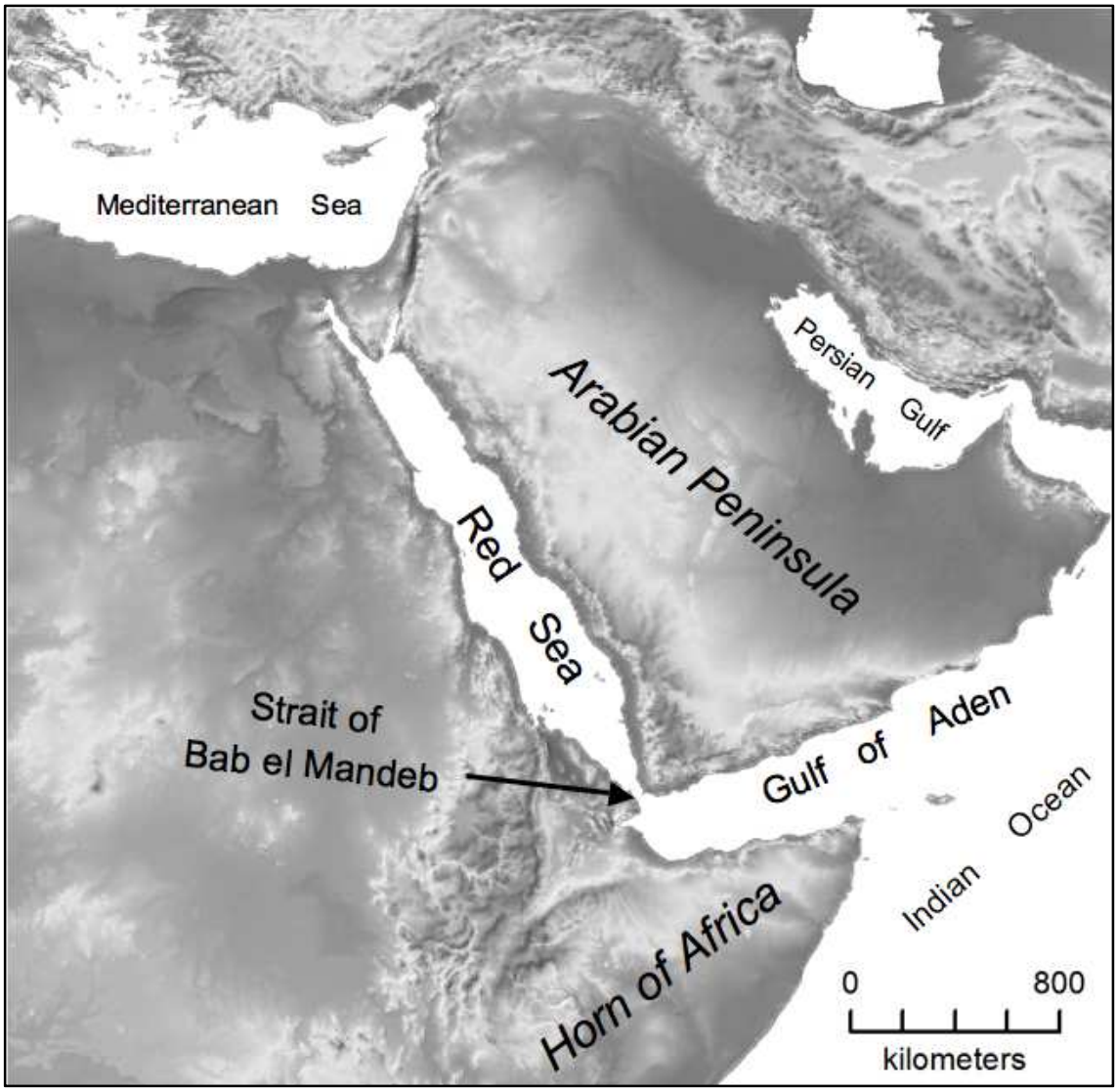

Figure 21. Proposed distribution maps of all six African monitors. Dashes show contact zones for subspecies within a species complex.

Figure 21

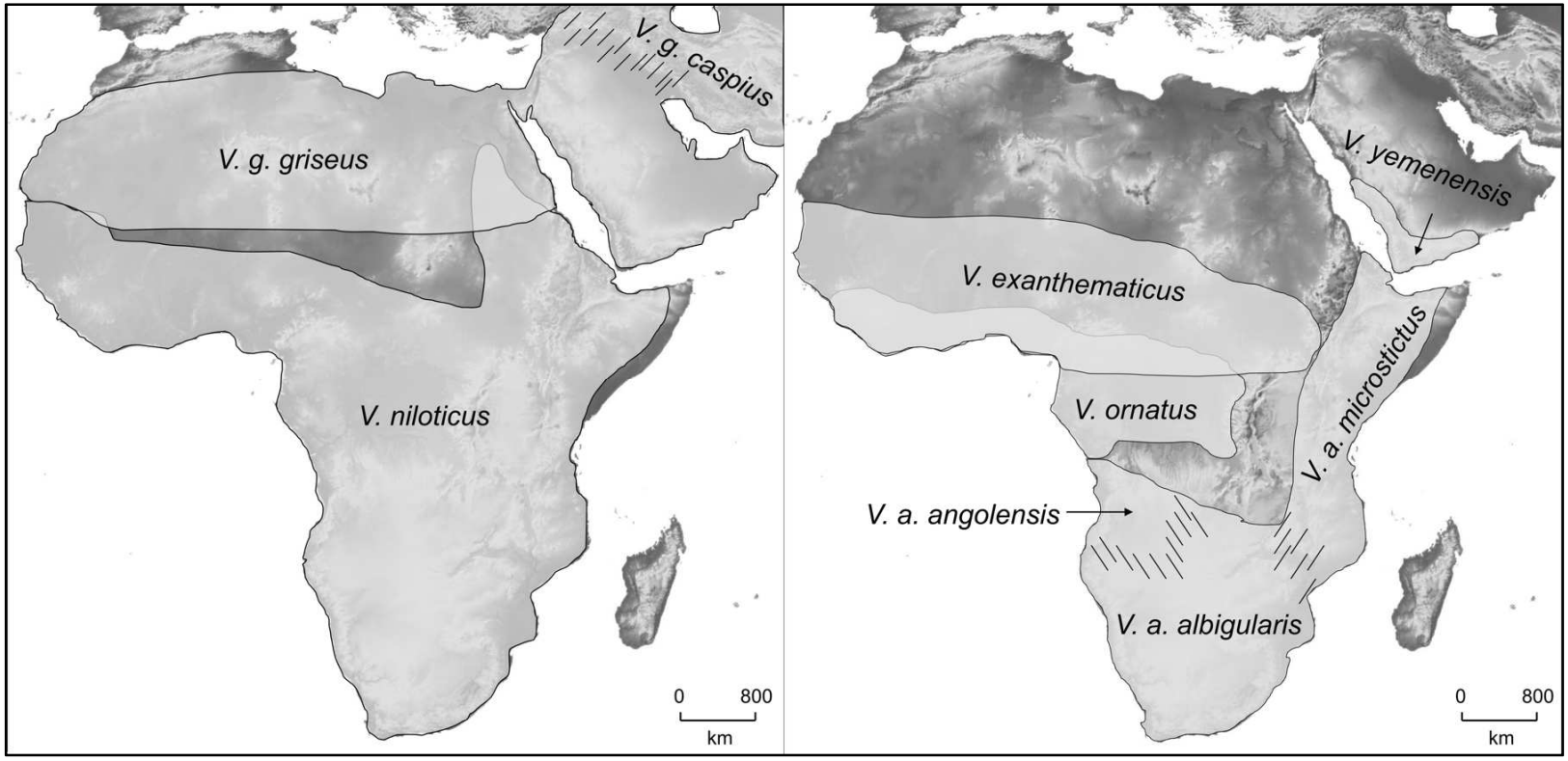

Figure 22. The BEAST maximum clade credibility tree based on RAG-1 data. Error bars on nodes depict 95% highest posterior density intervals for node ages, and median node age (mya) is presented as an italicized number. Posterior probability support values are given as bold numbers above branches. The scale bar below displays age in millions of years.

Figure 22

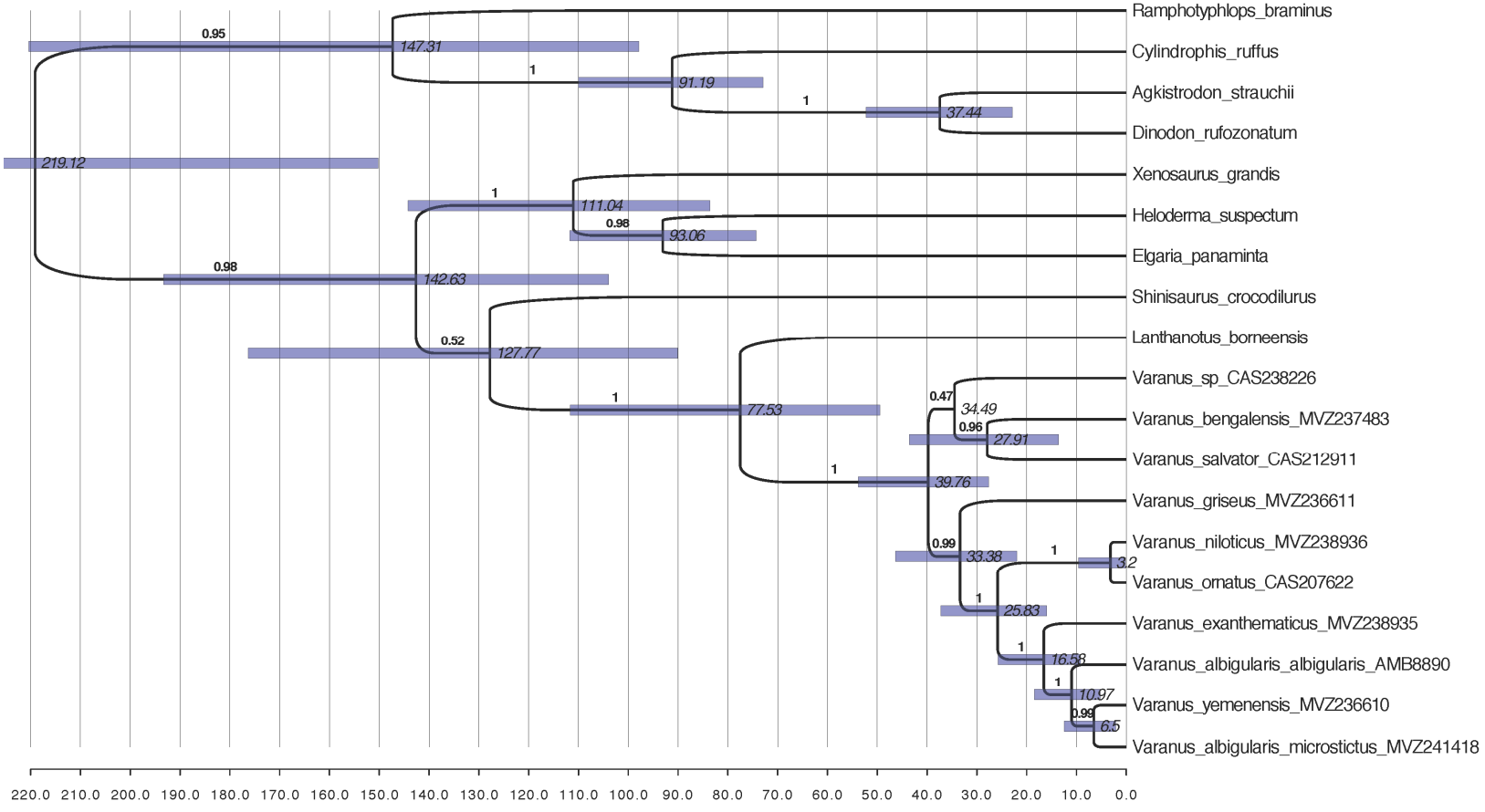

Figure 23. Phylogenies reconstructed using Bayesian analysis for both mtDNA (*ND2*) and nucDNA (*RAG-1*). Nodes receiving either >0.95 posterior probability support (top) or >70% likelihood bootstrap score (bottom) are indicated by a circle, whereas dashes indicate less than these threshold values. African distributed taxa are outlined in gray. Scale bar represents substitutions/site.

Figure 23

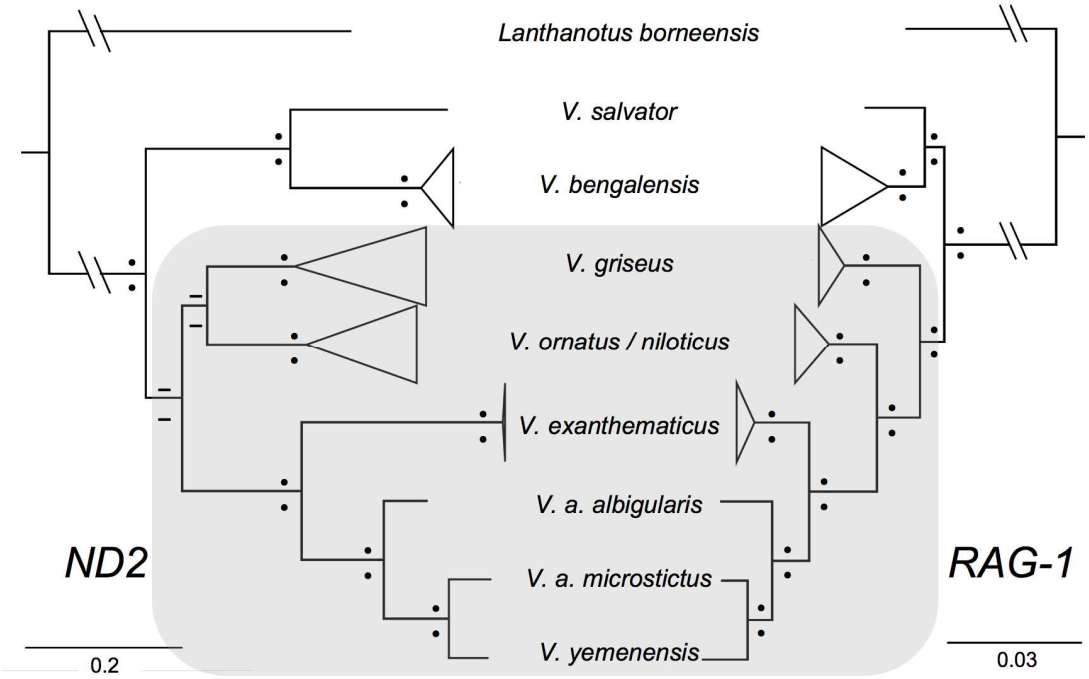

Table 14. A compilation of studies conducted on groups with disjunct Afro-Arabian distributions. The species or genus is given along with divergence date estimates as indicated by each study. We use the authors hypotheses of historical biogeography or interpret our own based on timings of our proposed hypotheses: 1, primary vicariance (initial rifting); 2, northern route; 3, southern route; 4, overwater sweepstakes; 5, human-mediated dispersal.

Table 14

Order	Family	Species or Genus	Divergence estimate	Hypothesis	Study	Molecular marker
Squamata	Agamidae	*Uromastyx ocellata* vs. *U. ornata*	7-8 mya	3	(Amer and Kumazawa 2005)	mtDNA
	Chameleonidae	*Chamaeleo chamaeleon*	<10 mya	3	(Macey et al. 2008)	mtDNA
	Elapidae	*Naja haje*	1.75 mya	2	(Pook et al. 2009)	mtDNA
	Gekkonidae	*Pristurus*	none	1, 2, or 3	(Papenfuss et al. 2009)	mtDNA
		Stenodactylus	27.9 mya (41.3-15.8 mya)	1	(Fujita and Papenfuss 2011)	mtDNA, nucDNA
	Scincidae	*Chalcides ocellatus*	n/a	5	Lavin and Papenfuss, in press	mtDNA, nucDNA
	Varanidae	*Varanus yemenensis* vs. *V. albigularis*	5.2 mya (9.5-1.6 mya)	4	This study	mtDNA, nucDNA
	Viperidae	*Echis pyramidum*	8.1 mya	3	(Pook et al. 2009)	mtDNA
		Echis pyramidum vs. *E. coloratus*	19.4 mya	1	(Pook et al. 2009)	mtDNA
		Bitis arietans	4 mya	4	(Pook et al. 2009)	mtDNA
Testudines	Pelomedusidae	*Pelomedusa subrufa*	none	3	(Vargas-Ramìrez et al. 2010)	mtDNA, nucDNA
			2.4 mya (14.1 – 1 mya)	3, 4	(Wong et al. 2010)	mtDNA
Primates	Cercopithecidae	*Papio hamadryas*	619 kya (577-660 kya)	4	(Newman et al. 2004)	mtDNA
			443- 156 kya	4	(Winney et al. 2004)	mtDNA
			320 kya (160-510 kya)	4	(Zinner et al. 2009)	mtDNA

Table 15. A list of *Varanus* species examined along with museum numbers and relevant geographic sampling information. AMB, Aaron M. Bauer field number; CAS, California Academy of Sciences; MVZ, Museum of Vertebrate Zoology.

Table 15

Species	Museum Number	Locality	Latitude	Longitude
V. a. albigularis	AMB 8890	Botswana	20°13'40" S	26°16'03" E
V. a. microstictus	MVZ 241418	Somalia	09°57'06" N	43°12'59" E
V. bengalensis	MVZ 237483	Laghman Province, Afghanistan	34°29'36" N	70°18'00" E
V. bengalensis	CAS 213887	Magwe Division, Myanmar	20°08'00" N	94°36'05" E
V. exanthematicus	MVZ 238934	Niger	14°00'40" N	05°19'18" E
V. exanthematicus	MVZ 238935	Niger	14°00'40" N	05°19'18" E
V. griseus	MVZ 235860	Nouakchott District, Mauritania	18°38'58" N	15°37'43" W
V. griseus	MVZ 236611	Al Hudaydah Governorate, Yemen	14°49'25" N	43°07'38" E
V. niloticus	MVZ 238936	Niger	13°37'48" N	01°52'42" E
V. niloticus	MVZ 245359	Volta Region, Ghana	08°19'30" N	00°33'22" E
V. ornatus	CAS 207622	Bioko Island, Equatorial Guinea	03°42'39" N	08°39'59" E
V. salvator	CAS 212911	Ayeyarwade Division, Myanmar		
V. yemenensis	MVZ 236610	Yemen	14°21'11" N	43°19'08" E

Supplementary Table 16. List of primers designed for this study.

Table 16

Gene	Primer Name	Source	Sequence (5' to 3')
ND2	ND2 MetVF1	This study	TCGGGCCCATACCCCGAAAA
	ND2 VR567	This study	TCAACCTAGGTGGGCGATTGA
RAG-1	RAG-1 VF16	This study	CTTGRAAAGCCACTTCCTGAA
	RAG-1 VF69	This study	AGTGGCAGCTTCTCTGGTCA
	RAG-1 VR901	This study	GCAAGGATAGCGACAGGATGG GCA

Table 17. Uncorrected pairwise differences among taxa studied. Below diagonal, mtDNA values; above diagonal, nucDNA values.

Table 17

	Museum No.	Species	1	2	3	4	5	6	7	8	9	10	11	12	13
1	AMB 8890	*V. a. albigularis*		0.018	0.052	0.047	0.024	0.023	0.041	0.041	0.039	0.039	0.038	0.045	0.012
2	MVZ 241418	*V. a. microstictus*	0.096		0.048	0.051	0.023	0.022	0.040	0.041	0.034	0.036	0.036	0.048	0.011
3	MVZ 237483	*V. bengalensis*	0.225	0.229		0.020	0.053	0.051	0.038	0.041	0.040	0.044	0.039	0.029	0.047
4	CAS 213887	*V. bengalensis*	0.217	0.210	0.044		0.048	0.048	0.038	0.038	0.038	0.042	0.036	0.026	0.045
5	MVZ 238934	*V. exanthematicus*	0.176	0.202	0.239	0.227		0.004	0.046	0.042	0.034	0.038	0.030	0.043	0.028
6	MVZ 238935	*V. exanthematicus*	0.178	0.201	0.236	0.234	0.000		0.045	0.042	0.033	0.036	0.030	0.043	0.028
7	MVZ 236611	*V. griseus*	0.197	0.214	0.226	0.214	0.233	0.239		0.006	0.032	0.033	0.033	0.032	0.037
8	MVZ 235860	*V. griseus*	0.195	0.211	0.224	0.215	0.232	0.236	0.034		0.032	0.034	0.035	0.035	0.037
9	MVZ 238936	*V. niloticus*	0.174	0.195	0.220	0.205	0.207	0.213	0.172	0.181		0.004	0.003	0.031	0.033
10	MVZ 245359	*V. niloticus*	0.185	0.212	0.230	0.224	0.212	0.219	0.192	0.182	0.102		0.007	0.035	0.034
11	CAS 207622	*V. niloticus*	0.179	0.200	0.219	0.203	0.204	0.211	0.177	0.175	0.068	0.095		0.033	0.034
12	CAS 212911	*V. salvator*	0.240	0.239	0.162	0.165	0.243	0.242	0.218	0.205	0.204	0.209	0.195		0.046
13	MVZ 236610	*V. yemenensis*	0.098	0.058	0.232	0.215	0.193	0.195	0.200	0.192	0.196	0.200	0.199	0.238	

Chapter 5

Historical biogeography resolves the origins of endemic Arabian toad lineages (Anura: Bufonidae): Evidence for ancient vicariance and dispersal events with the Horn of Africa and South Asia

This chapter was published prior to the filing of my dissertation as follows:
Portik, D.M., and T.J. Papenfuss. 2015. Historical biogeography resolves the origins of endemic Arabian toad lineages (Anura: Bufonidae): Evidence for ancient vicariance and dispersal events with the Horn of Africa and South Asia. BMC Evolutionary Biology 15:152.

Abstract
The Arabian Peninsula is home to a unique fauna that has assembled and evolved throughout the course of major geophysical events, including the separation of the Arabian Plate from Africa and subsequent collision with Eurasia. Opportunities for faunal exchanges with particular continents occurred in temporally distinct periods, and the presence of African, Western Eurasian, and South Asian derived taxa on the Arabian Peninsula signifies the complexity of these historical biogeographic events. The six true toad species (family Bufonidae) endemic to Arabian Peninsula present a considerable taxonomic and biogeographic challenge because they are part of a global bufonid radiation, including several genera surrounding the Arabian Peninsula, and difficult to discriminate morphologically. As they could be derived from African, Western Eurasian, or South Asian toad groups, elucidating their evolutionary relationships has important implications for historical biogeography. Here, we analyze a global molecular data set of 243 bufonid lineages, with an emphasis on new sampling from the Horn of Africa, Western Eurasia, South Asia, and the Arabian Peninsula, to reconstruct the evolutionary relationships of the Arabian species. We produce a robust time-calibrated phylogeny to infer the biogeographic history of this group on and around the Arabian Peninsula. Our phylogenetic analyses indicate two of the endemic Arabian toad species, "*Bufo*" *tihamicus* and "*Bufo*" *arabicus*, evolved independently within the African genus *Amietophrynus*. We confirm the Arabian species *Duttaphrynus dhufarensis* is of South Asian origin, but do not find evidence for the Asian genus *Duttaphrynus* being present in the Horn of Africa, discrediting a previously proposed Asian bufonid dispersal event to Africa. We also do not find evidence of the African genus *Amietophrynus* occurring in South Asia, suggesting that unlike many other vertebrate taxa, toads have not used the Arabian Peninsula as a stepping-stone for trans-continental dispersal. Our divergence dating estimates strongly suggest the formation of the Red Sea drove simultaneous divergences between two of the Arabian species (*A. tihamicus* comb. nov. and *A. arabicus* comb. nov.) and their closest mainland African relatives in the Early Miocene. We estimate the divergence of *D. dhufarensis* with its closest South Asian relatives occurred in the mid to Late Miocene, suggesting the temporary or permanent land connections between the Arabian plate and Eurasia facilitated dispersal of this lineage to the Arabian Peninsula. The Arabian bufonid assemblage, despite being comparatively depauperate with respect to surrounding continents, exemplifies the faunal pattern of the Arabian Peninsula, namely being a complex admixture of African, Western Eurasian, and

South Asian elements. The historical biogeographic patterns exhibited by Arabian toads and their allies are concordant with studies of other vertebrate taxa, building support for the role of major geological events in driving simultaneous vicariance and dispersal events around the Arabian Peninsula. Although many taxa or groups exhibiting disjunct Afro-Arabian distributions appear to have dispersed more recently from the Horn of Africa via a southern land bridge or overwater dispersal, both *Amietophrynus tihamicus* and *A. arabicus* likely represent true African relicts resulting from vicariance associated with the Red Sea formation, a pattern that so far is rare among the vertebrate species investigated.

Introduction

The Arabian Peninsula possesses a unique assemblage of plant and animal species that result from the dynamic geologic history and shifting climate of this region. The flora and fauna range from localized temperate endemics occurring in the disjunct extensions of the Horn of Africa Hotspot and the Eastern Afromontane Hotspot (Mittermeier et al. 1999; Myers et al. 2000) in the southwest regions (Fig. 24), to more broadly distributed arid-adapted clades occurring throughout much of the Arabian Peninsula (Fisher & Membery 1998; Gasperetti 1988; Guba and Glennie 1998; Joger 1987; Kürschner 1998; Mallon 2011; Meister et al. 2005; Meister et al. 2006; Miller and Cope 1996; Vesey-Fitzgerald 1955; Vesey-Fitzgerald 1957). Many large-scale biogeographic patterns for Arabian taxa have been profoundly shaped by the complex tectonic history of the Arabian Peninsula (Amer and Kumazawa 2005; Gvoždík et al. 2010; Metallinou et al. 2012; Pook et al. 2009; Portik and Papenfuss 2012; Šmíd et al. 2013; Wong et al. 2010; Yu et al. 2014). The tectonic activity ultimately responsible for the separation of the Arabian Plate from Africa began approximately 30 Ma, and rifting of the southern Red Sea occurred in the early Miocene (27 Ma), and a connection to the Neotethys Sea and was completed by 23 Ma (Bohannon 1986; Bosworth et al. 2005). The formation of the Red Sea represented a major vicariance event for taxa with formerly continuous Afro-Arabian distributions. The isolated Arabian plate collided with the Anatolian plate approximately 18–16 Ma, forming the temporary *Gomphotherium* bridge (Rögl 1998, 1999). A more permanent land bridge was established after the Arabian plate collided with Eurasia approximately 15 Ma (Bosworth et al. 2005), which allowed regular faunal exchanges between North Africa, the Arabian Peninsula, and Eurasia. These land bridges also created opportunities for trans-continental dispersals using the Arabian Peninsula as a stepping-stone. Sea levels dropped throughout the Red Sea around 10 Ma, providing evidence for the closing of the Strait of Bab el Mandeb and establishment of a southern land bridge between Africa and Arabia (Bosworth et al. 2005). Pliocene marine sediment deposits formed around 5.3 Ma, indicating this land bridge was subsequently lost, and although cyclical Pleistocene land bridges related to glacial cycles have been proposed (Rohling et al. 1998), more recent evidence does not support their existence (Fernandes et al. 2006). Conversely, there is evidence the Persian Gulf was reduced to a series of freshwater lakes (Lambeck 1996) or completely waterless (Uchupi et al. 1999) during the late Pleistocene, providing a connection between the Arabian Peninsula and Southwest Asia. Plio-Pleistocene climate change is hypothesized to have driven speciation in the southwest montane regions of the Arabian Peninsula (Fisher and Membery 1998; Miller and Cope

1996), and is linked to aridification in the Neogene (deMenocal 1995; deMenocal 2004). This may explain additional biogeographic patterns, but on a much more localized scale.

The geologic activity associated with the Arabian Peninsula altered the historical connectivity of this landmass to surrounding continents, and as a result opportunities for faunal exchanges with particular continents occurred in temporally distinct periods. For example, taxa exhibiting disjunct Afro-Arabian distributions either exhibit early Miocene divergences associated with vicariance resulting from the Red Sea formation (30–23 Ma), or dispersed more recently from the Horn of Africa via a southern landbridge or overwater dispersal (10–5.3 Ma or 5.3 Ma to present, respectively) (reviewed in Portik and Papenfuss 2012). If the phylogenetic placement of an Arabian focal group is recovered, predictions can be made concerning the timing of speciation and tested in a temporally explicit phylogenetic framework to infer historical biogeographic patterns. The limited phylogenetic studies focusing on or including Arabian species have demonstrated concordant divergence timings for taxa with similar biogeographic origins (Metallinou et al. 2012; Pook et al. 2009; Portik and Papenfuss 2012; Šmíd et al. 2013; Wong et al. 2010), however these biogeographic patterns remain largely understudied for most Arabian taxa.

Although overall levels of endemism among vertebrate groups across the Arabian Peninsula are moderate, Arabian amphibians exhibit a high degree of endemism (Gvoždík et al. 2010; Balletto et al. 1985). Of the nine described Arabian anuran species, six of these species are true toads belonging to the family Bufonidae, a clade that contains close to 600 described species and which exhibits a nearly worldwide distribution (AmphibiaWeb 2014). The dispersal ability of toads is limited by their reliance on freshwater habitat for breeding and their intolerance to saltwater, making them an interesting system for investigating historical biogeography in this region. Several broad-scale phylogenetic studies of bufonids have revealed the existence of largely discrete continental clades (Frost et al. 2006; Pramuk et al. 2008; Van Bocxlaer et al. 2009; Van Bocxlaer 2010; Pyron and Wiens 2011) that are now accepted as distinct genera or subgenera (AmphibiaWeb 2014; Frost 2014). Three such genera are currently distributed in geographic regions directly surrounding the Arabian Peninsula, including *Amietophrynus* (Africa), *Bufotes* (Northern Eurasia and North Africa), and *Duttaphrynus* (Southwest and South Asia), allowing the possibility of multiple biogeographic origins for the Arabian species.

Due to a lack of molecular sampling, the assignment of the Arabian bufonid species to genera, and therefore biogeographic origin, has long been problematic. One Arabian toad species, *Bufotes* cf. *variabilis*, is attributable to the Western Eurasian *Bufotes viridis* species complex on the basis of morphology (Balletto et al. 1985; Stöck et al. 2006). However, the affinities of the other Arabian toads are unresolved in part due to the similar morphology among the many species of toads occurring in regions surrounding the Arabian Peninsula. Based on molecular evidence the Arabian endemic *Duttaphrynus dhufarensis* was recently determined to be of Asian origin, with its closest relatives occurring on the Indian subcontinent (Van Bocxlaer et al. 2009), although several taxa occurring across Southwest Asia were not included in the analysis. This discovery, and the assignment of *D. dodsoni* (a species distributed throughout the Horn of Africa) to the genus (Van Bocxlaer et al. 2009; Frost 2014), implies the Arabian Peninsula has acted as a stepping-stone for Asian-derived bufonid species to colonize

mainland Africa. This scenario remains to be tested in a phylogenetic framework, but has important biogeographic implications. Other than *B*. cf. *variabilis* and *D. dhufarensis*, the remaining Arabian toads could not be allocated to genera, and remain in a non-taxon *Bufo* that is currently polyphyletic and is denoted as "*Bufo*". These species include *"Bufo" arabicus, "Bufo" tihamicus, "Bufo" hadramautinus*, and *"Bufo" scorteccii*. Both *"Bufo" arabicus* and *D. dhufarensis* possess the largest geographic distribution of the Arabian toads, whereas *"Bufo" tihamicus* is only distributed along the coastal Red Sea region. Although *"Bufo" tihamicus* is thought to be a close relative of the Sahelian distributed African taxon *"Bufo" pentoni*, the latter species is also currently not assigned to a genus, and therefore the phylogenetic relationships of this complex remains convoluted. The species *"Bufo" hadramautinus* and *scorteccii* are restricted to one or two localities, and their validity has been questioned as they may represent isolated phenotypically variable populations of the wider ranging *"Bufo" arabicus* and *D. dhufarensis*, respectively (Balletto et al. 1985). As a result of these hypotheses "*Bufo*" *scorteccii* has been tentatively assigned to the genus *Duttaphrynus* (Frost 2014), although this taxonomic assignment is certainly premature, and some authors have regarded *"Bufo" hadramautinus* as a synonym of *"Bufo" arabicus* (Schätti and Desvoignes 1999).

The evolution of Arabian bufonids and their close relatives remains a challenging biogeographic and taxonomic problem, and it is unknown if the unassigned Arabian toads are derived from African, Southwest Asian, or Western Eurasian lineages, or are the result of in-situ diversification on the Peninsula. In addition, the biogeographic patterns of the genus *Duttaphrynus* have not been investigated yet have important implications concerning whether amphibians successfully colonized the Horn of Africa using the Arabian Peninsula as a stepping-stone. A similar problem persists for *"Bufo" tihamicus* and *pentoni*, which could be derived from African, Southwest Asian, or Western Eurasian clades, each of which has implies different biogeographic scenarios for explaining the current geographic distributions of these two species.

Through several years of fieldwork, TJP obtained key samples of Arabian taxa (including *"Bufo" arabicus, "Bufo" tihamicus, D. dhufarensis*) as well as many biogeographically important species surrounding the Arabian Peninsula (including "*Bufo*" *pentoni* and *D. dodsoni*), finally allowing an assessment of their evolutionary relationships using a molecular phylogenetic framework. We seek to untangle the various biogeographic scenarios resulting in the evolution of the Arabian bufonids and their close relatives. We aim to identify the closest relatives of *"Bufo" tihamicus* and *"Bufo" arabicus* to determine if they are derived from African, Western Eurasian, or South Asian lineages. If they are African in origin, we predict they will have diverged from their closest mainland relatives either 1) following the formation of the Red Sea (23 Ma), or 2) as a result of dispersal across a southern Afro-Arabian land bridge (10–5.3 Ma). Alternatively, if they are Western Eurasian or South Asian in origin, we predict they will have colonized the Arabian Peninsula following establishment of Eurasian land bridge connections (18–15 Ma). We also seek to test the stepping-stone colonization of the Horn of Africa by the genus *Duttaphrynus*, and predict a sister-taxon relationship between *D. dodsoni* and *D. dhufarensis,* with a divergence time following permanent Eurasian land bridge connections (15 Ma). We use newly generated multi-locus molecular data and a comprehensive data set of published bufonid sequences to reconstruct phylogenetic relationships to identify the origins of the Arabian taxa. We use these molecular data to

estimate divergence times to identify key biogeographic events that could have led to the formation of these lineages. We examine our results in the context of prior studies that have investigated historical biogeographic patterns in other Arabian taxa, and highlight areas requiring further study.

Methods

Taxon Sampling

Due to the uncertainty of the phylogenetic placement of Arabian bufonid taxa, we sampled broadly across the family Bufonidae with an emphasis on regions directly surrounding the Arabian Peninsula (Table 2). We generated new multi-locus sequence data for 114 bufonid samples from 21 recognized species distributed across the Horn of Africa, the Arabian Peninsula, and Southwest Asia (Table 19, Appendix 6).

DNA extraction and amplification

Whole genomic DNA was extracted from liver samples using a high-salt DNA extraction (Aljanabi and Martinez 1997). We obtained a combination of sequence data from two mitochondrial markers, NADH dehydrogenase subunit 2 (*ND2*) and 16S ribosomal RNA (16S), and partial exonic sequence from the nuclear markers Recombination activating gene 1 (RAG1). The three loci (ND2, 16S, and RAG1) were amplified using the primer pairs 16SA and 16SB (Palumbi et al. 1991) for the 16S rRNA partial gene fragment, MET F1 L4437 and TRP R3 (Macey et al. 1997) for the ND2 partial gene fragment, and the primer pairs MartF1 and AmpR1 (Hoegg et al. 2004) were used to amplify RAG1.

Polymerase chain reactions (PCRs) were carried out in 12.5 µl volumes consisting of: 1.25 µl Roche 10x (500 mM Tris/HCl, 100 mM KCl, 50 mM $(NH_4)_2$ SO_4, 20 mM $MgCl_2$, pH = 8.3), 0.75 µl 25 mM $MgCl_2$, 0.75 µl 2 mM DNTPs, 0.25 µl 10.0 µM forward primer, 0.25 µl 10.0 µM reverse primer, 8.40 µl H_2O, 0.10 µl Taq, and 0.75 µl DNA. Amplification of both *ND2* and *16S* involved initial denaturation at 94°C for 4 min, followed by 35 cycles of 95°C for 60 s, 51°C for 60 s, 72°C for 90 s, and a final extension at 72°C for 7 min. The amplification of RAG1 followed the extended touchdown gradient reported by (Stöck et al. 2006), and involved initial denaturation at 95°C for 4 min, followed by a first program of 15 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 90 s (decreasing annealing temperature by -1°C per cycle), then a second program consisting of 20 cycles of 95°C for 30 s, 45°C for 30 s, and 72°C for 60 s, with a final extension at 72°C for 10 min.

The PCR amplifications were visualized on an agarose gel and cleaned using ExoSAP-IT (USB). Gene products were sequenced using BigDye v3.1 on an ABI3730 (Applied Biosystems). Newly generated sequences were edited using Geneious Pro (Drummond et al. 2011). All newly generated sequences are deposited in GenBank (Accession numbers: KT031406–KT031518 [16S]; KT031519–KT031601 [ND2]; KT031602–KT031708 [RAG-1]).

GenBank Sampling

To provide the most updated bufonid phylogeny for placing our focal taxa, we included GenBank data of representatives of all available unique bufonids (234 species, 39 genera) and several outgroups (seven species) (Appendix 7). The resulting data matrix is largely based on the alignment produced by Pyron and Wiens (2011) for their comprehensive

analysis of amphibian sequence data (Dryad repository doi:10.5061/dryad.vd0m7). Their matrix was composed of 12 loci: nine nuclear genes consisting of C-X-C chemokine receptor type 4 (CXCR4), histone 3a (H3A), sodium–calcium exchanger (NCX1), pro-opiomelanocortin (POMC), recombination-activating gene 1 (RAG1), rhodopsin (RHOD), seventh-in-absentia (SIA), solute-carrier family 8 (SLC8A3), and tyrosinase (TYR); and three mtDNA loci including cytochrome b (cyt-b), and the large and small sub-units of the mitochondrial ribosome genes (12S/16S; without tRNAs). We conducted GenBank searches for additional taxa not included in this data matrix, and we added sequence data for NADH dehydrogenase subunit 2 (ND2), which has been sequenced for many bufonid taxa. The final data matrix containing both GenBank and newly generated sequence data consists of 243 bufonid lineages and 7 outgroup species. The taxonomy of the family Bufonidae has been under revision, and names of several genera have been changed to reflect evolutionary relationships (AmphibiaWeb 2014; Frost et al. 2006). Species names were updated accordingly and are here presented using the newest taxonomy from AmphibiaWeb and the Amphibian Species of the World v6.0 (Frost 2014). GenBank numbers for all sequence data included are given in Appendix 7.

Sequence Alignment

All protein-coding genes were aligned using MUSCLE (Edgar 2004), and subsequently translated to ensure conservation of reading frame. The 12S and 16S sequences were initially aligned using Clustal Omega (Sievers et al. 2011), manually adjusted by eye, and poorly aligned regions were trimmed from the alignment. Trimmed sequences were then realigned using Clustal Omega, with some slight manual adjustments. The final concatenated alignment consists of 13 loci, 250 taxa, and 10,492 base pairs. The matrix is composed of the following data: 12S, 215 sequences (85%, 1,011bp); 16S, 236 sequences (93%, 1,223 bp); cyt-b, 91 sequences (36%, 1,122 bp); ND2, 78 sequences (30%, 1,035 bp); CXCR4, 114 sequences (45%, 732 bp); H3A, 38 sequences (15%, 328 bp); NCX1, 60 sequences (23%, 1,275 bp); POMC, 69 sequences (27%, 550 bp); RAG1, 101 sequences (40%, 840 bp); RHOD, 38 sequences (15%, 315 bp); SIA, 40 sequences (16%, 397 bp); SLC8A3, 14 sequences (5%, 1,132 bp); and TYR, 12 sequences (4%, 532 bp). The mean sequence length is 3,185 bp, and the range in length across taxa is 375 to 7,886 bp. The proportion of missing data across the matrix is approximately 70%.

Phylogenetic Analyses

We used PartitionFinder to simultaneously determine our best partitioning strategy and models for each partition subset (Lanfear et al. 2012). The greedy search algorithm was employed, and model selection was conducted using the Bayesian information criterion (BIC). Due to the large size and complexity of the molecular data set, we did not allow for partitioning of genes by codon position. The best partitioning scheme of the full data set includes four gene partitions: 12S and 16S: GTR+G+I; ND2 and cyt-b: GTR+G+I; RHOD, SIA, TYR: HKY+G; and CXCR4, H3A, NCX1, POMC, RAG1, and SLC8A3: GTR+G+I. We conducted Bayesian analyses using MrBayes v3.2 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003), and parallel runs utilizing four MCMC chains were allowed to run for $2x10^7$ million generations, with sampling every 1000 generations. Runs were assessed using Tracer v1.6 (Rambaut and Drummond 2009) to ensure key parameters had reached stationarity (ESS values >150). The first 25% of the

total number of generations were discarded as burn-in and a maximum clade credibility tree was calculated from the remaining trees (30,000) using TreeAnnotator v1.8.1 (Drummond et al. 2012). We performed maximum likelihood analyses of the partitioned data set using GARLI v2.0 (Zwickl 2006). Using default parameters in the ML search algorithm, 10 replicate searches for the best point estimate topology were conducted, and the tree with the best likelihood was selected as a guide tree for the bootstrap analyses. The Garli Web Service (Bazinet and Cummings 2011), functioning on the Lattice Project grid system (Bazinet and Cummings 2008), was used to execute 1000 nonparametric bootstrap replicates asynchronously in parallel using the default stopping criteria. A maximum clade credibility tree was generated from the 1000 replicates using TreeAnnotator v1.8.1 (Rambaut and Drummond 2009).

Divergence Dating Analyses

To infer the timing of lineage divergences of sampled Arabian amphibian taxa, we carried out dating analyses in BEAST v1.8.1 ()Drummond et al. 2012). We included four internal calibration points, following recommendations of (Pramuk et al. 2008; Van Bocxlaer et al. 2009):

(A) A minimum age of 20 million years (Myr) for the split between North- and Central America based on the fossil *Bufo praevis* (Tihen 1951). This was enforced using a lognormal distribution with real space mean of 10, log(stdev) of 1, offset of 19, and initial value of 21, creating the following credibility interval: 5% = 20.1, 95% = 50.4.

(B) A minimum age of 18 Mya for the stem origin of toads belonging to the *Bufotes viridis* complex (Rage and Rocek 2003; Böhme 2003). This was enforced using a lognormal distribution with real space mean of 10, log(stdev) of 1, offset of 17, and initial value of 19, creating the following credibility interval: 5% = 18.1, 95% = 48.4.

(C) A minimum age of 11 Myr for the origin of the *Rhinella marina* group (sensu Maceil et al. 2010) based on a fossil from the Middle Miocene (Sanchíz 1998). This was enforced using a lognormal distribution with real space mean of 10, log(stdev) of 1, offset of 10, and initial value of 12, creating the following credibility interval: 5% = 11.1, 95% = 41.4.

(D) A minimum age of 9.6 Myr for the origin of toads belonging to the *Bufo bufo* group based on the appearance of a *Bufo bufo* fossil from the Miocene of Europe (Rage and Rocek 2003). This was enforced using a lognormal distribution with real space mean of 10, log(stdev) of 1, offset of 8, and initial value of 9, creating the following credibility interval: 5% = 9.1, 95% = 39.4.

In addition, we used a minimal constraint on the family Bufonidae for the oldest known bufonid fossils recovered from Paleocene deposits in both Brazil and France (Estes and Reig 1973; Estes and Báez 1985; Rage 2003). This was enforced using three different approaches: 1) a lognormal distribution with a real space mean of 15, log(stdev) of 1, offset of 55, and initial value of 65, producing a credibility interval of (5% = 56.1, 95% = 102.1); 2) a normal distribution with a mean of 80, a standard deviation of 14, and initial value of 65, creating a credibility interval of (5% = 56.9, 95% = 103); and 3) an exponential distribution with a mean of 20, offset of 57, and initial value of 60, producing a credibility interval of (5% = 58.0, 95% = 116.9).

Dating analyses were initially run using the full molecular data set of 13 loci and 253 taxa; analyses behaved very poorly and failed to converge even after several hundred

million generations. Therefore, the data set was reduced to include only the most well sampled genes: 12S, 16S, ND2, CXCR4, and RAG1. Data were grouped into three partitions: 1) 12S and 16S, 2) ND2, and 3) CXCR4 and RAG1. Inclusion of taxa in the reduced data set required partial sequences for at least two of these three partitions. The final alignment for all subsequent dating analyses included 132 taxa, 5 loci, and 4,840 bp. The average sequence length was 3,051 bp, with a range of 811 to 4,640 bp, and the alignment contained 37% missing data.

Dating analyses were run for $2x10^7$ generations with sampling every 2000 generations. For all analyses, we used the Yule model of speciation as our tree prior, applied an uncorrelated relaxed lognormal clock, and unlinked clock and substitution models. The partitioning scheme and substitution models are as follows: 12S and 16S: GTR+G+I; ND2: GTR+G+I; CXCR4 and RAG1: GTR+G+I. Runs were assessed using Tracer v1.6 (Rambaut and Drummond 2009) to examine convergence. A burn-in of 25% was discarded and maximum clade credibility trees were created from a total of 7,500 trees for each analysis.

To explore the effects of calibration choices, we ran multiple analyses with different combinations of calibrations enforced. Fixing the age of the Bufonidae, we explored every permutation of the internal calibrations resulting in 14 sets of analyses (Additional File 4). These 14 analyses revealed that various combinations of calibrations had little overall effect on dating results throughout the tree (not shown). We therefore focused on exploring the effects of the shape of the calibration prior for the age of the Bufonidae in our final dating analyses. This produced three analyses: (A1) all four internal calibrations with a lognormal prior for the Bufonidae, (A2) all four internal calibrations with a normal distribution prior for the Bufonidae, and (A3) all four internal calibrations with an exponential prior for the Bufonidae.

Results

Phylogenetic Relationships

The overall phylogenetic relationships within the family Bufonidae are concordant across analyses and are presented in Figure 25. We find strong support for the monophyly of the Bufonidae, and relationships among the basal genera (*Melanophryniscus, Dendrophryniscus, Osornophryne, Atelopus, Amazophrynella, Nannophryne, Peltophryne*) are well resolved and consistent with previous studies (Frost et al. 2006; Pramuk et al. 2008; Van Bocxlaer et al. 2009; Van Bocxlaer et al. 2010; Pyron and Wiens 2011). Our phylogenetic analyses place *Anaxyrus* as the sister clade to *Incilius* with high support, and together these two genera are moderately supported as being sister to the genus *Rhinella* (Fig. 25). Beyond this grouping, the relationships among the other derived genera (*Rhaebo, Didynamipus, Poyntonophrynus, Nimbaphrynoides, Vandijkophrynus, Capensibufo, Mertensophryne, Amietophrynus, Wolterstorffina, Werneria, Nectophryne, Barbarophryne, Schismaderma, Churamiti, Nectophrynoides, Pedobistes, Adenomus, Xanthophryne, Duttaphrynus, Bufotes, Epidalea, Strauchbufo, Sabahphrynus, Bufo, Leptophryne, Ingerophrynus, Ghatophryne, Phrynoidis, Pelophryne, Ansonia*) are largely unresolved (Fig. 25). Genera sampled for more than one lineage are supported as monophyletic, with the exception of *Pedobistes* (as found by Van Bocxlaer et al. 2010). This is true for the speciose genera surrounding the Arabian Peninsula, including *Amietophrynus*, *Bufotes*, and *Duttaphrynus*, which are all recovered as independent

monophyletic groups with strong support (Fig. 25). In this global bufonid species data set with an emphasis on sampling surrounding the Arabian Peninsula, our analyses consistently place the Arabian toad species within particular genera.

The Arabian species "*Bufo*" *tihamicus* is strongly supported as the sister taxon of the African Sahelian distributed "*Bufo*" *pentoni*, and together these two lineages are recovered as sister to all other species of *Amietophrynus* (Figs. 25, 26). Our analyses recover "*Bufo*" *arabicus* as the sister taxon to a grouping of two clades. One clade consists of the more arid-adapted species *A. xeros* and *A. gutturalis*, and the other consists of *A. tuberosus*, *A. garmani*, *A. camerunensis*, *A. kisolensis*, and *A. gracilipes*. The resolution of the relationship between "*Bufo*" *arabicus* and these two clades is not strongly supported, with some results placing "*Bufo*" *arabicus* as sister to *A. xeros* and *A. gutturalis,* and others placing "*Bufo*" *arabicus* as sister to both clades (as in Fig. 25). Our results confirm the findings of Van Bocxlaer et al. (2009) in recovering the Arabian species *Duttaphrynus dhufarensis* within the Asian genus *Duttaphrynus*. Our improved sampling from Iran and Pakistan places *D. dhufarensis* as sister to a clade consisting of *D. hololius*, *D. stomaticus* and *D. olivaceus* with strong support (Figs. 25, 27). *Duttaphrynus dodsoni*, which is distributed across the Horn of Africa, is not recovered as a member of the genus *Duttaphrynus*, but rather is found deeply nested in the genus *Amietophrynus* and is a close relative of the East African species *Amietophrynus brauni* (Figs. 25, 26).

In addition to elucidating the relationships of Arabian bufonids, our phylogenetic analyses revealed a unique bufonid lineage restricted to the Horn of Africa, designated as "*Bufo*" sp. (Fig. 25, "Undescribed lineage"). The placement of this lineage is not well resolved and it does not appear to be a member of any currently described African genus (Fig. 25). Maximum likelihood and Bayesian consensus trees show this new genus is most closely related to a clade of African genera (*Capensibufo*, *Mertensophryne, Vandijkophrynus*, and *Nimbaphrynoides*) but this relationship is not strongly supported (Fig. 25). An additional extensive comparison with over 500 unpublished 16S sequences of mainly African bufonids also failed to resolve the placement of this taxon with any described genus (H.C. Liedtke, pers. comm.), and additional taxonomic work is therefore required for this unique lineage.

Divergence Dating Estimates

The three divergence dating analyses (A1, A2, A3) produced largely congruent phylogenetic topologies. Divergence dating estimates varied little between analyses, though the 95% highest posterior density regions are slightly broader in analyses A1 and A3 than those in A2 (Table 18). All dating analyses suggest a mid-Cretaceous origin of the family Bufonidae, with median estimates ranging between 93.5–100.3 Ma ([A1: 94.9 Ma, 95% HPD 64.6–129.8]; [A2: 93.5, 73.0–114.0 Ma]; [A3: 100.3, 70.1–133.1 Ma]. These dates are consistent with Pramuk et al. (2008), who also recover a mid-Cretaceous origin of the family with an estimate of 88.2 Ma (78.3–98.8 Ma), but are approximately 20–25 million years older than dates recovered by Van Bocxlaer et al. (2010) for the group (67.9 Ma, 95% HPD 52.7–92.7]. Similar to both Pramuk et al. (2008) and Van Bocxlaer et al. (2010), we find the origin and diversification of most derived (eg. formerly *Bufo*) bufonid genera occurred in a window of approximately 10 million years.

Our estimates place this rapid diversification in the Eocene, as found by Pramuk et al. (2008), though Van Bocxlaer et al. (2010) date this radiation to the Oligocene.

Based on our data, diversification of the genus *Amietophrynus* began around the Eocene-Oligocene boundary ([A1: 37.6, 25.6–51.5 Ma]; [A2: 36.5, 28.1–45.5 Ma]; [A3: 37.2, 27.1–50.4 Ma]) whereas comparatively diversification began more recently in *Duttaphrynus* during the mid-Oligocene ([A1: 29.3, 19.3–40.5 Ma]; [A2: 27.8, 20.9–35.1 Ma]; [A3: 29.2, 20.4–40.0 Ma]) (Table 18). The divergences of the Arabian species (*"Bufo" tihamicus* and *arabicus*, *D. dhufarensis*) are estimated to have occurred during the Miocene. An Early Miocene divergence is estimated between *"Bufo" tihamicus* and *"Bufo" pentoni* ([A1: 17.9, 9.0–26.6 Ma]; [A2: 16.8, 9.2–25.4 Ma]; [A3: 17.0, 9.3–25.9 Ma]) as well as the split between *"Bufo" arabicus* and its sister clade ([A1: 21.2, 14.2–29.2 Ma]; [A2: 20.5, 15.1–26.3 Ma]; [A3: 20.7, 14.8–29.2 Ma]) (Fig. 25, Table 18). Within *"Bufo" arabicus*, the allopatric populations occurring in Yemen and Oman are supported as genetically distinct and are estimated to have diverged in the Pliocene ([A1: 4.0, 1.5–6.7 Ma]; [A2: 3.5, 1.6–6.1 Ma]; [A3: 3.5, 1.5–7.3 Ma]) (Fig. 26). The divergence estimates for the split between *Duttaphrynus dhufarensis* and the clade consisting of *D. hololius*, *D. stomaticus* and *D. olivaceus* occur in the mid-Miocene, with median estimates ranging from 12.9–13.5 Ma ([A1: 13.5, 7.9–20.9 Ma]; [A2: 12.9, 8.2–18.6 Ma]; [A3: 13.5, 8.4–20.1 Ma]) (Fig. 27). These dates are slightly older than those recovered by Van Bocxlaer et al. (2009), who estimate this divergence at 8.5 Ma (95% HPD 5.9–12.3). Previously untested, the divergence event between closely related *D. stomaticus* and *D. olivaceus* is estimated to have occurred in the late Pliocene ([A1: 3.0, 1.0–5.1 Ma]; [A2: 2.6, 1.1–4.7 Ma]; [A3: 2.7, 1.0–5.2 Ma]; Fig. 27).

Discussion

Evolutionary relationships of Arabian bufonids

The relationships among Arabian toad species have long been problematic, owing in part to the generalized morphology of these species and of lineages in surrounding regions (Balletto et al. 1985). The molecular delimitation of at least three speciose toad genera differentially distributed around the Arabian Peninsula (*Amietophrynus, Bufotes, Duttaphrynus*) further highlighted the possibility of several alternative biogeographic origins for the Arabian species (Frost et al. 2006). Prior to our study, the Arabian bufonid species assemblage was recognized as having both Western Eurasian and Asian elements (*Bufotes* cf. *variabilis* (Balletto et al. 1985; Stöck et al. 2006), *Duttaphrynus dhufarensis* (Van Bocxlaer et al. 2009)), and here we identify a previously unrecognized, though not unexpected, African component. We find strong support for a sister relationship between "*Bufo*" *pentoni* and "*Bufo*" *tihamicus*, which together form a monophyletic assemblage with all other species in the African genus *Amietophrynus* (Figs. 25, 26). The Arabian lineage "*Bufo*" *arabicus* is also recovered in this clade (Figs. 25, 26), and our results do not support a close relationship between these Arabian lineages. Rather, these species have independent evolutionary origins in the genus.

We recover the Arabian species *Duttaphrynus dhufarensis* as part of a largely South Asian clade, represented by other species in the genus *Duttaphrynus* (Figs. 25, 27), consistent with Van Bocxlaer et al. (2009). In their study, Van Bocxlaer et al. (2009) recover *D. dhufarensis* as sister to a clade containing *D. stomaticus* and *D. hololius*. They report *D. hololius* as widespread on the Indian subcontinent, with *D. stomaticus* being

restricted to the Western Ghats (their Fig. 2). The true distributions of these species are actually converse, though their interpretation of the biogeographic results is sound. Regardless, both species were sampled from India, and together with the other species included in their analyses the westernmost sampling for the genus was limited to India. This left a vast sampling gap across Pakistan and Iran, a region that contains many widespread species in the genus *Duttaphrynus* and *Bufotes*. Our sampling includes six species from this region (*Bufotes oblongus, B. pseudoraddei*, *B. surdus, B. variabilis*, *D. olivaceus*, and *D. stomaticus*) and includes geographically relevant localities adjacent to the Gulf of Oman and Persian Gulf in Iran and Pakistan. With this biogeographically improved sampling, we find *D. dhufarensis* is the sister taxon to a clade containing *D. hololius*, *D. olivaceus* and *D. stomaticus*. We find *D. stomaticus* exhibits a distribution throughout Pakistan and India, and the *D. olivaceus* is distributed largely throughout Iran and into Pakistan (Figs. 25, 27).

With our broad species sampling of *Duttaphrynus*, *Bufotes*, and *Amietophrynus*, our analyses do not recover *Duttaphrynus dodsoni* as monophyletic with other members of the genus. Instead, this species is strongly supported as nested within the African distributed genus *Amietophrynus* as sister taxon to the East African species *Amietophrynus brauni* (Figs. 25, 26). We discuss the biogeographic implications of this discovery and the overall historical biogeography of the Arabian toad species below.

Historical biogeography of Arabian toads

Our phylogenetic analyses have resolved the origins of several biogeographically interesting toad lineages distributed throughout the Horn of Africa, the Arabian Peninsula, and Southwest Asia, and here we interpret major divergence events in the context of the geological history of this region.

The species *"Bufo" arabicus* and *"Bufo" tihamicus* are both estimated to have diverged from their closest African relatives in the Early Miocene, approximately 20.5–21.2 Ma (95% HPD range: 14.2–29.2 Ma) and 16.8–17.9 Ma (95% HPD range: 9.0–26.6 Ma), respectively (Fig. 26, Table 18). These timings are remarkably concordant considering these two lineages have evolved independently within the genus. The age estimates are quite ancient and rule out dispersal across a southern Afro-Arabian landbridge (10–5.3 Ma) and overwater dispersal (5.3 Ma to present), but are consistent with the separation of the Arabian Plate from mainland Africa as a result of the Red Sea formation during the Oligocene-Miocene boundary (27–23 Ma) (Fig. 3) (Bohannon 1986; Bosworth et al. 2005). The landbridge spanning the Strait of Bab el Mandeb, which created a dispersal route between Arabian Peninsula and the Horn of Africa, was not established until 10–5.3 Ma. Many of the African-derived terrestrial vertebrate taxa present in the southwestern Arabian Peninsula arrived by dispersing from the Horn of Africa beginning in the late Miocene via a southern land bridge or overwater dispersal events (Pook et al. 2009; Portik and Papenfuss 2012; Wong et al. 2010; Zinner et al. 2009) rather than originating as a result of the initial separation of the Arabian plate from Africa. This classifies the Arabian toad species *"Bufo" arabicus* and *"Bufo" tihamicus* as true African relicts, a pattern that has only been demonstrated for one other taxon, the viper species *Echis coloratus*, which also diverged in the Early Miocene (Pook et al. 2009).

The Asian derived species *Duttaphrynus dhufarensis* is estimated to have diverged in the mid-Miocene, 12.9–13.5 Ma (95% HPD range: 7.9–20.9 Ma) (Fig. 27, Table 18). The temporary connections between the Arabian plate and Eurasia established 18–16 Ma are thought to have allowed the first faunal exchanges between these distinct biogeographic regions, with a permanent Eurasian connection being established ~15 Ma (Bohannon 1986; Bosworth et al. 2005; Rögl 1998, 1999). The divergence estimates for *D. dhufarensis* are congruent with the temporal range of these events, and the establishment of land bridges likely served as dispersal routes for this lineage to colonize the Arabian Peninsula.

Following a stepping stone model of dispersal, if the genus *Duttaphrynus* occurred on mainland Africa, the temporal origins of these lineages would be expected to be younger than the age of *D. dhufarensis* (13 Ma; 7.9–20.9 Ma). However, with the discovery of the phylogenetic placement of *Duttaphrynus dodsoni* in the genus *Amietophrynus*, we find no support for the genus occurring on the Horn of Africa, indicating South Asian-derived bufonid lineages did not successfully complete trans-continental dispersals across the Arabian Peninsula. Additionally, based on our extensive sampling of the Horn of Africa, Arabian Peninsula, and South Asia, we find no evidence for the stepping-stone model of dispersal across the Arabian Peninsula for any African (*Amietophrynus*) or Asian (*Duttaphrynus*) bufonid species (Figs. 25-27), as no species of *Amietophrynus* are found in Iran or Pakistan. This is somewhat unexpected because this model has been invoked to explain ancient Asiatic-African dispersals in multiple groups of ranoid frogs (Kosuch et al. 2001). Additionally, there is evidence suggesting the Persian Gulf region was greatly reduced or dry in the Pleistocene (Lambeck 1996; Uchupi et al. 1999). The apparent lack of dispersals of toad lineages across the Persian Gulf region during this time period is surprising, as population exchanges of the viper species *Echis carinatus* likely occurred through this route (Pook et al. 2009).

Taxonomic Implications

Based on our sampling strategy and the phylogenetic placement of key species, we recommend several taxonomic changes. The species "*Bufo*" *tihamicus* and "*Bufo*" *pentoni* form a monophyletic group with all other *Amietophrynus*, and although they represent the most basal divergence in the group we recommend assignment to the genus, and recognize *Amietophrynus tihamicus* comb. nov. and *Amietophrynus pentoni* comb. nov. The assignment of these two species can be further tested through karyotyping, as the 20-chromosome condition is considered apomorphic for this group (Cunningham and Cherry 2004; Bogart 1968). If an alternative chromosome condition is discovered in these two species, additional insight into chromosome evolution among bufonids will be gained and their generic assignment can be reconsidered. The Arabian species "*Bufo*" *arabicus* can be confidently assigned to the genus *Amietophrynus*, as *Amietophrynus arabicus* comb. nov., and *Duttaphrynus dodsoni* is also transferred to the genus as *Amietophrynus dodsoni* comb. nov.

The origins of several Arabian bufonids have been investigated, however several lineages require further investigation. The Arabian population of *Bufotes* cf. *variabilis* warrants further study to determine its distinctiveness with respect to *B. variabilis* sensu stricto. This population shares a similar distribution with *Hyla felixarabica*, a recently described Arabian species that was previously thought to be an isolate of a wider ranging

Western Eurasian species complex (*Hyla arborea*) (Gvoždík et al. 2010). Additionally, the relationships of *"Bufo" hadramautinus* and *"Bufo" scorteccii* remain speculative. Although *"Bufo" hadramautinus* is morphologically similar to *Amietophrynus arabicus* (Balletto et al. 1985), *"Bufo" scorteccii* is morphologically intermediate between *A. arabicus* and *D. dhufarensis* (Balletto et al. 1985). On this basis, *"Bufo" hadramautinus* can be tentatively recognized as *Amietophrynus hadramautinus* comb. nov., but *"Bufo" scorteccii* remains problematic and should remain unassigned until further study. Although inconvenient, there are major biogeographic implications associated with assignment to particular genera, and this action would circumvent such issues. When population-level sampling becomes available, additional phylogenetic work can clarify if these two geographically restricted lineages are: 1) intraspecific populations of one of the widespread Arabian bufonid species, 2) distinct lineages derived independently from surrounding continental faunas, or 3) distinct lineages resulting from in-situ speciation of an Arabian species, and further taxonomic assessments can be made.

Conclusions

The Arabian Peninsula is home to a unique fauna that has assembled and evolved throughout the course of major geophysical events. The Arabian species assemblage represents an admixture of African, Western Eurasian, and South Asian elements, and this pattern is exemplified even in the relatively depauperate Arabian bufonids. In addition to having South Asian and Western Eurasian lineages, we have identified two lineages independently derived from continental Africa. Our dating estimates strongly suggest *Amietophrynus arabicus* and *A. tihamicus* did not colonize the Arabian Peninsula through overwater dispersal or a southern land bridge from the Horn of Africa, rather the formation of the Red Sea likely drove simultaneous divergences in these species. In this sense, they represent true African relicts in their current distribution on the Arabian Peninsula. More importantly, across all dating analyses the relative timing of divergence for these species is considerably older than that for the South Asian derived *Duttaphrynus dhufarensis*. These results conform to predictions based on geological events that species dispersing to the Arabian Peninsula across Eurasian land bridges should be younger in origin than true African relicts. Our investigation has revealed the stepping-stone hypothesis for trans-continental Afro-Asian bufonid dispersals is not accurate, and we find no evidence for *Amietophrynus* or *Duttaphrynus* species distributed outside of their main continental range and the Arabian Peninsula. Further studies of the remaining Arabian amphibian species can test if these biogeographic scenarios hold true for not only bufonids but also other anuran families. These amphibian studies and comparative studies of other terrestrial vertebrates can provide a clearer picture of the diversification of the unique faunal assemblage present on the Arabian Peninsula.

Acknowledgments

T.J.P. thanks Ali Al Kiyumi, Ministry of Environment and Climate Affairs issued Oman for collecting and export permits, the American Institute for Yemeni Studies for facilitating fieldwork in Yemen, Abdurahman A. Osman, Director of the Puntland Development Research Center for arranging fieldwork in Puntland State, Somalia, Suleiman Ahmed Gulaid, President of Amoud University, Borama, Somaliland for facilitating fieldwork in Somaliland. Partial funding for T.J.P's fieldwork was provided

by a grant from the George Lindsay Field Research Fund of the California Academy of Sciences. Molecular work was funded in part by the Museum of Vertebrate Zoology of the University of California, Berkeley. D.M.P. was funded by teaching assistantships awarded by the Integrative Biology Department of the University of California, Berkeley. We thank J.A. McGuire, D.C. Blackburn, R.C.K. Bowie, S.M. Rovito, S. Hykin, P. Skipwith, and R. von May for comments on a prior version of a combined manuscript, R.C. Bell for comments on this manuscript, and H.C. Liedtke for comparisons of unpublished bufonid molecular data. We thank Václav Gvoždík, Ines Van Bocxlaer, and one anonymous reviewer for helpful comments on the first submission and resubmission of this manuscript.

References

Aljanabi, S., and I. Martinez. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research 25:4692–4693.

Amer, S.A.M., and Y. Kumazawa. 2005. Mitochondrial DNA sequences of the Afro-Arabian spiny-tailed lizards (genus *Uromastyx*; family Agamidae): phylogenetic analyses and evolution of gene arrangements. Biological Journal of the Linnean Society 85:247–260.

AmphibiaWeb. 2014. Information on amphibian biology and conservation. AmphibiaWeb. Berkeley, California: [http://amphibiaweb.org]

Balletto, E., M.A. Cherchi, and J. Gasperetti. 1985. Amphibians of the Arabian Peninsula. Fauna of Saudi Arabia, vol. 7. Edited by Büttiker W, Krupp F. Basle: Pro Entomologia, Natural History Museum; 318–392.

Bazinet, A.L., and M.P. Cummings. 2011. Computing the Tree of Life — Leveraging the power of desktop and service grids (IPDPSW). Pp. 1896–1902 in 2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum.

Bazinet, A.L., and M.P. Cummings. 2008. The Lattice Project: a grid research and production environment combining multiple grid computing models. Pp. 2–13 in M.H.W. Weber (ed.), Distributed & Grid Computing - Science Made Transparent for Everyone. Principles, Applications and Supporting Communities. Marburg: Rechenkraft.net.

Bogart, J.P. 1968. Chromosome number difference in the amphibian genus *Bufo*: the *Bufo regularis* species group. Evolution 22:42–45.

Bohannon, R.G. 1986. Tectonic configuration of the western Arabian continental margin, southern Red Sea. Tectonics 5:477–499.

Böhme, M. 2003. The Miocene climatic optimum: Evidence from ectothermic vertebrates of Central Europe. Palaeogeography Palaeoecology 195:389–401.

Bosworth, W., P. Huchon, and K. McClay. 2005. The Red Sea and Gulf of Aden Basins. Journal of African Earth Sciences 43:334–378.

Cunningham, M., and M.I. Cherry. 2004. Molecular systematics of African 20-chromosome toads (Anura: Bufonidae). Molecular Phylogenetics and Evolution 32:671–685.

deMenocal, P.B. 1995. Plio-Pleistocene African Climate. Science 270:53–59.

deMenocal, P.B. 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. Earth and Planetary Science Letters 220:3–24.

Drummond, A.J., B. Ashton, S. Buxton, M. Cheung, A. Cooper, C. Duran, M. Field, J. Heled, M. Kearse, S. Markowitz, R. Moir, S. Stones-Havas, S. Sturrock, T. Thierer, and A. Wilson. 2011. Geneious v5.4. [http://www.geneious.com]

Drummond, A.J., M.A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29:1969–1973.

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

Estes, R., and A.M. Báez. 1985. Herpetofaunas of North and South America during the Late Cretaceous and Cenozoic: Evidence for interchange? Pp. 139–197 in F.G.

Stehli and S.D. Webb (eds.), The Great American Biotic Interchange. New York: Springer Science and Business Media LLC.
Estes, R., O.A. Reig. 1973. The early fossil record of frogs: A review of the evidence. Pp. 11–63 in J.L. Vial (ed.), Evolutionary Biology of the Anurans. Columbia: University of Missouri Press.
Fernandes, C.A., E.J. Rohling, and M. Siddall. 2006. Absence of post-Miocene Red Sea land bridges: biogeographic implications. Journal of Biogeography 33:961–966.
Fisher, M., and D.A. Membery. 1998. Climate. Pp. 5–38 in S.A. Ghazanfar, F.M. Dordrecht (eds.), Vegetation of the Arabian Peninsula. Kluwer Academic Publishers.
Frost, D.R. 2014. Amphibian Species of the World: an Online Reference. Version 6.0. American Museum of Natural History, New York: [http://research.amnh.org/herpetology/amphibia/index.html]
Frost, D.R., T. Grant, J. Faivovich, R.H. Bain, A. Haas, C.F.B. Haddad, A. Channing, M. Wilkinson, S.C. Donnellan, C.J. Raxworthy, J.A. Campbell, B.L. Blotto, P. Moler, R.C. Drewes, R.A. Nussbaum, J.D. Lynch, D.M. Green, and W.C. Wheeler. 2006. The amphibian tree of life. Bulletin of the American Museum of Natural History 297:1–370.
Gasperetti, J. 1988. Snakes of Arabia. Pp. 169–450 in W. Büttiker, K.F. Basle (eds.), Fauna of Saudi Arabia, Vol. 9. Pro Entomologia, Natural History Museum.
Guba, I., and K. Glennie. 1998. Geology and geomorphology. Pp. 39–62 in S.A. Ghazanfar, F.M. Dordrecht (eds.), Vegetation of the Arabian Peninsula. Kluwer Academic Publishers.
Gvoždík, V., J. Moravec, C. Klüsch, and P. Kotlík. 2010. Phylogeography of the Middle Eastern tree frogs (*Hyla*, Hylidae, Amphibia) as inferred from nuclear and mitochondrial DNA variation, with a description of a new species. Molecular Phylogenetics and Evolution 55:1146–1166.
Hoegg, S., M. Vences, H. Brinkmann, and A. Meyer. 2004. Phylogeny and comparative substitution rates of frogs inferred from sequences of three nuclear genes. Molecular Biology and Evolution 21:1188–1200.
Huelsenbeck, J.P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.
Joger, U. 1987. An interpretation of reptile zoogeography in Arabia, with special reference to Arabian herpetofaunal relations with Africa. Pp. 257–271 in F. Krupp, W. Schneider, K. Göttingen (eds.), Proceedings of the Symposium on the Fauna and Zoogeography of the Middle East. R. Hubert & Co.
Kosuch, J., M. Vences, A. Dubois, A. Ohler, and W. Böhme. 2001. Mitochondrial DNA evidence for an Oriental origin of Tiger Frogs, genus *Hoplobatrachus*. Molecular Phylogenetics and Evolution 21:398–407.
Kürschner, H. 1998. Biogeography and introduction to vegetation. Pp. 63–98 in S.A. Ghazanfar, F.M. Dordrecht (eds.), Vegetation of the Arabian Peninsula. Kluwer Academic Publishers.
Lambeck, K. 1996. Shoreline reconstructions for the Persian Gulf since the last glacial maximum. Earth and Planetary Science Letters 142:43–57.

Lanfear, R., B. Calcott, S.Y.W. Ho, and S. Guindon.: 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29:1695–1701.

Macey, R.J., A. Larson, N.B. Ananjeva, Z. Fang, and T.J. Papenfuss. 1997. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Molecular Biology and Evolution 14:91–104.

Maciel, N.M., R.G. Collevatti, G.R. Colli, and E.F. Schwartz. 2010. Late Miocene diversification and phylogenetic relationships of the huge toads in the *Rhinella marina* (Linnaeus, 1758) species group (Anura: Bufonidae). Molecular Phylogenetics and Evolution 57:787–797.

Mallon, D.P. 2011. Global hotspots in the Arabian Peninsula. Zoology in the Middle East 54: Supplementum 3, 13–20.

Meister, J., M.A. Hubaishan, N. Kilian, and C. Oberprieler. 2005. Chloroplast DNA variation in the shrub *Justicia areysiana* (Acanthaceae) endemic to the monsoon affected coastal mountains of the southern Arabian Peninsula. Botanical Journal of the Linnean Society 148:437–444.

Meister, J., M.A. Hubaishan, N. Kilian, and C. Oberprieler. 2006. Temporal and spatial diversification of the shrub *Justicia areysiana* Deflers (Acanthaceae) endemic to the monsoon affected coastal mountains of the southern Arabian Peninsula. Plant Systematics and Evolution 262:153–171.

Metallinou, M., N.E. Arnold, P.A. Crochet, P. Geniez, J.C. Brito, P. Lymberakis, S.B. El Din, R. Sindaco, M. Robinson, and S. Carranza. 2012. Conquering the Sahara and Arabian deserts: Systematics and biogeography of *Stenodactylus* geckos (Reptilia: Gekkonidae). BMC Evolutionary Biology 12:258.

Miller, A.G., and T.A. Cope. 1996. Flora of the Arabian Peninsula and Socotra, vol. 1. Edinburgh: Edinburgh University Press.

Mittermeier, R.A., N. Myers, P.R. Gil and C.G. Mittermeier. 1999. Hotspots: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions. Monterrey, Mexico: Cemex, Conservation International and Agrupacion Sierra Madre.

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

Palumbi, S., A. Martin, S. Romano, W.O. McMillan, L. Stice, and G. Grabowski. 1991. The Simple Fool's Guide to PCR. Version 2. Honolulu: University of Hawaii.

Pook, C.E., U. Joger, N. Stümpel, and W. Wüster. 2009. When continents collide: phylogeny, historical biogeography and systematics of the medically important viper genus *Echis* (Squamata: Serpentes: Viperidae). Molecular Phylogenetics and Evolution 53:792– 807.

Portik, D.M., and T.J. Papenfuss. 2012. Monitors cross the Red Sea: the biogeographic history of *Varanus yemenensis*. Molecular Phylogenetics and Evolution 62:561–565.

Pramuk, J.B., T. Robertson, J.W. Sites, and B.P. Noonan. 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, R.A., and J.J. Wiens. 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.

Rage, J.-L.: 2003. Oldest Bufonidae (Amphibia, Anura) from the Old World: A bufonid from the Paleocene of France. Journal of Vertebrate Paleontology 23:462–463.
Rage, J.-C., and Z. Rocek. 2003. Evolution of anuran assemblages in the Tertiary and Quaternary of Europe, in the context of palaeoclimate and palaeogeography. Amphibia-Reptilia 24:133–167.
Rambaut, A., and A.J. Drummond. 2009. Tracer v1.5.0 and higher [http://beast.bio.ed.ac.uk]
Ronquist, F., and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.
Rögl, F. 1998. Paleogeographic considerations for Meditteranean and Paratethys seaways (Oligocene to Miocene). Annalen des Naturhistorischen Museums in Wien 99A:279–310.
Rögl, F. 1999. Mediterranean and Paratethys. Facts and hypotheses of an Oligocene to Miocene paleogeography (short overview). Geologica Carpathia 50:339–349.
Rohling, E.J., M. Fenton, F.J. Jorissen, P. Bertrand, G. Ganssen, and J.P. Caulet. 1998. Magnitudes of sea-level lowstands of the past 500,000 years. Nature 394:162–165.
Sanchíz, B. 1998. Salientia. Pp. 1–275 in W.P. Munich (ed.), Handbuch der paläoherpetologie Volume Part 4. Verlag Dr. Friedrich Pfeil.
Schätti, B., and A. Desvoignes. 1999. The herpetofauna of southern Yemen and the Socotran Archipelago. Instrumenta Biodiversitatis 4:1–178.
Sievers, F., A. Wilm, D.G. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, and D. Higgins. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology 7:Article 539.
Šmíd, J., S. Carranza, L. Kratochvíl, V. Gvoždík, A.K. Nasher, and J. Moravec. 2013. Out of Arabia: A complex biogeographic history of multiple vicariance and dispersal events in the Gecko genus *Hemidactylus* (Reptilia: Gekkonidae). PLoS ONE 8(5):e64018.
Stöck, M., C. Moritz, M. Hickerson, D. Frynta, T. Dujsebayeva, V. Eremchenko, R.J. Macey, T.J. Papenfuss, and D.B. Wake. 2006. Evolution of mitochondrial relationships and biogeography of Palearctic green toads (*Bufo viridis* subgroup) with insights in their genomic plasticity. Molecular Phylogenetics and Evolution 41:663–689.
Tihen, J.A. 1951. Anuran remains from the Miocene of Florida, with the description of a new species of *Bufo*. Copeia 3:230–235.
Uchupi, E., S.A. Swift, and D.A. Ross. 1999. Late Quaternary stratigraphy, Paleoclimate and neotectonism of the Persian (Arabian) Gulf region. Marine Geology 160:1–23.
Van Bocxlaer, I., S.D. Biju, S.P. Loader, and F. Bossuyt. 2009. Toad radiation reveals into-India dispersal as a source of endemism in the Western Ghats-Sri Lanka biodiversity hotspot. BMC Evolutionary Biology 9:131.
Van Bocxlaer, I., S.P. Loader, K. Roelants, S.D. Biju, M. Menegon, and F. Bossuyt. 2010. Gradual adaptation toward a range-expansion phenotype initiated the global radiation of toads. Science 327:679–682.

Vesey-Fitzgerald, D.F. 1955. Vegetation of the Red Sea coast south of Jedda, Saudi Arabia. Journal of Ecology 43:477–489.
Vesey-Fitzgerald, D.F. 1957. The vegetation of central and eastern Arabia. Journal of Ecology 45:779–798.
Wong, R.A., J.J. Fong, and T.J. Papenfuss. 2010. Phylogeography of the African helmeted terrapin, *Pelomedusa subrufa*: genetic structure, dispersal, and human introduction. Proceedings of the California Academy of Sciences 61:575–585.
Yu, X.-Q., M. Maki, B.T. Drew, A.J. Paton, H.-W. Li, J.-L. Zhao, J.G. Conran, and J. Li. 2014. Phylogeny and historical biogeography of *Isodon* (Lamiaceae): Rapid radiation in south-west China and Miocene overland dispersal into Africa. Molecular Phylogenetics and Evolution 77:183–194.
Zinner, D., L.F. Groeneveld, C. Keller, and C. Roos. 2009. Mitochondrial phylogeography of baboons (*Papio* spp.) – indication for introgressive hybridization? BMC Evolutionary Biology 9:1–15.
Zwickl, D.J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD thesis, The University of Texas at Austin, Austin.

Figure 24. The Arabian Peninsula. A map of the Arabian Peninsula showing disjunct extensions of the Horn of Africa Hotspot (medium grey) and the Eastern Afromontane Hotspot (dark grey), along with pertinent seas and country borders. Shape files for hotspots were obtained through the "Biodiversity Hotspots" package produced by Conservation International (2011).

Figure 24

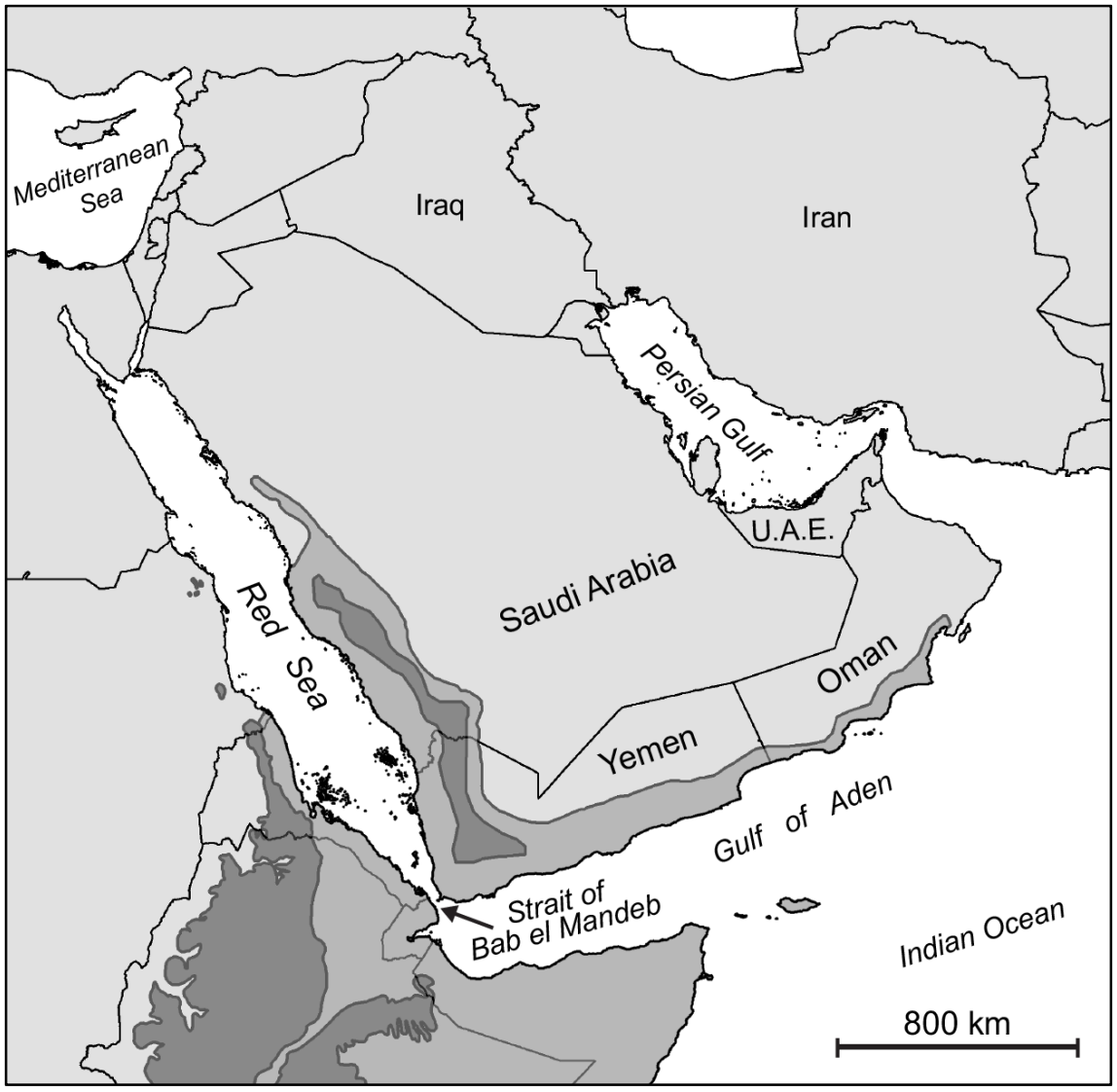

Figure 25. Phylogenetic relationships of global bufonids. The maximum likelihood tree of the Bufonidae, based on 13 loci and 243 taxa. Filled circles on nodes represent high support (Bayesian posterior probabilities [BPP] > 0.95 and maximum likelihood bootstrap scores [MLBS] > 70%); circles with top half fill: MLBS > 70% and BPP < 0.95; circles with bottom half fill: MLBS < 70% and BPP > 0.95; open circles represent support values less than given threshold for both analysis types. Branch lengths are proportional to substitutions/site, indicated by scale bar below. Genera are outlined in grey shading, genera containing Arabian taxa are colored and Arabian species are denoted by red bold text.

Figure 25

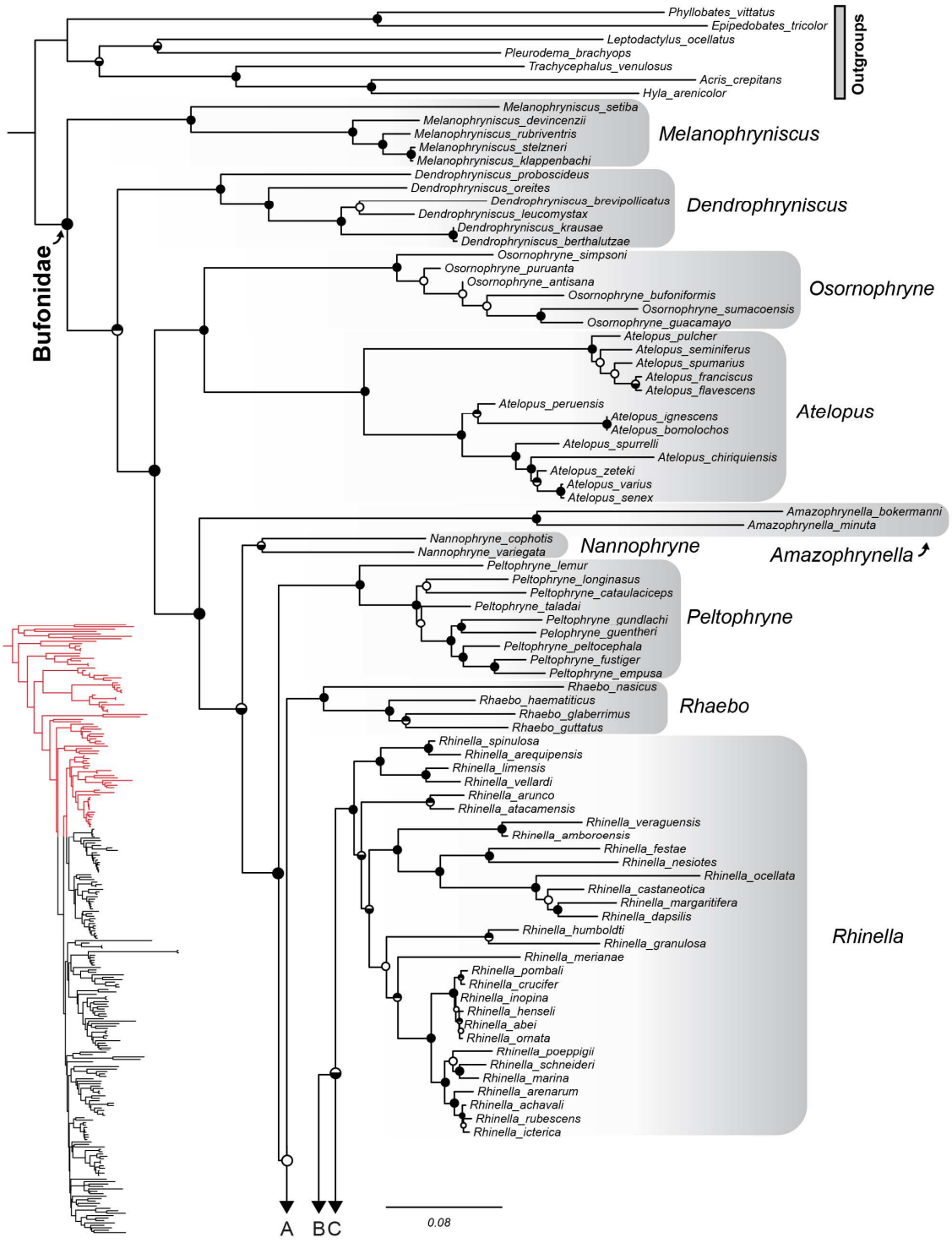
Phyllobates_vittatus
Epipedobates_tricolor
Leptodactylus_ocellatus
Pleurodema_brachyops
Trachycephalus_venulosus
Acris_crepitans
Hyla_arenicolor
Outgroups
Melanophryniscus_setiba
Melanophryniscus_devincenzii
Melanophryniscus_rubriventris
Melanophryniscus_stelzneri
Melanophryniscus_klappenbachi
Melanophryniscus
Dendrophryniscus_proboscideus
Dendrophryniscus_oreites
Dendrophryniscus_brevipollicatus
Dendrophryniscus_leucomystax
Dendrophryniscus_krausae
Dendrophryniscus_berthalutzae
Dendrophryniscus
Bufonidae
Osornophryne_simpsoni
Osornophryne_puruanta
Osornophryne_antisana
Osornophryne_bufoniformis
Osornophryne_sumacoensis
Osornophryne_guacamayo
Osornophryne
Atelopus_pulcher
Atelopus_seminiferus
Atelopus_spumarius
Atelopus_franciscus
Atelopus_flavescens
Atelopus_peruensis
Atelopus_ignescens
Atelopus_bomolochos
Atelopus_spurrelli
Atelopus_chiriquiensis
Atelopus_zeteki
Atelopus_varius
Atelopus_senex
Atelopus
Amazophrynella_bokermanni
Amazophrynella_minuta
Amazophrynella
Nannophryne_cophotis
Nannophryne_variegata
Nannophryne
Peltophryne_lemur
Peltophryne_longinasus
Peltophryne_cataulaciceps
Peltophryne_taladai
Peltophryne_gundlachi
Pelophryne_guentheri
Peltophryne_peltocephala
Peltophryne_fustiger
Peltophryne_empusa
Peltophryne
Rhaebo_nasicus
Rhaebo_haematiticus
Rhaebo_glaberrimus
Rhaebo_guttatus
Rhaebo
Rhinella_spinulosa
Rhinella_arequipensis
Rhinella_limensis
Rhinella_vellardi
Rhinella_arunco
Rhinella_atacamensis
Rhinella_veraguensis
Rhinella_amboroensis
Rhinella_festae
Rhinella_nesiotes
Rhinella_ocellata
Rhinella_castaneotica
Rhinella_margaritifera
Rhinella_dapsilis
Rhinella_humboldti
Rhinella_granulosa
Rhinella_merianae
Rhinella_pombali
Rhinella_crucifer
Rhinella_inopina
Rhinella_henseli
Rhinella_abei
Rhinella_ornata
Rhinella_poeppigii
Rhinella_schneideri
Rhinella_marina
Rhinella_arenarum
Rhinella_achavali
Rhinella_rubescens
Rhinella_icterica
Rhinella
A
BC
0.08

Figure 25 (continued).

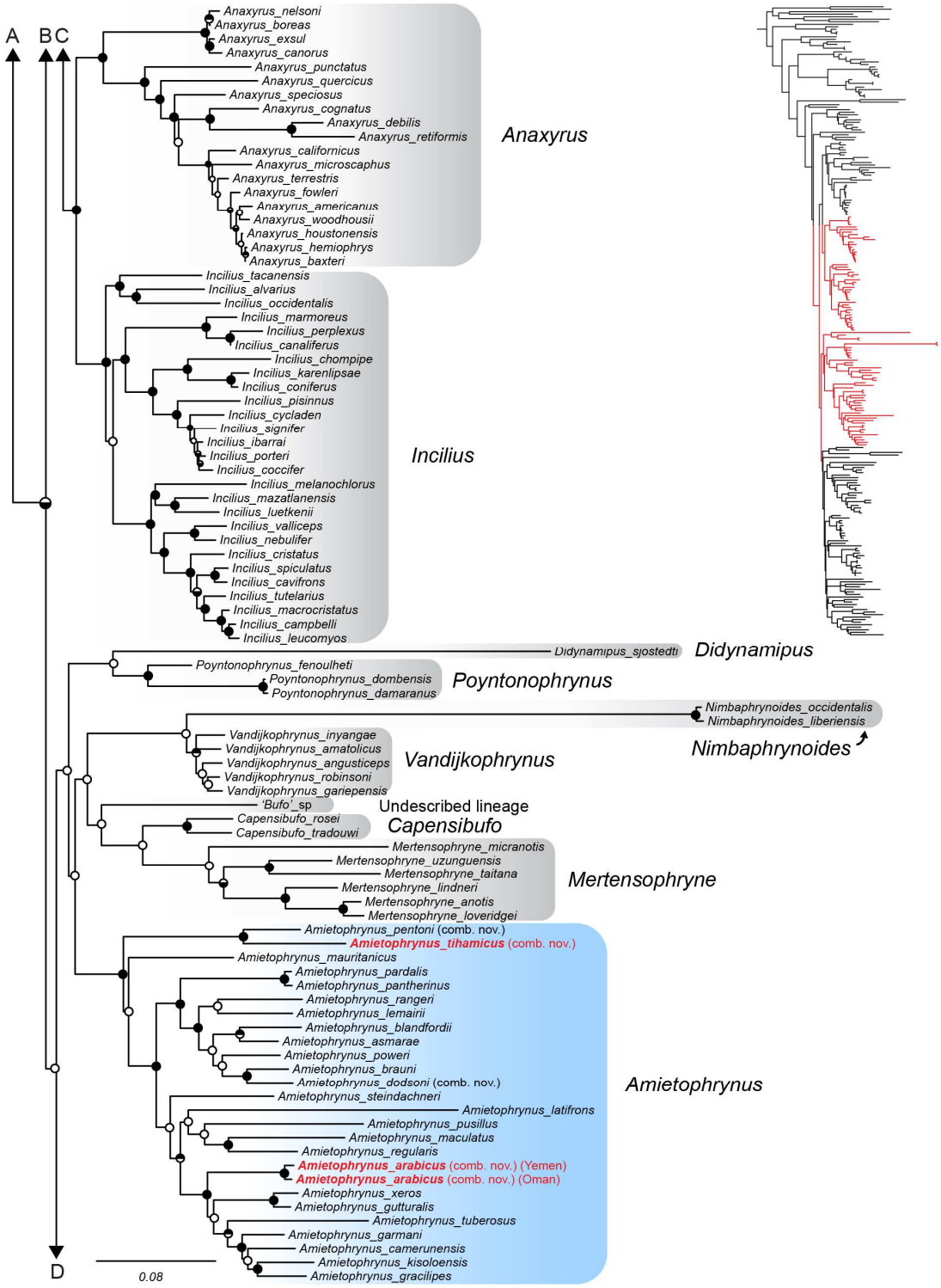

Figure 25 (continued).

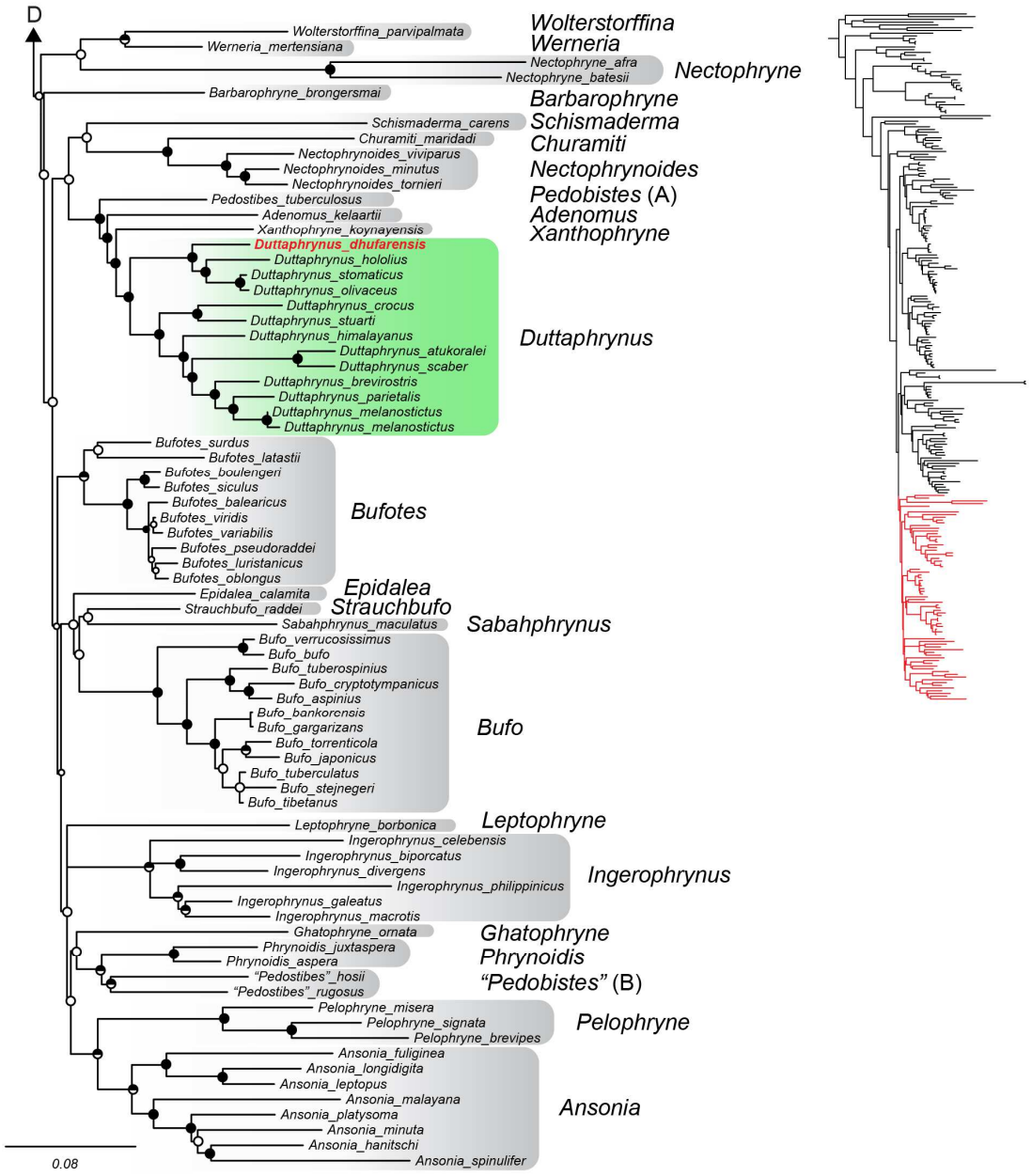

Figure 26. Divergence timings of the genus *Amietophrynus*. Bayesian maximum clade credibility chronogram of the genus *Amietophrynus* inferred in BEAST (Analysis 2) with median divergence times and associated 95% highest posterior distributions of dates in brackets. Timings of relevant geologic events are illustrated: A – Red Sea formation (27–23 Ma); B – *Gomphotherium* bridge (18–16 Ma); C – permanent Eurasian landbridge (15 Ma); D – southern Afro-Arabian landbridge (10–5.3 Ma). Ranges of Arabian species and, if relevant, corresponding sister taxon are illustrated. Filled circles on nodes represent Bayesian posterior probabilities > 0.95, open circles represent support values < 0.95.

Figure 26

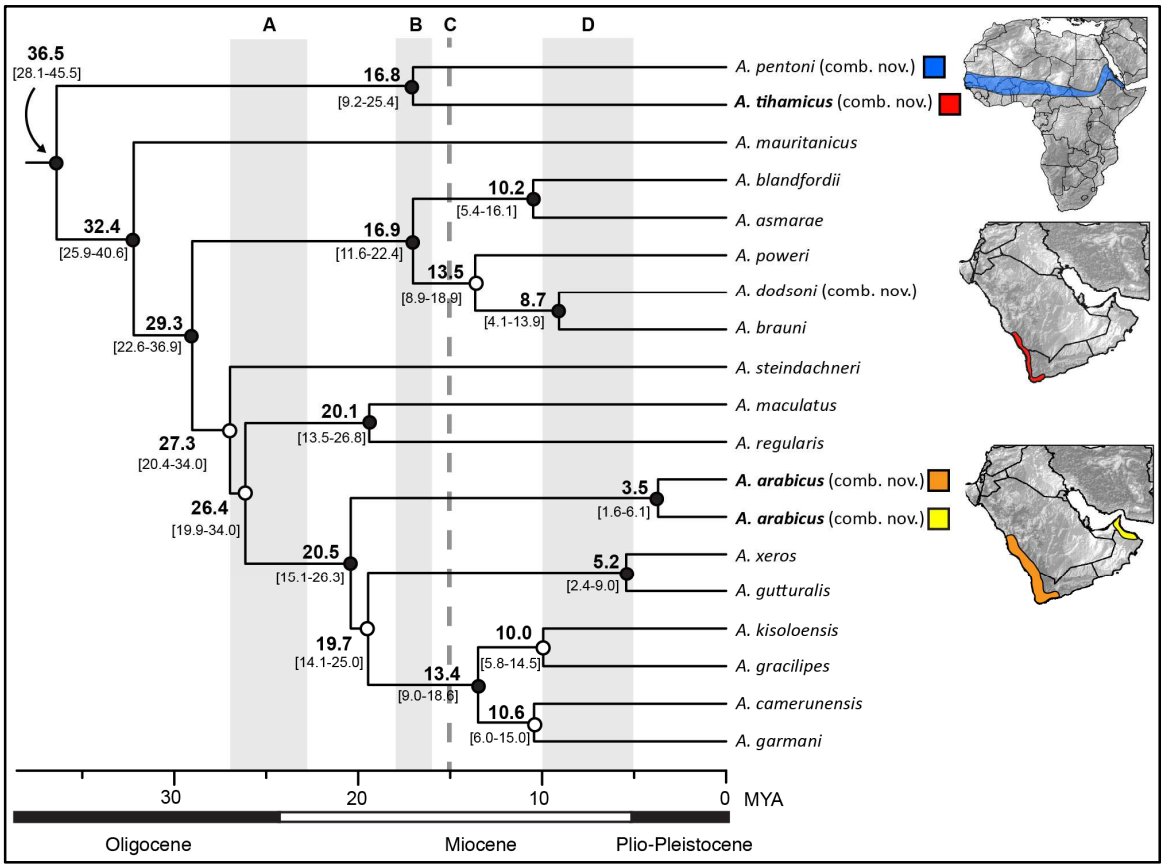

Figure 27. Divergence timings of the genus *Duttaphrynus*. Bayesian maximum clade credibility chronogram of the genus *Duttaphrynus* inferred in BEAST (Analysis 2) with median divergence times and associated 95% highest posterior distributions of dates in brackets. Timings of relevant geologic events are illustrated: A – Red Sea formation (27–23 Ma); B - *Gomphotherium* bridge (18–16 Ma); C – permanent Eurasian landbridge (15 Ma); D – southern Afro-Arabian landbridge (10–5.3 Ma). Ranges of Arabian species and, if relevant, corresponding sister taxon are illustrated. Filled circles on nodes represent Bayesian posterior probabilities > 0.95, open circles represent support values < 0.95.

Figure 27

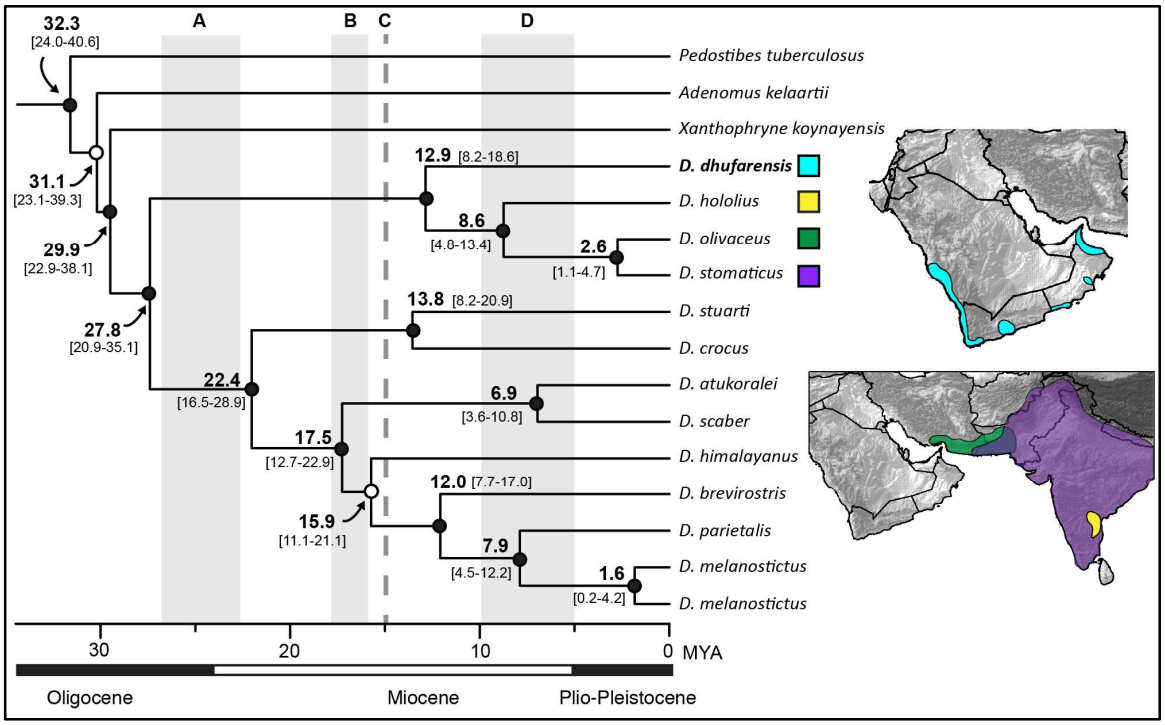

Table 18. Key results of divergence dating analyses.

Table 18

	Dating Analysis		
Node Label	A1	A2	A3
Root Age	124.8 [84.1–177.1]	121.5 [89.3–160.3]	128.0 [85.7–182.1]
Bufonidae	94.9 [64.6–129.8]	93.5 [73.0–114.0]	100.3 [70.1–133.1]
Origin *Amietophrynus*	37.6 [25.6–51.5]	36.5 [28.1–45.5]	37.2 [27.1–50.4]
TMRCA *A. arabicus*	21.2 [14.2–29.2]	20.5 [15.1–26.3]	20.7 [14.8–29.2]
TMRCA *A. tihamicus*	17.9 [9.0–26.6]	16.8 [9.3–25.4]	17.0 [9.3–25.9]
Origin *Duttaphrynus*	29.3 [19.3–40.5]	27.8 [20.9–35.1]	29.2 [20.4–40.0]
TMRCA *D. dhufarensis*	13.5 [7.9–20.9]	12.9 [8.2–18.6]	13.5 [8.4–20.1]

For relevant nodes, median age and 95% highest posterior density region are given (Ma). Analysis 1 (A1): inclusion of four internal calibrations with lognormal calibration prior for age of the Bufonidae; Analysis 2 (A2): inclusion of four internal calibrations with normal calibration prior for age of the Bufonidae; Analysis 3 (A3): inclusion of four internal calibrations with exponential calibration prior for age of the Bufonidae.

Table 19. New sampling of taxa included for this study.

Table 19

Species	Museum No.	Latitude	Longitude	Country
*Amietophrynus arabicus**	CAS 250888	25.2651	56.3068	Oman
	CAS 250889	25.2651	56.3068	Oman
	CAS 250890	25.2651	56.3068	Oman
	CAS 250907	25.387	56.2649	Oman
	CAS 250908	25.387	56.2649	Oman
	CAS 251024	22.6052	59.0886	Oman
	CAS 251026	22.6052	59.0886	Oman
	CAS 251027	22.6052	59.0886	Oman
	CAS 251126	20.6888	58.2949	Oman
	CAS 251127	20.6888	58.2949	Oman
	CAS 251130	20.6888	58.2949	Oman
	CAS 251147	23.0548	57.4679	Oman
	CAS 251148	23.0548	57.4679	Oman
	CAS 251149	23.0548	57.4679	Oman
	CAS 251166	23.0713	57.6042	Oman
	CAS 251167	23.0713	57.6042	Oman
	MVZ 236403	15.3443	44.217	Yemen
	MVZ 236407	15.4693	44.2618	Yemen
	MVZ 236866	24.2633	56.1633	Oman
	MVZ 241304	23.0525	57.4691	Oman
	MVZ 241305	23.0525	57.4691	Oman
	MVZ 241306	23.0713	57.6043	Oman
	MVZ 241307	23.0713	57.6043	Oman
Amietophrynus asmarae	MVZ 257839			Ethiopia
	MVZ 257840			Ethiopia
	MVZ 257843			Ethiopia
	MVZ 257844			Ethiopia
	MVZ 257847			Ethiopia
	MVZ 257848			Ethiopia
	MVZ 257849			Ethiopia
	MVZ 257850			Ethiopia
	MVZ 257852			Ethiopia
Amietophrynus blandfordii	MVZ 241309	9.9493	43.2193	Somalia
	MVZ 241313	9.9698	43.4325	Somalia
	MVZ 241314	9.9698	43.4325	Somalia
	MVZ 241316	9.9698	43.4325	Somalia
	MVZ 241317	9.9698	43.4325	Somalia
	MVZ 241318	9.9698	43.4325	Somalia

	MVZ 242725	11.0261	49.193	Somalia
*Amietophrynus dodsoni**	MVZ 241310	9.9493	43.2193	Somalia
	MVZ 241312	9.9493	43.2193	Somalia
Amietophrynus garmani	**MVZ 257841**			Ethiopia
	MVZ 257846			Ethiopia
Amietophrynus mauritanicus	MVZ 235679	36.9641	8.8822	Tunisia
*Amietophrynus pentoni**	MVZ 235732	16.5796	-15.8735	Mauritania
	MVZ 249297	9.24191	-1.84415	Ghana
	MVZ 249298	9.24191	-1.84415	Ghana
Amietophrynus regularis	MVZ 235735	16.5165	-15.8135	Mauritania
	MVZ 238858	13.5036	2.1135	Niger
	MVZ 238859	13.5036	2.1135	Niger
	MVZ 249302	9.259	-1.8541	Ghana
Amietophrynus steindachneri	MVZ 234101	-3.1999	40.0077	Kenya
	MVZ 234102	-3.1999	40.0077	Kenya
*Amietophrynus tihamicus**	**MVZ 236409**	14.8238	43.1273	Yemen
	MVZ 236413	14.8238	43.1273	Yemen
Amietophrynus xeros	MVZ 235737	16.5165	-15.8135	Mauritania
	MVZ 238867	17.3868	7.9563	Niger
	MVZ 238868	17.3868	7.9563	Niger
"Bufo" sp.	**MVZ 242731**	11.2493	49.2678	Somalia
	MVZ 242732	11.2493	49.2678	Somalia
	MVZ 242733	11.2493	49.2678	Somalia
	MVZ 242776	11.0261	49.193	Somalia
Bufotes oblongus	MVZ 241548	36.2797	60.548	Iran
	MVZ 245904	32.8198	59.2171	Iran
	MVZ 245905	32.8198	59.2171	Iran
	MVZ 245906	32.8198	59.2171	Iran
	MVZ 245910	33.631	57.1616	Iran
	MVZ 245911	33.631	57.1616	Iran
	MVZ 245912	33.631	57.1616	Iran
	MVZ 245917	35.9666	56.0684	Iran
	MVZ 248374	36.2688	60.5357	Iran
	MVZ 248376	37.7194	55.9001	Iran
	MVZ 249177	37.7332	55.9006	Iran
	MVZ 249178	37.7332	55.9006	Iran
Bufotes pseudoraddei	MVZ 241550	36.5	74.8666	Pakistan
	MVZ 241551	36.5	74.8666	Pakistan
	MVZ 241552	36.5	74.8666	Pakistan
	MVZ 241553	35.8833	71.7833	Pakistan
	MVZ 241554	36.0663	72.5166	Pakistan
	MVZ 248375	36.5	74.87	Pakistan
Bufotes surdus	MVZ 234217	29.4441	60.5136	Iran

	MVZ 234218	29.4441	60.5136	Iran
	MVZ 234219	29.4441	60.5136	Iran
	MVZ 234238	27.8763	60.0955	Iran
	MVZ 234239	28.6066	61.0771	Iran
	MVZ 234240	28.6066	61.0771	Iran
Bufotes variabilis	**MVZ 234222**	32.9395	48.2558	Iran
	MVZ 234223	32.9395	48.2558	Iran
	MVZ 234241	31.2133	49.2173	Iran
	MVZ 234242	31.2133	49.2173	Iran
	MVZ 238503	29.615	52.5386	Iran
Duttaphrynus dhufarensis	MVZ 241308	26.1503	56.1606	Oman
	MVZ 242729	23.0655	57.4701	Oman
	MVZ 242774	17.1001	54.284	Oman
	MVZ 242775	23.0655	57.4701	Oman
Duttaphrynus himalayanus	**MVZ 241543**	28.598	83.6469	Himalayas
	MVZ 241544	28.598	83.6469	Himalayas
Duttaphrynus melanostictus	MVZ 226298	21.4536	105.6436	Vietnam
	MVZ 239140	-3.957	122.5315	Indonesia
Duttaphrynus olivaceus	**MVZ 234225**	27.2035	60.6785	Iran
	MVZ 234226	27.1853	60.5895	Iran
	MVZ 234227	27.1853	60.5895	Iran
	MVZ 234228	27.1853	60.5895	Iran
	MVZ 234235	25.2563	60.8326	Iran
	MVZ 234236	25.2563	60.8326	Iran
	MVZ 234237	25.2703	60.7553	Iran
Duttaphrynus stomaticus	**MVZ 237424**	34.4366	70.4483	Afghanistan
	MVZ 248377	25.7768	66.6256	Pakistan
	MVZ 248378	25.7768	66.6256	Pakistan
	MVZ 248379	25.7768	66.6256	Pakistan
	MVZ 248380	24.3516	70.7573	Pakistan
	MVZ 248382	24.3516	70.7573	Pakistan
	MVZ 248383	24.3516	70.7573	Pakistan
Duttaphrynus stuarti	CAS 242587	27.7746	98.3354	China

Museum numbers in bold represent samples included in the dating analyses. Asterisks indicate new name combinations for species.

Appendix 6**.** Voucher, locality information, and gene sampling of new specimens.

Species	**Museum No.**	**16S**	**ND2**	**RAG1**	**Dating Analysis**	**Lat**	**Long**	**Country**
Amietophrynus arabicus	CAS 250888	x	x	x		25.2651	56.3068	Oman
Amietophrynus arabicus	CAS 250889	x	x	x	YES	25.2651	56.3068	Oman
Amietophrynus arabicus	CAS 250890	x	x	x		25.2651	56.3068	Oman
Amietophrynus arabicus	CAS 250907	x	x	x		25.387	56.2649	Oman
Amietophrynus arabicus	CAS 250908	x	x	x		25.387	56.2649	Oman
Amietophrynus arabicus	CAS 251024	x	x	x		22.6052	59.0886	Oman
Amietophrynus arabicus	CAS 251026	x	x	x		22.6052	59.0886	Oman
Amietophrynus arabicus	CAS 251027	x	x	x		22.6052	59.0886	Oman
Amietophrynus arabicus	CAS 251126	x	x	x		20.6888	58.2949	Oman
Amietophrynus arabicus	CAS 251127	x	x	x		20.6888	58.2949	Oman
Amietophrynus arabicus	CAS 251130	x	x	x		20.6888	58.2949	Oman
Amietophrynus arabicus	CAS 251147	x	x	x		23.0548	57.4679	Oman
Amietophrynus arabicus	CAS 251148	x	x	x		23.0548	57.4679	Oman
Amietophrynus arabicus	CAS 251149	x	x	x		23.0548	57.4679	Oman
Amietophrynus arabicus	CAS 251166	x	x	x		23.0713	57.6042	Oman
Amietophrynus arabicus	CAS 251167	x	x	x		23.0713	57.6042	Oman
Amietophrynus arabicus	MVZ 236403	x	x	x		15.3443	44.217	Yemen
Amietophrynus arabicus	MVZ 236407	x	x	x	YES	15.4693	44.2618	Yemen
Amietophrynus arabicus	MVZ 236866	x	x	x		24.2633	56.1633	Oman
Amietophrynus arabicus	MVZ 241304	x	x	x		23.0525	57.4691	Oman
Amietophrynus arabicus	MVZ 241305	x	x	x		23.0525	57.4691	Oman
Amietophrynus arabicus	MVZ 241306	x	x	x		23.0713	57.6043	Oman
Amietophrynus arabicus	MVZ 241307	x	x	x		23.0713	57.6043	Oman
Amietophrynus asmarae	MVZ 257839	x		x				Ethiopia
Amietophrynus asmarae	MVZ 257840	x		x				Ethiopia

Amietophrynus asmarae	MVZ 257843	x		x				Ethiopia
Amietophrynus asmarae	MVZ 257844	x		x				Ethiopia
Amietophrynus asmarae	MVZ 257847	x						Ethiopia
Amietophrynus asmarae	MVZ 257848	x	x	x	YES			Ethiopia
Amietophrynus asmarae	MVZ 257849	x	x	x				Ethiopia
Amietophrynus asmarae	MVZ 257850	x						Ethiopia
Amietophrynus asmarae	MVZ 257852	x						Ethiopia
Amietophrynus blandfordii	MVZ 241309	x		x		9.9493	43.2193	Somalia
Amietophrynus blandfordii	MVZ 241313	x		x		9.9698	43.4325	Somalia
Amietophrynus blandfordii	MVZ 241314	x		x		9.9698	43.4325	Somalia
Amietophrynus blandfordii	MVZ 241316	x		x		9.9698	43.4325	Somalia
Amietophrynus blandfordii	MVZ 241317	x		x		9.9698	43.4325	Somalia
Amietophrynus blandfordii	MVZ 241318	x		x		9.9698	43.4325	Somalia
Amietophrynus blandfordii	MVZ 242725	x	x	x	YES	11.0261	49.193	Somalia
Amietophrynus dodsoni	MVZ 241310	x	x	x		9.9493	43.2193	Somalia
Amietophrynus dodsoni	MVZ 241312	x	x	x	YES	9.9493	43.2193	Somalia
Amietophrynus garmani	MVZ 257841	x		x	YES			Ethiopia
Amietophrynus garmani	MVZ 257846	x		x				Ethiopia
Amietophrynus mauritanicus	MVZ 235679	x	x	x		36.9641	8.8822	Tunisia
Amietophrynus pentoni	MVZ 235732	x	x	x		16.5796	-15.8735	Mauritania
Amietophrynus pentoni	MVZ 249297	x	x	x	YES	9.24191	-1.84415	Ghana
Amietophrynus pentoni	MVZ 249298	x	x	x		9.24191	-1.84415	Ghana
Amietophrynus regularis	MVZ 235735	x	x	x		16.5165	-15.8135	Mauritania
Amietophrynus regularis	MVZ 238858	x	x	x		13.5036	2.1135	Niger
Amietophrynus regularis	MVZ 238859	x	x	x		13.5036	2.1135	Niger
Amietophrynus regularis	MVZ 249302	x				9.259	-1.8541	Ghana
Amietophrynus steindachneri	MVZ 234101	x				-3.1999	40.0077	Kenya

Amietophrynus steindachneri	MVZ 234102	x				-3.1999	40.0077	Kenya
Amietophrynus tihamicus	MVZ 236409	x	x	x	YES	14.8238	43.1273	Yemen
Amietophrynus tihamicus	MVZ 236413	x	x	x		14.8238	43.1273	Yemen
Amietophrynus xeros	MVZ 235737	x				16.5165	-15.8135	Mauritania
Amietophrynus xeros	MVZ 238867	x	x	x		17.3868	7.9563	Niger
Amietophrynus xeros	MVZ 238868	x	x	x		17.3868	7.9563	Niger
Bufo sp	MVZ 242731	x	x	x	YES	11.2493	49.2678	Somalia
Bufo sp	MVZ 242732	x		x		11.2493	49.2678	Somalia
Bufo sp	MVZ 242733	x	x	x		11.2493	49.2678	Somalia
Bufo sp	MVZ 242776	x		x		11.0261	49.193	Somalia
Bufotes oblongus	MVZ 241548	x	x	x		36.2797	60.548	Iran
Bufotes oblongus	MVZ 245904	x		x		32.8198	59.2171	Iran
Bufotes oblongus	MVZ 245905	x	x	x		32.8198	59.2171	Iran
Bufotes oblongus	MVZ 245906	x	x	x		32.8198	59.2171	Iran
Bufotes oblongus	MVZ 245910	x		x		33.631	57.1616	Iran
Bufotes oblongus	MVZ 245911	x	x	x		33.631	57.1616	Iran
Bufotes oblongus	MVZ 245912	x	x	x		33.631	57.1616	Iran
Bufotes oblongus	MVZ 245917	x	x	x		35.9666	56.0684	Iran
Bufotes oblongus	MVZ 248374	x		x		36.2688	60.5357	Iran
Bufotes oblongus	MVZ 248376	x	x	x		37.7194	55.9001	Iran
Bufotes oblongus	MVZ 249177	x	x	x		37.7332	55.9006	Iran
Bufotes oblongus	MVZ 249178	x	x	x		37.7332	55.9006	Iran
Bufotes pseudoraddei	MVZ 241550	x	x	x		36.5	74.8666	Pakistan
Bufotes pseudoraddei	MVZ 241551	x		x		36.5	74.8666	Pakistan
Bufotes pseudoraddei	MVZ 241552	x	x	x		36.5	74.8666	Pakistan
Bufotes pseudoraddei	MVZ 241553	x	x	x	YES	35.8833	71.7833	Pakistan
Bufotes pseudoraddei	MVZ 241554	x	x	x		36.0663	72.5166	Pakistan

Bufotes pseudoraddei	MVZ 248375	x	x	x		36.5	74.87	Pakistan
Bufotes surdus	MVZ 234217	x	x	x		29.4441	60.5136	Iran
Bufotes surdus	MVZ 234218	x	x	x		29.4441	60.5136	Iran
Bufotes surdus	MVZ 234219	x	x	x		29.4441	60.5136	Iran
Bufotes surdus	MVZ 234238	x	x	x	YES	27.8763	60.0955	Iran
Bufotes surdus	MVZ 234239	x	x	x		28.6066	61.0771	Iran
Bufotes surdus	MVZ 234240	x	x	x		28.6066	61.0771	Iran
Bufotes variabilis	MVZ 234222	x	x	x	YES	32.9395	48.2558	Iran
Bufotes variabilis	MVZ 234223	x		x		32.9395	48.2558	Iran
Bufotes variabilis	MVZ 234241	x	x	x		31.2133	49.2173	Iran
Bufotes variabilis	MVZ 234242	x	x	x		31.2133	49.2173	Iran
Bufotes variabilis	MVZ 238503	x	x	x		29.615	52.5386	Iran
Duttaphrynus dhufarensis	MVZ 241308	x		x		26.1503	56.1606	Oman
Duttaphrynus dhufarensis	MVZ 242729	x	x	x	YES	23.0655	57.4701	Oman
Duttaphrynus dhufarensis	MVZ 242774	x	x	x		17.1001	54.284	Oman
Duttaphrynus dhufarensis	MVZ 242775	x	x	x		23.0655	57.4701	Oman
Duttaphrynus himalayanus	MVZ 241543	x	x	x	YES	28.598	83.6469	Himalayas
Duttaphrynus himalayanus	MVZ 241544	x	x	x		28.598	83.6469	Himalayas
Duttaphrynus melanostictus	MVZ 226298	x	x	x		21.4536	105.6436	Vietnam
Duttaphrynus melanostictus	MVZ 239140	x	x	x	YES	-3.957	122.5315	Indonesia
Duttaphrynus olivaceus	MVZ 234225	x	x	x	YES	27.2035	60.6785	Iran
Duttaphrynus olivaceus	MVZ 234226	x	x	x		27.1853	60.5895	Iran
Duttaphrynus olivaceus	MVZ 234227	x	x	x		27.1853	60.5895	Iran
Duttaphrynus olivaceus	MVZ 234228	x		x		27.1853	60.5895	Iran
Duttaphrynus olivaceus	MVZ 234235	x	x	x		25.2563	60.8326	Iran
Duttaphrynus olivaceus	MVZ 234236	x	x	x		25.2563	60.8326	Iran
Duttaphrynus olivaceus	MVZ 234237	x	x	x		25.2703	60.7553	Iran

Duttaphrynus stomaticus	MVZ 237424	x	x	x	YES	34.4366	70.4483	Afghanistan
Duttaphrynus stomaticus	MVZ 248377	x	x	x		25.7768	66.6256	Pakistan
Duttaphrynus stomaticus	MVZ 248378	x		x		25.7768	66.6256	Pakistan
Duttaphrynus stomaticus	MVZ 248379	x	x	x		25.7768	66.6256	Pakistan
Duttaphrynus stomaticus	MVZ 248380	x		x		24.3516	70.7573	Pakistan
Duttaphrynus stomaticus	MVZ 248382	x		x		24.3516	70.7573	Pakistan
Duttaphrynus stomaticus	MVZ 248383	x	x	x		24.3516	70.7573	Pakistan
Duttaphrynus stuarti	CAS 242587	x	x	x	YES	27.7746	98.3354	China

Appendix 7. Information for Genbank data included.

Species	12S	16S	cytb	ND2	CXCR4	H3A	NCX1	POMC	RAG1	RHOD	SIA	SLC8A3	TYR
Outgroup Sampling													
Acris crepitans	FJ882758	EF107181	EF988131	---	EF107468	DQ284107	EF107244	EF988231	EF107304	---	AY844762	EF107403	EF988305
Epipedobates tricolor	AY364576	EU342578	AF128550	---	EF107458	---	EF107233	---	EF107295	---	---	EF107381	---
Hyla arenicolor	DQ347054	FJ882776	AY831027	---	AY364190	---	EF107241	AY819112	AY844391	AY364401	AY844802	EF107393	DQ347187
Leptodactylus ocellatus	AY843688	AY843688	AY843934	---	DQ306492	DQ284104	---	DQ158259	DQ158343	---	AY844890	---	---
Phyllobates vittatus	FJ882833	EU342635	AF324039	---	EF107459	---	EF107234	---	DQ503355	---	DQ503105	EF107382	DQ503166
Pleurodema brachyops	AY843733	AY843733	AY843979	---	---	DQ284111	---	---	AY844503	---	AY844926	---	---
Trachycephalus venulosus	DQ347027	FJ882779	EU034077	---	GQ366030	---	AY948824	AY819132	EU034147	AY364396	AY844912	AY948880	DQ347161
Ingroup Sampling													
Adenomus kelaartii	FJ882780	FJ882780	---	---	EF107447	---	EF107221	---	EF107284	---	---	EF107348	---
Amazophrynella bokermanni	JN867532	JN867559	---	---	---	---	---	---	---	---	---	---	---
Amazophrynella minuta	AY843582	AY843582	AY843804	---	DQ306496	DQ284096	---	AY819081	DQ503337	---	---	---	EF364362
Amietophrynus brauni	FJ882822	FJ882822	---	---	DQ306514	---	EF107272	DQ158279	EF107331	DQ284021	DQ282873	EF107435	DQ283011
Amietophrynus camerunensis	DQ158439	DQ158439	---	AF463789	DQ306555	DQ284345	---	DQ158281	DQ158363	DQ283979	DQ282830	---	---
Amietophrynus garmani	FJ882823	FJ882823	---	---	DQ306547	---	FJ882668	DQ158294	DQ158375	---	---	---	---
Amietophrynus gracilipes	FJ882824	DQ158456	---	---	DQ306522	---	FJ882669	DQ158297	DQ158378	---	---	---	---
Amietophrynus gutturalis	DQ158460	DQ158460	---	---	FJ882725	---	FJ882670	DQ158301	DQ158382	DQ284035	DQ282890	---	---
Amietophrynus kisoloensis	AY325995	AY325995	L10975	AF463788	DQ306560	---	---	DQ158305	---	---	---	---	---
Amietophrynus latifrons	DQ283343	DQ283343	---	---	---	DQ284332	---	---	---	DQ283970	DQ282819	---	---
Amietophrynus lemairii	AF220847	AF220895	---	---	---	---	---	---	---	---	---	---	---
Amietophrynus maculatus	DQ158469	DQ158469	L10979	AF463782	DQ306533	DQ284374	---	DQ158310	DQ158389	DQ284005	DQ282850	---	---
Amietophrynus mauritanicus	FJ882826	AY680265	---	---	FJ882727	---	FJ882672	---	---	---	---	---	---
Amietophrynus pantherinus	AF220848	AF220896	AF210087	GQ994942	---	---	---	---	---	---	---	---	---
Amietophrynus pardalis	AF220850	AF220897	AF210083	---	---	---	---	---	---	---	---	---	---

Amietophrynus poweri	FJ882771	AF220882	---	AF463781	FJ882722	---	---	DQ158324	DQ158401	---	---	---	---
Amietophrynus pusillus	AF220837	AF220883	---	---	---	---	---	---	---	---	---	---	---
Amietophrynus rangeri	AF220828	AF220868	---	AF463773	---	---	---	---	---	---	---	---	---
Amietophrynus regularis	DQ158485	DQ158485	AY028511	---	DQ306523	DQ284201	---	DQ158326	FJ227103	DQ283858	DQ282716	---	---
Amietophrynus steindachneri	FJ882825	AY325981	---	---	DQ306546	---	FJ882671	DQ158329	DQ158406	---	---	---	---
Amietophrynus tuberosus	DQ283362	DQ283362	---	---	---	---	---	---	---	DQ283984	DQ282832	---	---
Amietophrynus xeros	DQ158499	DQ158499	---	---	DQ306561	---	---	DQ158340	DQ158414	---	---	---	---
Anaxyrus americanus	FJ882827	DQ158426	AB159264	---	DQ306520	---	FJ882676	DQ158268	DQ158352	---	---	---	---
Anaxyrus baxteri	AY680207	AY680207	---	---	---	---	---	---	---	---	---	---	---
Anaxyrus boreas	DQ283180	---	L10959	JN868005	DQ306499	DQ284215	FJ882678	DQ158278	DQ158360	DQ283871	---	---	---
Anaxyrus californicus	FJ882828	FJ882828	---	---	---	---	---	---	---	---	---	---	---
Anaxyrus canorus	EF532005	AY680239	L10967	---	---	---	---	---	---	---	---	---	---
Anaxyrus cognatus	DQ158444	DQ158444	L10968	---	DQ306498	DQ284197	---	DQ158285	DQ158367	---	DQ282713	---	---
Anaxyrus debilis	DQ158449	AY680233	AY010165	---	DQ306507	---	---	DQ158290	DQ158371	---	---	---	---
Anaxyrus exsul	AY325990	AY325990	L10971	AF004524	DQ306550	---	---	DQ158291	DQ158372	---	---	---	---
Anaxyrus fowleri	DQ158451	AY680224	---	---	DQ306505	---	---	DQ158292	DQ158373	---	---	---	---
Anaxyrus hemiophrys	EF532234	AY680213	---	---	---	---	---	GQ443070	GQ443081	---	---	---	---
Anaxyrus houstonensis	EF532236	AY680208	---	---	---	---	---	---	---	---	---	---	---
Anaxyrus microscaphus	AY325989	AY325989	L10978	---	DQ306563	---	---	DQ158318	DQ158395	---	---	---	---
Anaxyrus nelsoni	EF532010	AY680240	---	---	---	---	---	---	---	---	---	---	---
Anaxyrus punctatus	EF532246	DQ283160	DQ085753	---	---	DQ284198	---	---	---	DQ283855	DQ282714	---	---
Anaxyrus quercicus	DQ158484	DQ158484	---	---	DQ306549	DQ284192	---	DQ158325	DQ158403	---	DQ282708	---	---
Anaxyrus retiformis	AY325982	AY325982	AY010166	---	---	---	---	---	---	---	---	---	---
Anaxyrus speciosus	EF532242	AY680229	L10980	---	---	---	---	---	---	---	---	---	---
Anaxyrus terrestris	FJ882829	FJ882829	---	---	DQ306537	DQ284196	FJ882677	DQ158330	---	DQ283854	DQ282712	---	---
Anaxyrus woodhousii	AY325986	AY325986	AY288067	---	DQ306551	DQ284222	---	DQ158339	DQ158413	DQ283875	DQ282733	---	---
Ansonia fuliginea	AB331709	AB331709	---	---	---	---	---	---	---	---	---	---	---

Ansonia hanitschi	AB331710	FJ882794	---	---	FJ882695	---	FJ882642	---	---	---	---	---	---
Ansonia leptopus	FJ882795	FJ882795	---	---	FJ882697	---	FJ882644	---	---	---	---	---	---
Ansonia longidigita	FJ882796	FJ882796	---	---	FJ882698	---	FJ882645	---	---	DQ283968	DQ282817	---	---
Ansonia malayana	AB331712	AB331712	---	---	---	---	---	---	---	---	---	---	---
Ansonia minuta	AB435283	AB435283	---	---	---	---	---	---	---	---	---	---	---
Ansonia platysoma	AB435272	AB435272	---	---	---	---	---	---	---	---	---	---	---
Ansonia spinulifer	FJ882798	FJ882798	---	---	FJ882696	---	FJ882643	---	---	---	---	---	---
Atelopus bomolochos	AF375482	AF375508	AY995955	---	---	---	---	---	---	---	---	---	---
Atelopus chiriquiensis	U52735	U52780	---	---	---	---	---	---	---	---	---	---	---
Atelopus flavescens	DQ283259	DQ283259	AY995960	---	---	DQ284282	---	---	---	DQ283928	DQ282780	---	EF364359
Atelopus franciscus	---	---	AY995991	---	---	---	---	---	---	---	---	---	---
Atelopus ignescens	---	---	AY995950	---	---	---	---	---	---	---	---	---	---
Atelopus peruensis	DQ158419	DQ158419	---	---	DQ306495	---	---	AY819078	---	---	---	---	---
Atelopus pulcher	---	EU672973	AY995953	---	---	---	---	---	---	---	---	---	---
Atelopus seminiferus	---	EU672976	---	---	---	---	---	---	---	---	---	---	---
Atelopus senex	---	---	EF494950	---	---	---	---	---	---	---	---	---	---
Atelopus spumarius	DQ283260	DQ283260	AY995954	---	GU183852	DQ284283	GU183854	---	---	DQ283929	DQ282781	---	---
Atelopus spurrelli	DQ502200	DQ502200	---	---	---	---	---	---	DQ503380	---	---	---	---
Atelopus varius	AY325996	AY325996	EF494953	---	---	---	---	---	---	---	---	---	---
Atelopus zeteki	---	DQ283252	EF494920	---	---	---	---	---	---	---	---	---	---
Barbarophryne brongersmai	FJ882817	FJ882817	---	---	FJ882718	---	FJ882663	---	---	---	---	---	---
Bufo aspinius	AF160770	AF160787	AF171198	---	---	---	---	---	---	---	---	---	---
Bufo bankorensis	AB159476	AB159589	AB159260	---	---	---	---	---	---	---	---	---	---
Bufo bufo	AY325988	FJ882806	AB159262	DQ629612	DQ306504	---	FJ882650	DQ158280	EU497611	---	---	---	---
Bufo cryptotympanicus	AF160771	AF160789	---	---	---	---	---	---	---	---	---	---	---
Bufo gargarizans	NC_008410	NC_008410	NC_008410	AF004530	FJ882707	---	FJ882653	---	---	---	---	---	---
Bufo japonicus	AB159452	AB159565	AB159257	---	---	---	---	---	---	---	---	---	---

Bufo stejnegeri	AF218710	---	---	---	---	---	---	---	---	---	---	---	---
Bufo tibetanus	AF160766	AF160784	AF171193	AY936839	---	---	---	---	---	---	---	---	---
Bufo torrenticola	AB159465	AB159578	AB270963	---	---	---	---	---	---	---	---	---	---
Bufo tuberculatus	AF160765	AF160783	AF171192	---	---	---	---	---	---	---	---	---	---
Bufo tuberospinius	AF160769	AF160788	AF174510	---	---	---	---	---	---	---	---	---	---
Bufo verrucosissimus	FJ882807	FJ882807	---	AF004526	FJ882705	---	FJ882651	---	---	---	---	---	---
Bufotes balearicus	---	EU497437	---	---	---	---	---	---	EU497608	---	---	---	---
Bufotes baturae	---	---	---	DQ629604	---	---	---	---	---	---	---	---	---
Bufotes boulengeri	---	EU497492	---	DQ629602	---	---	---	---	EU497613	---	---	---	---
Bufotes latastii	---	---	---	DQ629599	---	---	---	---	---	---	---	---	---
Bufotes luristanicus	---	---	---	DQ629610	---	---	---	---	---	---	---	---	---
Bufotes oblongus	AF160777	AF160796	AF171211	DQ629601	---	---	---	---	---	---	---	---	---
Bufotes pewzowi	GQ489032	---	---	---	---	---	---	---	GU064322	---	---	---	---
Bufotes siculus	---	EU497446	---	DQ629608	---	---	---	---	EU497609	---	---	---	---
Bufotes variabilis	GQ489041	---	---	DQ629600	---	---	---	---	EU497606	---	---	---	---
Bufotes viridis	AY680267	FJ882813	AB159263	DQ629606	FJ882714	DQ284297	FJ882659	---	EU497603	DQ283940	DQ282791	---	---
Capensibufo rosei	AF220864	AF220911	---	AF463795	---	---	---	---	---	---	---	---	---
Capensibufo tradouwi	AF220865	AF220912	---	FN650115	---	---	---	---	---	---	---	---	---
Churamiti maridadi	FJ882769	FJ882769	---	---	FJ882715	---	FJ882660	---	---	---	---	---	---
Dendrophryniscus berthalutzae	JN867524	JN867551	---	---	---	---	---	---	---	---	---	---	---
Dendrophryniscus brevipollicatus	JN867526	JN867553	---	---	---	---	---	---	---	---	---	---	---
Dendrophryniscus krausae	JN867523	JN867550	---	---	---	---	---	---	---	---	---	---	---
Dendrophryniscus leucomastyx	JN867530	JN867557	---	---	---	---	---	---	---	---	---	---	---
Dendrophryniscus oreites	JN867540	JN867567	---	---	---	---	---	---	---	---	---	---	---
Dendrophryniscus proboscidea	JN867537	JN867564	---	---	---	---	---	---	---	---	---	---	---
Didynamipus sjostedti	AY325991	AY325991	---	---	---	---	---	---	---	---	---	---	---

Duttaphrynus atukoralei	FJ882835	FJ882835	---	---	FJ882682	---	FJ882629	---	---	---	---	---	---
Duttaphrynus brevirostris	FJ882786	FJ882786	---	---	FJ882686	---	FJ882633	---	---	---	---	---	---
Duttaphrynus crocus	FJ882789	FJ882789	---	---	FJ882690	---	FJ882637	---	---	---	---	---	---
Duttaphrynus dhufarensis	FJ882837	FJ882837	---	---	FJ882679	---	FJ882626	---	---	---	---	---	---
Duttaphrynus himalayanus	FJ882790	FJ882790	AF171191	---	FJ882688	---	FJ882635	---	---	---	---	---	---
Duttaphrynus hololius	FJ882781	FJ882781	---	---	FJ882680	---	FJ882627	---	---	---	---	---	---
Duttaphrynus melanostictus	AB331714	FJ882791	AF249082	---	DQ306508	DQ284324	AY948805	DQ158317	EU712821	AF249097	DQ282815	AY948851	---
Duttaphrynus parietalis	FJ882784	FJ882784	---	---	FJ882684	---	FJ882631	---	---	---	---	---	---
Duttaphrynus scaber	FJ882785	FJ882785	---	---	FJ882683	---	FJ882630	---	---	---	---	---	---
Duttaphrynus stomaticus	FJ882787	FJ882787	AY028513	---	FJ882681	---	FJ882628	---	---	---	---	---	---
Duttaphrynus stuarti	FJ882788	FJ882788	---	---	FJ882689	---	FJ882636	---	---	---	---	---	---
Epidalea calamita	FJ882809	FJ882809	L10964	DQ629607	FJ882709	---	FJ882655	---	EU497610	---	---	---	---
Ghatophryne ornata	FJ882797	FJ882797	---	---	FJ882694	---	FJ882641	---	---	---	---	---	---
Incilius alvarius	AY325984	DQ283269	L10956	JN868006	DQ306516	DQ284289	---	AY819079	DQ158351	DQ283933	DQ282785	---	---
Incilius bocourti	AY680246	AY680246	---	JN868008	---	---	---	---	---	---	---	---	---
Incilius campbelli	---	AY008253	AY008254	JN868010	---	---	---	---	---	---	---	---	---
Incilius canaliferus	AY680251	AY680251	---	JN868013	---	---	---	---	---	---	---	---	---
Incilius cavifrons	HM563827	HM563868	HM563942	JN868014	HM563900	---	---	---	HM563986	---	---	---	---
Incilius chompipe	HM563817	HM563859	HM563932	JN868000	HM563890	---	---	---	HM563976	---	---	---	
Incilius coccifer	DQ158443	DQ158443	AY929304	JN868015	DQ306526	---	---	DQ158284	DQ158366	---	---	---	---
Incilius coniferus	DQ158445	DQ283166	AY927866	JN868001	DQ306534	DQ284204	---	DQ158286	DQ158368	DQ283860	DQ282719	---	---
Incilius cristatus	HM563830	HM563869	HM563946	JN868017	HM563903	---	---	---	HM563989	---	---	---	---
Incilius cycladen	---	AY927858	AY927865	JN868018	---	---	---	---	---	---	---	---	---
Incilius ibarrai	AY680249	AY680249	AY927862	JN868020	---	---	---	---	---	---	---	---	---
Incilius karenlipsae	---	GU552454	GU552455	---	---	---	---	---	---	---	---	---	---
Incilius leucomyos	HM563834	HM563871	HM563951	JN868022	HM563907	---	---	---	HM563993	---	---	---	---
Incilius luetkenii	DQ158467	DQ158467	AY028509	JN868025	DQ306565	---	---	DQ158308	DQ158387	---	---	---	---

Incilius macrocristatus	AY680256	AY680256	AY008252	JN868026	---	---	---	---	---	---	---	---	---
Incilius marmoreus	AY680250	AY680250	---	JN868003	---	---	---	---	---	---	---	---	---
Incilius mazatlanensis	AY680254	AY680254	DQ415638	JN868029	---	---	---	---	---	---	---	---	---
Incilius melanochlorus	AY680255	AY680255	---	JN868030	---	---	---	---	---	---	---	---	---
Incilius nebulifer	AY325985	AY680252	DQ415628	JN868031	---	---	---	---	---	---	---	---	---
Incilius occidentalis	AY680257	AY680257	---	JN868033	---	---	---	---	---	---	---	---	---
Incilius perplexus	HM563846	HM563880	HM563963	JN868034	HM563919	---	---	---	HM564005	---	---	---	---
Incilius pisinnus	HM563847	HM563881	HM563964	JN868004	HM563920	---	---	---	HM564006	---	---	---	---
Incilius porteri	HM563848	HM563882	HM563965	JN868035	HM563921	---	---	---	HM564007	---	---	---	---
Incilius signifer	HM563849	HM563883	HM563966	JN868036	HM563922	---	---	---	HM564008	---	---	---	---
Incilius spiculatus	HM563850	HM563884	HM563967	JN868037	HM563923	---	---	---	HM564009	---	---	---	---
Incilius tacanensis	AY680258	AY680258	---	JN868038	---	---	---	---	---	---	---	---	---
Incilius tutelarius	---	---	---	JN868040	---	---	---	---	---	---	---	---	---
Incilius valliceps	DQ158493	DQ158493	DQ415637	JN868041	DQ306545	---	---	DQ158333	DQ158409	---	---	---	---
Ingerophrynus biporcatus	AY325987	AY325987	---	---	---	---	---	---	---	---	---	---	---
Ingerophrynus celebensis	EU713068	AF375513	---	---	---	---	---	---	EU712916	---	---	---	---
Ingerophrynus divergens	FJ882802	FJ882802	---	---	FJ882701	DQ284189	FJ882648	---	EU712818	DQ283849	DQ282705	---	---
Ingerophrynus galeatus	FJ882767	FJ882768	---	---	DQ306506	DQ284362	FJ882649	DQ158293	DQ158374	DQ283995	DQ282839	---	DQ282987
Ingerophrynus macrotis	DQ158468	FJ882803	---	---	DQ306525	---	---	DQ158309	DQ158388	---	---	---	---
Ingerophrynus philippinicus	AF375488	---	---	---	---	---	---	---	---	---	---	---	---
Leptophryne borbonica	FJ882799	AB331716	---	---	EF107450	---	EF107224	---	EF107287	---	---	EF107357	---
Melanophryniscus devincenzii	JX961678	JX961678	---	---	---	---	---	---	---	JX961676	---	---	---
Melanophryniscus klappenbachi	AY843699	AY843699	AY843944	---	---	DQ284060	---	---	DQ503299	DQ283765	AY844899	---	---
Melanophryniscus rubriventris	AF375493	AF375518	---	---	---	---	---	---	---	---	---	---	---
Melanophryniscus setiba	JX961679	JX961679	---	---	JX961675	---	---	---	---	JX961677	---	---	---
Melanophryniscus stelzneri	FJ882853	FJ882853	---	---	DQ306494	---	AY948822	DQ158263	---	---	---	AY948878	---
Mertensophryne anotis	AF220863	AF220910	---	---	---	---	---	---	---	---	---	---	---

Mertensophryne lindneri	FJ882847	FJ882847	---	---	---	---	FJ882667	---	---	---	---	---	---
Mertensophryne loveridgei	FJ882820	FJ882820	---	---	FJ882721	---	FJ882666	---	---	---	---	---	---
Mertensophryne micranotis	FJ882821	FJ882821	---	---	EF107491	---	EF107271	---	---	---	---	EF107434	---
Mertensophryne taitana	FJ882845	FJ882845	---	---	---	---	---	---	---	---	---	---	---
Mertensophryne uzunguensis	FJ882819	FJ882819	---	---	FJ882720	---	FJ882665	---	---	---	---	---	---
Nannophryne cophotis	DQ158446	DQ158446	---	---	DQ306540	---	---	DQ158287	DQ158369	---	---	---	---
Nannophryne variegata	DQ158494	DQ158494	---	---	DQ306515	---	---	DQ158335	DQ158410	---	---	---	---
Nectophryne afra	EU394535	EU394535	---	---	---	DQ284347	---	---	---	DQ283981	DQ282831	---	DQ282977
Nectophryne batesii	EU394537	EU394537	---	---	---	DQ284207	---	---	---	---	DQ282721	---	---
Nectophrynoides minutus	FJ882814	FJ882814	---	---	---	---	---	---	---	---	---	---	---
Nectophrynoides tornieri	FJ882815	FJ882815	---	---	EF107490	DQ284394	EF107270	---	EF107329	DQ284018	DQ282870	EF107433	---
Nectophrynoides viviparus	FJ882816	FJ882816	---	---	FJ882716	---	FJ882661	---	---	---	---	---	---
Nimbaphrynoides liberiensis	GU322821	GU322838	GU322850	---	---	---	---	---	---	---	---	---	---
Nimbaphrynoides occidentalis	GU322831	GU322843	GU322852	---	---	---	---	---	---	---	---	---	---
Osornophryne antisana	AF375496	EU672983	---	---	---	---	---	---	---	---	---	---	---
Osornophryne bufoniformis	AF375497	AF375524	---	---	---	---	---	---	---	---	---	---	---
Osornophryne guacamayo	AY819334	AY326036	---	---	---	---	---	AY819083	---	---	---	---	---
Osornophryne puruanta	---	EU672982	---	---	---	---	---	---	---	---	---	---	---
Osornophryne simpsoni	JF907477	JX412035	---	---	---	---	---	---	JX412109	---	---	---	---
Osornophryne sumacoensis	AF375500	AF375527	---	---	---	---	---	---	---	---	---	---	---
Pedostibes hosii	FJ882804	AY325993	---	---	EF107449	DQ284202	EF107223	---	EF107286	DQ283859	DQ282717	EF107356	---
Pedostibes rugosus	AB331719	AB331719	---	---	---	---	---	---	---	---	---	---	---
Pedostibes tuberculosus	FJ882793	FJ882793	---	---	FJ882693	---	FJ882640	---	---	---	---	---	---
Pelophryne brevipes	AB331720	AB331720	---	---	---	---	---	---	---	---	---	---	---
Pelophryne guentheri	AY028478	AY028491	AY028503	---	---	---	---	---	---	---	---	---	---
Pelophryne misera	FJ882800	FJ882800	---	---	FJ882700	---	FJ882647	---	---	---	---	---	---
Pelophryne signata	FJ882801	FJ882801	---	---	FJ882699	---	FJ882646	---	---	---	---	---	---

Peltophryne cataulaciceps	---	JF434574	---	---	JF342406	---	---	---	---	---	---	---	---
Peltophryne empusa	AF361695	AY028489	AY028501	---	---	---	---	---	---	---	---	---	---
Peltophryne fustiger	AF361696	AF361697	AF361698	---	---	---	---	---	---	---	---	---	---
Peltophryne gundlachi	AY028479	AY028492	AY028504	---	---	---	---	---	---	---	---	---	---
Peltophryne lemur	DQ158465	DQ158465	AY028506	---	DQ306513	---	---	DQ158306	DQ158386	---	---	---	---
Peltophryne longinasus	AY028480	AY028493	AY028505	---	---	---	---	---	---	---	---	---	---
Peltophryne peltocephala	AY028477	AY028490	AY028502	---	---	---	---	---	---	---	---	---	---
Peltophryne taladai	AY028482	AY028495	AY028507	---	---	---	---	---	---	---	---	---	---
Phrynoidis aspera	DQ158432	AF124109	---	---	DQ306503	DQ284188	---	DQ158273	DQ158356	DQ283848	DQ282704	---	DQ282939
Phrynoidis juxtaspera	FJ882805	FJ882805	---	---	DQ306542	---	FJ882656	DQ158304	DQ158385	---	---	---	---
Poyntonophrynus damaranus	---	AF220906	---	AF463793	---	---	---	---	---	---	---	---	---
Poyntonophrynus dombensis	AF220857	AF220907	---	AF463794	---	---	---	---	---	---	---	---	---
Poyntonophrynus fenoulheti	AF220859	AF220908	---	---	---	---	---	---	---	---	---	---	---
Rhaebo glaberrimus	DQ158455	DQ158455	---	---	DQ306510	---	---	DQ158295	DQ158377	---	---	---	---
Rhaebo guttatus	DQ158459	DQ158459	---	---	DQ306497	DQ284361	---	DQ158300	DQ158381	DQ283994	---	---	EF364361
Rhaebo haematiticus	DQ158461	DQ158461	L10974	JN868019	DQ306501	DQ284205	---	DQ158277	DQ158383	DQ283861	DQ282720	---	---
Rhaebo nasicus	DQ158477	DQ158477	---	---	DQ306512	---	---	DQ158319	DQ158396	---	---	---	---
Rhinella abei	---	---	---	KC198125	---	---	---	---	---	---	---	---	---
Rhinella achavali	GU178787	GU178787	GU178809	---	---	---	---	---	---	---	---	---	---
Rhinella amboroensis	DQ283386	DQ283386	---	---	---	---	---	---	---	DQ284003	DQ282848	---	---
Rhinella arenarum	AY843573	AY843573	AY843795	---	DQ306529	DQ284103	---	DQ158271	DQ158354	---	AY844775	---	---
Rhinella arequipensis	DQ158430	DQ158430	---	---	DQ306564	---	---	DQ158272	DQ158355	---	---	---	---
Rhinella arunco	DQ158442	DQ158442	---	---	DQ306552	---	---	DQ158283	DQ158365	---	---	---	---
Rhinella atacamensis	DQ158433	DQ158433	---	---	DQ306541	---	---	DQ158275	DQ158357	---	---	---	---
Rhinella castaneotica	DQ158440	DQ158440	---	---	DQ306539	---	---	DQ158282	DQ158364	---	---	---	---
Rhinella crucifer	DQ158447	DQ158447	DQ415596	KC198090	---	---	---	DQ158288	---	---	---	---	---
Rhinella dapsilis	DQ158448	DQ158448	---	---	DQ306532	---	---	DQ158289	DQ158370	---	---	---	---

Rhinella festae	DQ158423	DQ158423	---	---	DQ306521	---	---	DQ158265	---	---	---	---	---
Rhinella granulosa	FJ882774	AY028496	AY028508	---	DQ306557	DQ284323	FJ882673	DQ158298	DQ158380	DQ283966	---	---	---
Rhinella henseli	---	---	---	KC198218	---	---	---	---	---	---	---	---	---
Rhinella humboldti	EF532251	---	---	---	---	---	---	---	---	---	---	---	---
Rhinella icterica	DQ158462	DQ158462	GU178808	GU907269	---	---	---	DQ158303	DQ158384	---	---	---	---
Rhinella inopina	---	---	---	KC198367	---	---	---	---	---	---	---	---	---
Rhinella limensis	DQ158466	DQ158466	---	---	DQ306509	---	---	DQ158307	---	---	---	---	---
Rhinella margaritifera	AY819331	AF375514	---	JN868028	---	---	GQ345223	AY819080	---	---	---	GQ345321	EF364358
Rhinella marina	AY325994	AY325994	DQ415597	GU907290	DQ306544	DQ284092	---	AF194966	DQ158393	DQ283789	---	---	---
Rhinella merianae	---	---	---	GU907291	---	---	---	---	---	---	---	---	---
Rhinella nesiotes	DQ158478	DQ158478	---	---	DQ306500	---	---	DQ158320	DQ158397	---	---	---	---
Rhinella ocellata	DQ158479	DQ158479	---	---	DQ306538	---	---	DQ158321	DQ158398	---	---	---	---
Rhinella ornata	---	---	---	KC198091	---	---	---	---	---	---	---	---	---
Rhinella poeppigii	DQ158481	DQ158481	---	---	DQ306517	---	---	DQ158323	DQ158400	---	---	---	---
Rhinella pombali	---	---	---	KC198198	---	---	---	---	---	---	---	---	---
Rhinella rubescens	---	---	---	GU907265	---	---	---	---	---	---	---	---	---
Rhinella schneideri	FJ882831	DQ158480	DQ415598	---	DQ306528	DQ284102	FJ882674	DQ158322	DQ158399	---	---	---	---
Rhinella spinulosa	DQ158487	DQ158487	---	---	DQ306566	DQ284077	---	DQ158328	DQ158405	DQ283775	DQ282658	---	---
Rhinella vellardi	DQ158495	DQ158495	---	---	DQ306527	---	---	DQ158336	DQ158411	---	---	---	---
Rhinella veraguensis	DQ158497	DQ158497	---	---	DQ306524	---	---	DQ158337	DQ158412	---	---	---	---
Sabahphrynus maculatus	AB331718	AB331718	---	---	---	---	---	---	---	---	---	---	---
Schismaderma carens	AY325997	AY325997	---	---	FJ882717	DQ284403	FJ882662	DQ158266	---	DQ284027	DQ282882	---	---
Strauchbufo raddei	AF160776	GU183855	AF171210	---	---	---	---	---	---	---	---	---	---
Vandijkophrynus amatolicus	AF220851	AF220898	---	---	---	---	---	---	---	---	---	---	---
Vandijkophrynus angusticeps	AF220852	AF220899	---	---	---	---	---	---	---	---	---	---	---
Vandijkophrynus gariepensis	AF220853	AF220902	---	AF463792	---	---	---	---	---	---	---	---	---
Vandijkophrynus inyangae	AF220856	AF220904	---	---	---	---	---	---	---	---	---	---	---

Vandijkophrynus robinsoni	GU183857	GU183857	---	---	---	---	---	---	---	---	---	---	---
Werneria mertensiana	DQ283348	DQ283348	---	---	---	DQ284338	---	---	---	DQ283974	DQ282824	---	---
Wolterstorffina parvipalmata	FJ882818	FJ882818	---	---	FJ882719	DQ284334	FJ882664	---	---	DQ283972	DQ282822	---	---
Xanthophryne koynayensis	FJ882782	FJ882782	---	---	FJ882691	---	FJ882638	---	---	---	---	---	---

Chapter 6

Historical biogeography of the Arabian Peninsula: A review of geologic and climatic-based biogeographic hypotheses and the role of the Arabian Hotspot Area

Abstract

Aim

We summarize the various geologic and climatic-based hypotheses influencing the diversification of Arabian taxa and highlight the role of an understudied biodiversity hotspot, the Arabian Hotspot Area (AHA), in driving patterns of endemism. We evaluate these biogeographic hypotheses and patterns utilizing Arabian amphibians as a study system, and contrast these results with a comprehensive literature review of other Arabian taxa.

Location

The Arabian Peninsula, Horn of Africa, Levant, and South Asia

Methods

We investigated biogeographic hypotheses using a temporally explicit phylogenetic framework and by compiling divergence-dating estimates for all available extant Arabian taxa. To examine the effects of climate change, we generated distribution models for Arabian amphibians under current and historical climate regimes to identify historically stable refugial areas for montane and lowland species.

Results

We find a striking concordance of divergence times for taxa with similar biogeographic histories, providing strong support for hypotheses derived from geologic events. Most in-situ speciation events occurring on the Arabian Peninsula are concentrated in the late Miocene, with intraspecific divergences of species occurring mainly in the Plio-Pleistocene.

Main Conclusions

Changes in connectivity between the Arabian Peninsula and surrounding continents throughout geologic time allows explicit temporal predictions to be made about biogeographic events for Arabian taxa derived from African, Asian, and Palearctic lineages. We find the AHA is in many ways similar to the montane systems of East Africa, and is characterized by habitat stability and the long-term persistence of localized lineages and the diversification of cyclically fragmented lineages. Historically stable refugial regions for amphibian species mainly occur within the AHA, demonstrating the importance of this region for the persistence of both montane and lowland species.

Introduction

The Arabian Peninsula harbors a unique and rich assemblage of species that has been sculpted by the complex geologic and climatic history of this region. The diversity of landscapes occurring across the Arabian Peninsula, ranging from high altitude montane woodlands to hyperarid gravel plains, results from a combination of distinct

topographical features and a variable climate. This landscape diversity is reflected in a high richness of vascular plants (>3500 species) that exhibit African, Palearctic, and Asian influences (Vesey-Fitzgerald, 1955, 1957; Miller & Cope, 1996; Fisher *et al.*, 1998; Fisher & Membery, 1998; Kürschner, 1998). Both the greatest number of plant species and the highest concentration of endemic species of the Arabian Peninsula occur in the mountainous areas along the southern and southwestern coasts (Miller & Cope, 1996; Fisher *et al.*, 1998). As a result these mountains and the surrounding coastal areas were defined as disjunct extensions of the Horn of Africa Hotspot and the Eastern Afromontane Hotspot (Mittermeier *et al*., 1999, 2004; Myers *et al*., 2000), two globally recognized sectors defining areas of high species diversity and endemism across mainland Africa (Fig. 28). In an effort to recognize the distinctiveness of the Arabian Peninsula hotspot extensions, the collective area these two hotspots encompass on the Peninsula has more recently been referred to as the Arabian Hotspot Area (AHA) (Mallon, 2011).

The general patterns of plant species richness and endemism are mirrored by the vertebrate community of the Arabian Peninsula. Although the proportion of vertebrate species endemic to the Peninsula varies across groups (6–78%), over half of the peninsular endemics are restricted specifically to the Arabian Hotspot Area (Mallon, 2011). This pattern is most striking in lizards, snakes, amphibians, and freshwater fish, which account for 80% of the endemic vertebrate species occurring in the AHA. The Arabian Hotspot Area has long been recognized as an important center of diversity for reptiles, amphibians, and freshwater fish (Balletto *et al*., 1985; Arnold, 1987; Joger, 1987; Gasperetti, 1988; Schatti & Desvoignes, 1999), however the role of this region in promoting diversification or persistence has not been investigated.

Diversification on the Arabian Peninsula

The species diversity and localized endemism in the AHA has been driven by a variety of biotic and abiotic factors, but the complex geological history of the Arabian Peninsula has had the most prominent effect in shaping major biogeographic patterns. Tectonic activity driving the separation of the Arabian Plate from Africa began approximately 30 Ma (Fig. 29a), and initial rifting and formation of the southern Red Sea occurred in the early Miocene (27 Ma), which resulted in partial embayment from the connection of the Eritrean Red Sea to the Gulf of Aden through the Strait of Bab el Mandeb (Bohannon, 1986; Bosworth *et al*., 2005). Through a second phase of volcanism, the northward extension of the Red Sea established a connection to the Neotethys Sea by 23 Ma, along with uplifting of the continental shoulders (Fig. 29b). A counterclockwise rotation of the Arabian plate caused a collision with the Anatolian plate approximately 18–16 Ma and established a temporary land bridge, termed the *Gomphotherium* bridge, allowing limited Palearctic and Afro-Arabian faunal exchanges (Rögl, 1998, 1999) (Fig. 29c). A subsequent collision with Eurasia established a more permanent northern land bridge by 15 Ma (Bosworth *et al*., 2005), allowing faunal exchanges between North Africa, the Arabian Peninsula, and Eurasia (Fig. 29d). Massive halite deposition across the entire Red Sea occurred 10 Ma, indicating a likely closing of the Strait of Bab el Mandeb and establishment of a southern land bridge between Africa and Arabia (Bosworth *et al*., 2005) that would have facilitated dispersal between the Horn of Africa and the Arabian Peninsula (Fig. 29e). This land bridge was maintained until around 5.3 Ma, when

Pliocene marine sediment deposits formed, indicating a marine connection was reestablished through Bab el Mandeb (Bosworth *et al*., 2005) (Fig. 29f). Although temporary Pleistocene land bridges across the Strait of Bab el Mandeb resulting from glacial cycles have been proposed (Rohling *et al*., 1998), their existence is not strongly supported (Fernandes *et al*., 2006). However, there is evidence that during the late Pleistocene the Persian Gulf was reduced to a series of freshwater lakes (Lambeck, 1996) or completely waterless (Uchupi *et al*., 1999), allowing for more recent dispersal events.

In addition to large-scale geologic events, it has also been proposed that Plio-Pleistocene climate changes have driven speciation in the montane regions of the Arabian Hotspot Area (Miller & Cope, 1996; Fisher *et al.*, 1998). The overall aridification of this region began during the Late Neogene (deMenocaol, 1995, 2004) and caused fragmentation of the once continuous montane habitats into smaller refugial areas (Miller & Cope, 1996; Kürschner, 1998; Fisher *et al.*, 1998). Climatic fluctuations in the Pleistocene affected monsoonal patterns in the AHA (Sirocko, 1996; Guba & Glennie, 1998; Lézine *et al*., 1998; Luckge *et al*., 2001; Glennie & Singhvi, 2002; Jung *et al*., 2004), causing repeated expansion and contraction of these refugial habitats. These recent cyclical changes in connectivity between refugial habitats are hypothesized to have driven a majority of the speciation in freshwater fish (Krupp, 1983), whereas the fragmentation of montane habitats at both ancient (Miocene) and recent (Plio-Pleistocene) time scales is postulated to be a main driver of speciation in plants of the AHA (Miller & Cope, 1996; Fisher *et al.*, 1998; Kürschner, 1998). The limited number of empirical studies of plants supports this pattern (Meister *et al*., 2005, 2006; Oberprieler *et al*., 2009), but the effects of these events on vertebrate groups have yet to be examined.

The accumulation of species diversity on the Arabian Peninsula may therefore be a consequence of isolation of relict populations, build-up via colonization, and within system diversification, but the relative contributions of these processes remain largely unexplored. These patterns are particularly important to understand for the Arabian Hotspot Area, where montane regions and associated coastal habitats could act as centers for lineage diversification and lineage persistence across different time scales. In this review, we first focus on a particular vertebrate system, utilizing amphibians as a focal group, to investigate the temporal, biogeographic, and climatic factors driving the accumulation of species in the AHA. We then examine these results in the context of other Arabian taxa that have been investigated in a temporally explicit phylogenetic framework, and further evaluate the various hypotheses proposed.

Study System

The amphibians of the Arabian Peninsula exhibit a high level of endemism, their ranges are either restricted to or broadly overlap the AHA, and they occur at both low and high elevations (Balletto *et al*., 1985). The dispersal ability of amphibians in the AHA is limited by their reliance on freshwater habitat for breeding and their intolerance to saltwater, and amphibians are also highly sensitive to environmental changes (Stuart *et al*., 2004). For these reasons, amphibians are an appropriate system to study in the proposed biogeographic framework.

There are nine described anuran species in four families present on the Arabian Peninsula, with seven confirmed endemics and two lineages of debated taxonomic status. Following the taxonomy of AmphibiaWeb (2014) and Frost (2014), these species are:

Pelophylax cf. *bedriagae* (Ranidae), *Hyla felixarabica* (Hylidae), *Euphlyctis ehrenbergii* (Dicroglossidae), and six true toads (Bufonidae) including *Duttaphrynus dhufarensis, Bufotes* cf. *variabilis*, *Amietophrynus arabicus, Amietophrynus tihamicus, Amietophrynus hadramautinus*, and *"Bufo" scortecci*. These species vary greatly in their geographic range size and elevation, with *E. ehrenbergii*, *D. dhufarensis*, and *Amietophrynus arabicus* possessing the largest geographic distributions and broadest elevational ranges (sea level to 2400 m, Fig. 30). Other species, such as *P.* cf. *bedriagae*, *H. felixarabica*, and *Bufotes* cf. *variabilis* are restricted to higher elevations throughout the Scarp Mountains (1400 to 2800 m, Fig. 30), whereas *Amietophrynus tihamicus* is a lowland species (< 350 m) distributed along the coastal Red Sea region. Both *Amietophrynus hadramautinus* and *B. scortecci* are restricted to one or two localities, and their validity has been questioned as they may represent isolated phenotypically variable populations of the wider ranging *Amietophrynus arabicus* and *D. dhufarensis*, respectively (Balletto *et al*., 1985). The complex topography of the AHA facilitates the formation of different amphibian communities, and up to four species have been recorded in sympatry at low and high elevation sites, with all sites of co-occurring species including at least one toad lineage (Balletto *et al*., 1985).

Three of the Arabian species are inferred to be geographical isolates and have been attributed to a continental origin on the basis of their presumed generic affinites (*P.* cf. *bedriagae* and *H. felixarabica* = Palearctic, *E. ehrenbergii* = Asian). These patterns were confirmed in part by Gvoždík *et al*. (2010), who described *H. felixarabica* as distinct from the Middle Eastern *H. savignyi*, to which it was formerly attributed, and by Wiens *et al.* (2009), who identified a sister relationship between *E. ehrenbergii* and the widespread Asian species *E. cyanophlyctis*. One Arabian toad species, *Bufotes* cf. *variabilis*, is attributable to the Palearctic *Bufotes viridis* species complex on the basis of morphology (Balletto *et al*., 1985; Stöck *et al*., 2006). The Arabian endemic *Duttaphrynus dhufarensis* was recently determined to be of Asian origin (Van Bocxlaer *et al*., 2009), and Portik and Papenfuss (2015) recently demonstrated the African origins of *Amietophrynus arabicus* and *tihamicus*.

Using Arabian amphibians as a study system, our main objective is to test the various geologic and climate-driven hypotheses influencing species accumulation on the Arabian Peninsula. To do this, we employ a combination of analyses to infer phylogenetic relationships, dates of divergence, and distributional models. We use multi-locus molecular data to reconstruct the phylogenetic relationships of hylids and dicroglossids with expanded sampling. We estimate divergence times for these Arabian taxa and use published estimates for Arabian toads (Portik & Papenfuss 2015) to identify the historical events or processes that may have promoted speciation. We model the potential distributions of all Arabian amphibian species under current and historical climate regimes to examine the effects of past climate change on range size and connectivity. These models are compared to identify historically stable refugial areas that may have facilitated long-term lineage persistence for taxa with differing elevational distributions. Based on our results, we identify key biogeographic events and historical processes that have influenced amphibian species diversity on the Arabian Peninsula. Lastly, we summarize all available divergence date estimates for all other Arabian taxa in an effort to explore overall diversification patterns of the Arabian Peninsula and identify key differences between patterns of amphibians and other major groups.

Methods

Sampling

Sampling for hylid frogs relied on published data, and includes 57 samples from 20 ingroup taxa. Samples were chosen based on the availability of sequence data for three loci, including the nucDNA genes Tyrosinase (*tyr*), and Rhodopsin (*rho*), and the mtDNA genes 16S and 12S. Sampling for dicroglossids includes newly generated sequence data for *Euphlyctis ehrenbergii* and published sequence data. We incorporate 15 ingroup taxa, for a total of 32 terminals. Samples were chosen based on the availability of sequence data for three loci: Tyrosinase (*tyr*), Rhodopsin (*rho*), and 16S. Voucher and locality information for all specimens is provided in Tables 20–22.

Phylogenetic Analyses

We used PartitionFinder (Lanfear *et al.*, 2012) to simultaneously determine our best partitioning strategy and models for each partition subset. The greedy search algorithm was employed, and model selection was conducted using BIC. The best partitioning scheme for the hylid data set includes three partitions: *tyr* cp3: HKY+G; *rho* cp1-3, *tyr* cp 1&2: K80+I+G; 12S & 16S: GTR+I+G; and the dicroglossid data set includes three partitions: *tyr* cp 1&2, *rho* cp 1&2: HKY+I; : *tyr* cp 3, *rho* cp 3: HKY+G; 16S: GTR+I+G. To resolve the phylogenetic placement of Arabian focal taxa, we analyzed our molecular data using maximum likelihood and Bayesian analyses. Using MrBayes v3.2 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003), parallel runs utilizing four MCMC chains were allowed to run for 2 x 10^7 million generations, with sampling every 1000 generations. Runs were assessed using Tracer v1.5.0 (Rambaut & Drummond, 2009) to ensure key parameters had reached stationarity (ESS values >150). The first 25% of the total number of generations were discarded as burn-in and a 50% majority rule consensus tree was calculated from the remaining trees (30,000). We performed maximum likelihood analyses of the partitioned data set using GARLI v2.0 (Zwickl, 2006). Using default parameters in the ML search algorithm, 10 replicate searches for the best point estimate topology were conducted, and the tree with the best likelihood was selected as a guide tree for the bootstrap analyses. The Garli Web Service (Bazinet & Cummings, 2011), functioning on the Lattice Project grid system (Bazinet & Cummings, 2008), was used to execute 1000 nonparametric bootstrap replicates asynchronously in parallel using the default stopping criteria. A 50% majority rule consensus tree was generated from the 1000 replicates.

Divergence Dating Analyses

To infer the timing of lineage divergences of sampled Arabian amphibian taxa, we carried out dating analyses in Beast v1.7.5 (Drummond *et al.*, 2012). We conducted separate analyses for hylids and dicroglossids using calibration strategies described below. Dating analyses for all focal groups were run for 6 x 10^7 generations, with sampling every 3000 generations. For all analyses, we used the Yule model of speciation as our tree prior, applied an uncorrelated relaxed lognormal clock, and unlinked clock and substitution models. Substitution models and partitioning schemes were identical to those used in phylogenetic analyses. Runs were assessed using Tracer v1.5.0 (Rambaut & Drummond, 2009) to examine convergence. A burn-in of 25% was discarded and

maximum clade credibility trees were created from a total of 15,000 trees for each analysis.

Divergence dates for *H. felixarabica* were previously estimated by Gvoždík *et al*. (2010) using a mitochondrial substitution rate for hylid frogs (Moriarty Lemmon *et al*., 2007), which resulted in large confidence intervals for nodes of interest. We attempted to improve on this strategy by incorporating nuclear loci and by using internal calibration points in our analysis. We employed two internal calibrations recommended by Smith *et al.* (2007) that place minimum constraints on divergences between European and Asian *Hyla* (16 Ma) and divergences between *H. squirella* and *H. cinerea* (15 Ma) based on fossil evidence (Sanchiz 1998; Holman 2003). Calibrations to relevant nodes were applied using an offset lognormal distribution (real space mean: 17 Ma, stdev 1.5, offset 16; 16 Ma, stdev 1.5, offset 15, respectively).

There are no internal fossil calibrations available for the family Dicroglossidae, and we therefore relied on using previously published divergence estimates to place a broad prior on the divergence of dicroglossids from included outgroups. Based on Roelants *et al*. (2004), Bossuyt *et al*. (2006) and Wiens *et al*. (2009), we used a calibration of ~73 Ma for the divergence of Dicroglossidae. This node was given a normal distribution and we tested several standard deviation values (1–5) to incorporate the full range of dates retrieved in prior work (Roelants *et al*., 2004; Bossuyt *et al*., 2006; Wiens *et al*., 2009).

Species Distribution Modelling

Species distribution models (SDMs) were developed for each species using locality data compiled primarily from Balletto *et al*. (1985) and Gvoždík *et al*. (2010) using MaxEnt v. 3.3 (Phillips *et al*., 2006). For all but three species (*Amietophrynus hadramautinus*, *"Bufo" scortecci, Pelophylax* cf. *bedriagae*) we were able to find a sufficient number of unique localities for model building (n > 10; Hernandez *et al*., 2006). Current climatic conditions were estimated using the WorldClim global database (Hijmans *et al*., 2005), and paleoclimatic conditions (LGM) were generated using the MIROC 3.2 gcm model (Braconnot *et al*., 2007). Climate data were utilized at 30 arc-second resolution. SDMs were constructed using the maximum entropy algorithm implemented in MaxEnt v. 3.3 (Phillips *et al*., 2006). Highly correlated climatic parameters were identified using a Pearson's correlation and then removed, leaving ten variables for model construction. These variables include annual mean temperature, temperature seasonality, mean temperature of the wettest quarter, mean temperature of the driest quarter, mean temperature of the warmest quarter, annual precipitation, precipitation seasonality, precipitation of driest quarter, precipitation of warmest quarter, and precipitation of coldest quarter. For each model 10,000 random background points were generated. To assess model performance, occurrence data were randomly partitioned into training (75%) and testing (25%) data sets and then 100 replicate bootstrap runs were run. The area under the curve (AUC) score was used to assess model performance and interpreted as suggested by Swets (1988). In order to produce binary presence/absence assignments for each grid cell, the threshold that maximized Kappa was applied. The resulting presence-absence maps were then summed to estimate range stability through time with values for each grid cell ranging from two (predicted presence during both times) to zero (predicted absence during both times). Areas where a species was predicted to be present

at both time periods were considered potential climatic refugia. To ascertain if there were common climatic refugia for species that currently occupy similar elevation limits, we also looked for stability surfaces in common among highland species and low to mid-elevation distributions. Analyses were performed using ArcGIS 9.3 (ESRI, Redlands, CA, USA).

Results

Phylogenetic Relationships and Divergence Estimates

The phylogenetic affinities of *Hyla felixarabica*, based on increased sampling and multi-locus data, are consistent with the results of Gvoždík *et al*. (2010), which place *H. savignyi* and *H. felixarabica* as sister taxa (Fig. 31). Our dating analysis for the genus *Hyla* estimates the main split between Old and New World lineages occurred approximately 26.0 Ma (95% HPD: 20.5–34.3 Ma; Fig. 31). The divergence between *Hyla felixarabica* and *H. savignyi* is estimated to have occurred around 6.5 Ma (95% HPD: 3.9–9.5 Ma), a slightly younger date than that retrieved by Gvoždík *et al*. (2010) (8.4 Ma; 95% HPD: 3.2–18.2 Ma). The disjunct populations of *H. felixarabica* occurring in the Levant versus Saudia Arabia and Yemen are estimated to have diverged in the Late Pliocene, approximately 2.1 Ma (95% HPD: 0.9–3.7 Ma).

Similar to Wiens *et al.* (2009), but with additional molecular data for the genus *Euphlyctis*, we also recover a sister taxon relationship between the Arabian species *Euphlyctis ehrenbergii* and the Asian species *E. cyanophlyctis* (Fig. 31). Based on our dicroglossid data set, the age estimate recovered in our analyses for the origin of the subfamily Dicroglossinae is approximately 66.0 Ma (95% HPD: 58.3–72.5 Ma; Fig. 31). The divergence between the Arabian species *Euphlyctis ehrenbergii* and the broadly distributed Asian *E. cyanophlyctis* is estimated at 14.7 Ma (95% HPD: 10.2–19.8 Ma).

Distribution Modelling

SDMs had high AUC scores (>0.8) indicating that they had high discriminatory power, and were able to capture environmental variables distinguishing presence from absence. Bootstrap replicates were largely concordant, with most replicates providing similar predictive estimates. We were able to identify areas that are predicted to have remained climatically stable and suitable for each species, and these varied in extent and location. Among high elevation species, refugia were identified in the mountains of Yemen and Saudi Arabia following the Eastern Afromontane Hotspot, while for low to mid-elevation species refugia primarily fall within the Horn of Africa Hotspot (Fig. 32).

Discussion

Biogeographic and temporal diversification of Arabian amphibians

Our phylogenetic and divergence dating analyses reveal a complex biogeographic scenario that has resulted in the punctuated accumulation of amphibian lineages on the Arabian Peninsula. We find evidence for both major geological events and Plio-Pleistocene climate change affecting the diversification of Arabian amphibians, although at different phylogenetic scales.

Effects of geological events. We find that amphibian species with similar biogeographic origins display concordance in divergence timings, which can be interpreted in the

context of major historical geologic events. The species *Amietophrynus arabicus* and *A. tihamicus* are both estimated to have diverged from their closest African relatives in the Early Miocene, approximately 20.5–21.2 Ma (95% HPD range: 14.2–29.2 Ma) and 16.8–17.9 Ma (95% HPD range: 9.0–26.6 Ma), respectively (Portik & Papenfuss 2015). These timings rule out dispersal across a southern Afro-Arabian landbridge (10–5.3 Ma), but are consistent with the separation of the Arabian Plate from mainland Africa as a result of the Red Sea formation (27–23 Ma; Fig. 29) (Bosworth *et al*., 2005), suggesting this major vicariance event drove the initial divergences between *A. arabicus* and *A. tihamicus* and their African relatives. The Asian derived species *Duttaphrynus dhufarensis* and *Euphlyctis ehrenbergii* are both estimated to have diverged in the mid-Miocene, 16.0 Ma (95% HPD: 10.6–22.4 Ma) and 14.7 Ma (95% HPD: 10.2–19.8 Ma), respectively (Portik & Papenfuss 2015). The temporary connections between the Arabian plate and Eurasia established 18–16 Ma are thought to have allowed the first faunal exchanges between these distinct biogeographic regions, with a permanent Eurasian connection being established ~15 Ma (Rögl, 1998, 1999; Bosworth *et al*., 2005). The divergence estimates for *D. dhufarensis* and *E. ehrenbergii* are congruent with temporal range of these events, and the initial land bridges likely served as dispersal routes for these lineages to colonize the Arabian Peninsula (Fig. 29).

Of all the Arabian species investigated, *Hyla felixarabica* is the most recent arrival to the Peninsula, having diverged from its Middle Eastern relative *H. savignyi* approximately 6.5 Ma (95% HPD: 3.9–9.5 Ma). By this time the connection between Arabia and Eurasia was well established, along with the mountain corridors extending along the Red Sea coast to the Jordan Rift Valley, which are postulated to have served as a dispersal corridor for Palearctic montane taxa (Balletto *et al*., 1985). Material for *Bufotes* cf. *variabilis* and *Pelophylax* cf. *bedriagae* from Arabian localities is currently lacking, but if they are indeed conspecific with the nominotypical species then they closely match the distribution of *H. felixarabica* in displaying a disjunct Levantine-Arabian distribution. Dating the split between the Levantine-Arabian populations is not possible, but estimating the divergence of the typical species and its closest relative places an upper age limit on the Arabian populations. The divergence estimates of typical *B. variabilis* and *P. bedriagae* sampled from the Levant and Anatolia and their closest relatives are similar to dates retrieved for *H. felixarabica*. Our bufonid dating analysis estimates a break between *B. variabilis* and *B. viridis* occurred 4.5 Ma (95% HPD: 2.5–11.8 Ma; not shown), whereas Akın *et al.* (2010) estimate the divergence between *P. bedriagae* and the *P. ridibundus* complex occurred 7.19 (95% HPD: 5.8–8.5 Ma). The strong overlap in divergence dates among Palearctic taxa with Arabian populations suggests repeated invasions of Arabian montane regions occurred in the late Miocene and early Pliocene.

Plio-Pleistocene climate effects. The accumulation of amphibian species on the Arabian Peninsula appears to be mainly driven by major geologic events, but the effects of climate change on the Peninsula in the Late Neogene are detectable in the form of intraspecific divergences between isolated populations. Geographically isolated populations have been sampled for *Hyla felixarabica* and *Amietophrynus arabicus*, allowing for an initial comparison of intraspecific divergences in both high-elevation and wide-ranging species, respectively.

The montane species *Hyla felixarabica* is hypothesized to have had a formerly continuous range throughout the mountains stretching from Yemen to the Jordan Rift Valley (Balletto *et al*., 1985), but currently displays a disjunct Levantine-Arabian distribution (Gvoždík *et al*., 2010). We estimate the divergence between these regional populations occurred during the Pliocene, approximately 2.1 Ma (95% HPD: 0.9–3.7 Ma). This substantial break in range connectivity was most likely a result of continued aridification of the intervening mid-elevation mountains and foothills occurring between these two regions during the Late Neogene. The current and paleoclimate distribution models for *H. felixarabica* demonstrate a lack of suitable habitat in this intervening region, supporting the notion that large-scale range fragmentation occurred before the Pleistocene. A similar break in this region is also apparent in the current and paleodistribution models for the montane species *Bufotes* cf. *variabilis* and *Pelophylax* cf. *bedriagae*, suggesting Pleistocene climatic fluctuations did not facilitate dispersals between the Levantine and Arabian populations of these Palearctic-derived species.

The bufonid species *Amietophrynus arabicus* has a broad elevational range (sea level to 2800 m) and is distributed mainly along the Red Sea region throughout the AHA, although a disjunct population also occurs in the North Oman Highlands (NOH). We estimate the divergence of these regional populations also occurred during the Pliocene, approximately 4.3 Ma (95% HPD: 2.7–6.6 Ma). The vast geographic area between the AHA and NOH is comprised mainly of xeromorphic shrublands, grasslands, and gravel plains. Despite exhibiting many climatic characteristics found across the range of *A. arabicus*, this region does not demonstrate the hydrological stability afforded by the AHA and, to a lesser extent, the NOH (Fisher & Membery, 1998; Kürschner, 1998). The expansion of these drier habitats is closely tied to aridification occurring in the Late Neogene, and our divergence estimates suggest this process has driven the geographic isolation of the AHA and NOH populations of *A. arabicus*.

Comparative Biogeographic Patterns

The Arabian amphibian community has been strongly influenced by geologic events and to a lesser extent climate change, but how do these temporal and biogeographic patterns contrast with other groups of vertebrates and plants? To investigate this topic, we compiled all available molecular divergence-dating estimates for inter- and intraspecific divergences in reptiles (Amer & Kumazawa, 2005; Pook *et al*., 2009; Wong *et al*., 2010; Metallinou *et al*., 2012; Portik & Papenfuss, 2012; Šmíd *et al*., 2013), mammals (Zinner *et al*., 2009; Fernandes, 2011), and plants (Meister *et al.*, 2006; Yu *et al.*, 2014). The resulting data can be classified into two broad categories: events relating to continental exchanges, and events related to in-situ diversification on the Arabian Peninsula (Fig. 33, Table 23), which we summarize below.

Continental faunal exchanges. There are several species divergences that are concordant with dating estimates recovered for Arabian amphibians, and there are also additional biogeographic events represented in our full data set of amphibians, reptiles, and mammals (Fig. 33). The oldest Arabian species divergences estimated are early Miocene in age and occur in African-derived taxa, including bufonids as well as the viper *Echis coloratus* (Pook *et al*., 2009). These speciation events likely result from vicariance driven by the formation of the Red Sea (27–24 Ma), splitting species with formerly

continuous Afro-Arabian distributions (Fig. 29b). The resulting faunal assemblage occurring on the Arabian shield during this time was largely isolated until the formation of temporary and permanent land bridges with Eurasia 18–15 Ma allowed initial dispersals between Asia and Arabia (Figs. 29, 32). Our data include examples of amphibian dispersals from Asia to Arabia, and these timings are concordant with the dispersal of the plant genus *Isodon* from South Asia across the Arabian Peninsula to the Horn of Africa (Yu *et al.*, 2014). Additionally, the gekkonid genus *Hemidactylus* dispersed from Arabia to Asia during this time period (Šmíd *et al*., 2013). A second pulse of African fauna was introduced to Arabia through the formation of the southern land bridge between the Horn of Africa and the southwestern Arabian Peninsula (10–5.3 Ma; Fig. 29e). This land bridge allowed a series of exchanges between Africa and Arabia, introducing additional African taxa to the southwestern Arabian Peninsula, and also allowed opportunities for the diversified Arabian fauna to colonize the Horn of Africa (Fig. 33). Our data demonstrate that during this time Palearctic-derived taxa also began to colonize the Arabian Peninsula (Fig. 33). After the break up of the southern land bridge, faunal exchanges between the Horn of Africa and the Arabian Peninsula continued as Plio-Pleistocene overwater dispersals (5.3 Ma to present; Fig. 29f), and have been proposed for snakes, turtles, large mammals, and potentially monitor lizards (Fig. 33) (Pook *et al*., 2009; Zinner *et al*., 2009; Wong *et al*., 2010; Fernandes, 2011; Portik & Papenfuss, 2012). During the Pleistocene the Persian Gulf region was greatly reduced or dry (Lambeck, 1996; Uchupi *et al*., 1999), and there is evidence that the South Asian viper species *Echis carinatus* established an Arabian population through this route (Pook *et al*., 2009).

Compared with other taxa with similar biogeographic origins, the amphibians appear to have been unsuccessful in utilizing late Miocene and Plio-Pleistocene dispersal routes to establish permanent populations on the Arabian Peninsula. Many of the African-derived terrestrial vertebrate taxa present in the southwestern Arabian Peninsula arrived by dispersing from the Horn of Africa beginning in the late Miocene (Pook *et al*., 2009; Zinner *et al*., 2009; Wong *et al*., 2010; Fernandes, 2011; Portik & Papenfuss, 2012), rather than originating as a result of the initial separation of the Arabian plate from Africa (Fig. 33). The African-derived Arabian bufonids (*Amietophrynus arabicus*, *A. tihamicus*) currently represent two of the oldest lineages occurring on the Arabian Peninsula (Portik & Papenfuss 2015), and the viper species *Echis coloratus* is the only other investigated example of an ancient African relict, having also diverged in the early Miocene (Pook *et al*., 2009). We do not find evidence for amphibian dispersals between South Asia and the Arabian Peninsula occurring in the Pleistocene, during which the Persian Gulf was greatly reduced or dry (Lambeck, 1996; Uchupi *et al*., 1999). This is surprising given this region is predicted as historically suitable habitat by the paleodistribution models for *A. arabicus* and *Duttaphrynus dhufarensis*, and the South Asian viper species *Echis carinatus* appears to have colonized the Arabian Peninsula through this route (Pook *et al*., 2009). Future comparative work on taxa with disjunct South Asian-Arabian distributions may help clarify the importance of the Persian Gulf region in facilitating Pleistocene dispersals.

In-situ diversification. Based on data from gekkonids, agamids, and vipers (Amer & Kumazawa, 2005; Pook *et al*., 2009; Metallinou *et al*., 2012; Šmíd *et al*., 2013),

speciation events occurring on the Arabian Peninsula appear to be primarily concentrated in the late Miocene and early Pliocene (Fig. 33). The continued aridification of the Peninsula during this time facilitated the expansion of xeric habitats, which drove diversification in the larger reptile groups through both the expansion of suitable habitats for arid clades (gekkonid genus *Stenodactylus*) and by the fragmentation and isolation of mesic habitats affecting more temperate clades (gekkonid genus *Hemidactylus*) (Metallinou *et al*., 2012; Šmíd *et al*., 2013). Intraspecific divergences occurring within Arabian species are concentrated in the Plio-Pleistocene era, and include examples from gekkonids, amphibians, vipers, and coastal shrubs (Meister *et al.*, 2006; Pook *et al*., 2009; Šmíd *et al*., 2013; Fig. 33).

In contrast to the Arabian amphibian species, the late Miocene appears to be a time for the rapid diversification of both arid and temperate Arabian gekkonids. A particularly interesting pattern is displayed by the genus *Hemidactylus*, in which sharp species breaks occur over short geographic distances throughout the AHA, coastal South Arabia, and the NOH (Šmíd *et al*., 2013). In addition, several *Hemidactylus* species exhibit high levels of intraspecific variation, particularly in the AHA. The pattern of high species turnover, coupled with signatures of intraspecific divergences, is suggestive of Plio-Pleistocene climatic fluctuations driving diversification at multiple levels in *Hemidactylus* geckos.

Diversification or persistence in the Arabian Hotspot Area?

In general, the effects of overall climate change and Plio-Pleistocene climatic oscillations are expected to affect montane and lowland systems in different ways, with montane regions typically experiencing long-term habitat stability and lowland regions displaying differential habitat expansions and contractions (Voelker *et al*., 2010; Voelker *et al*., 2012). In the montane regions of East Africa, the combination of stable refugial habitat with interspersed periods of fragmentation has resulted in both the long-term persistence of localized lineages and the diversification of cyclically fragmented lineages (Fjeldså & Lovett, 1997; Bowie *et al.*, 2004; Jetz *et al.*, 2004; Bowie *et al*., 2006; Fjeldså *et al.*, 2007; Fjeldså & Bowie, 2008; Voelker *et al*., 2010). Similar patterns could be predicted for the montane regions of the Arabian Peninsula, and in particular the Arabian Hotspot Area, which exhibits localized endemism across vertebrate groups.

Our results demonstrate that Arabian amphibians are characterized by a high degree of lineage persistence, which has lead to the accumulation of several ancient species occurring in the AHA. Although diverse in biogeographic origin, these amphibian species all diverged in the Miocene, and consist of ancient African relicts as well as Asian and Palearctic dispersers. Based on current and paleoclimate distribution models, we identified historically stable refugial regions for montane, lowland, and wide-ranging amphibian species (Fig. 32). These refugial areas occur within the AHA but are largely non-overlapping, demonstrating the importance of this region not only for the persistence of montane species, but for mid-elevation and lowland species as well.

The patterns of lineage persistence and endemism in Afromontane systems are sometimes interpreted as resulting from locally decreased extinction rates, rather than increased diversification rates (Fjeldså & Bowie, 2008). Although an explicit test of this hypothesis is beyond the scope of our review, the role of extinction in shaping diversification patterns on the Arabian Peninsula must not be overlooked. The extinct

aquatic pipid frog, *Xenopus arabiensis*, was recently discovered in the Yemeni highlands, and dates to the late Oligocene (30.2–26.5 Ma) (Henrici & Báez, 2001). During this time, the region was experiencing uplift and volcanism associated with the formation of the Red Sea (Bosworth *et al*., 2005), which may have contributed to the extinction of this African-derived species. The Arabian amphibian community therefore may have been more diverse historically, making it difficult to assess levels of amphibian persistence in the Arabian Peninsula. The number of examples of long-term persistence, represented by extant species, may be far outnumbered by extinction events or high turnover rates. In either scenario the AHA has acted primarily as a center for lineage persistence for amphibians, rather than promoting diversification at the species level.

There is strong evidence that several lineages in the gekkonid genus *Hemidactylus* diversified within the Arabian Hotspot Area during the late Miocene and Pliocene (Šmíd *et al*., 2013), in contrast to the patterns exhibited by Arabian amphibians. These species of *Hemidactylus* are estimated to be younger in age than the Arabian amphibians and typically display restricted ranges occurring at different elevational zones, suggesting diversification has occurred in both montane and lowland forms (Šmíd *et al*., 2013). Several species also display a high degree of intraspecific genetic divergence, often occurring on a very small geographic scale within the AHA, and it is likely there are undescribed species occurring in this region. Given these patterns, the AHA appears to have acted as a center of diversification for several lineages of *Hemidactylus* geckos.

The contrasting patterns of lineage persistence and diversification across different vertebrate groups in the AHA illustrate the importance of this region in sculpting patterns of biodiversity and endemism. The AHA is in many ways similar to the montane systems of East Africa, which are also characterized by habitat stability and the long-term persistence of localized lineages and the diversification of cyclically fragmented lineages (Fjeldså & Lovett, 1997; Bowie *et al.*, 2004; Jetz *et al.*, 2004; Bowie *et al*., 2006; Fjeldså *et al.*, 2007; Fjeldså & Bowie, 2008; Blackburn & Measey, 2009; Lawson, 2010; Voelker *et al*., 2010). However, in contrast to East African systems, we find evidence for stable habitat occurring in both montane areas and lowland coastal areas of the AHA. Despite these initial findings, the overall patterns of persistence and diversification in both lowland and montane regions of the AHA remain largely unexplored, and there is a clear need for continued study in this biodiversity hotspot.

Conclusions

The varying levels of connectivity between the Arabian Peninsula and surrounding continents throughout geologic time allows explicit temporal predictions to be made about diversification and speciation events for Arabian taxa derived from African, Asian, and Palearctic lineages (Fig. 29). Among the time-calibrated phylogenetic studies conducted on Arabian taxa, there is a striking concordance of divergence times for taxa with similar biogeographic histories, providing support for many of the biogeographic hypotheses derived from geologic events (Fig. 33).

Biogeographic patterns have been investigated in several groups that display disjunct Afro-Arabian distributions, revealing two distinct time periods in which African-derived lineages accumulated on the Arabian Peninsula. The completion of the formation of the Red Sea in the Early Miocene (by 23 Ma) appears to have driven vicariant speciation in lineages with formerly continuous Afro-Arabian distributions. These true

African relicts include examples of toads and sawscale vipers (Pook *et al*., 2009), and are the most ancient lineages currently detected on the Arabian Peninsula. The establishment of a temporary southern Afro-Arabian landbridge 10-5.3 Ma facilitated dispersals of many additional African taxa to the Arabian Peninsula, while simultaneously allowing Arabian taxa to invade the Horn of Africa (Amer & Kumazawa, 2005; Pook *et al*., 2009; Wong *et al*., 2010; Metallinou *et al*., 2012; Portik & Papenfuss, 2012; Šmíd *et al*., 2013). The dispersal of African taxa to the southwestern Arabian Peninsula appears to have continued in the form of overwater dispersal for several vagile lineages, most notably in large mammals (Zinner *et al*., 2009; Fernandes, 2011).

Divergence timings of terrestrial vertebrate exchanges between Arabia and Asia have revealed the importance of land bridge connections in facilitating dispersals. The collision of the Arabian and Anatolian plates approximately 18–16 Ma places an upper limit on dispersals between Asia and the Arabian Peninsula, and is supported by amphibian and gekkonid lineages and the plant genus *Isodon* (Šmíd *et al*., 2013; Yu *et al*., 2014). A second dispersal opportunity resulted from the drying of the Persian Gulf in the Pleistocene and is a pattern currently only supported by South Asian vipers (Pook *et al*., 2009), however additional groups exhibiting South Asian-Arabian disjunct distributions remain to be investigated.

The Palearctic-derived taxa occurring on the Arabian Peninsula require additional study, but evidence from montane amphibian lineages suggests dispersal along with the mountain corridors extending along the Red Sea coast to the Jordan Rift Valley occurred during the Late Miocene, after the establishment of the permanent Arabian-Eurasian land connection. The disjunct Levantine-Arabian distribution of several amphibian species is an intriguing biogeographic pattern, and prevalence of this distribution among other vertebrate taxa remains to be explored.

The aridification of the Arabian Peninsula in the Late Miocene facilitated the expansion of xeric habitats and the fragmentation of mesic habitats, which may have driven diversification in both arid and temperate climate reptile groups (Metallinou *et al*., 2012; Šmíd *et al*., 2013). Evidence based on amphibians suggests large populations became isolated during the Pliocene, resulting in the disjunct Levantine-AHA pattern and Northern Oman Highlands-AHA pattern. The Arabian Hotspot Area is similar to East African montane systems in facilitating both long-term lineage persistence and diversification of lineages, although the AHA appears to have afforded protection to both montane and lowland species. The contrasting patterns of persistence of Miocene-age amphibian species and the rapid diversification of Arabian geckos in the AHA demonstrates the clear need for continued study in this region. In addition, the coastal lowland and montane regions of the AHA are appealing systems for investigating the demographic effects of Plio-Pleistocene climate change, a topic that has received very little attention for the Arabian Peninsula. Phylogeographic studies of coastal lowland and montane Arabian species can test the effects of Pleistocene habitat fragmentation and expansion by examining range-wide patterns of population structuring and genetic diversity along with relevant paleodistribution models, and may ultimately reveal the presence of additional species in this biodiversity hotspot.

Acknowledgments
Anticipated coauthorship: Theodore J. Papenfuss, Guin O.U. Wogan, Jamie Oaks.

T.J.P. thanks Ali Al Kiyumi, Ministry of Environment and Climate Affairs issued Oman for collecting and export permits, the American Institute for Yemeni Studies for facilitating fieldwork in Yemen, Abdurahman A. Osman, Director of the Puntland Development Research Center for arranging fieldwork in Puntland State, Somalia, Suleiman Ahmed Gulaid, President of Amoud University, Borama, Somaliland for facilitating fieldwork in Somaliland. Partial funding for T.J.P's fieldwork was provided by a grant from the George Lindsay Field Research Fund of the California Academy of Sciences. Molecular work was funded in part by the Museum of Vertebrate Zoology of the University of California, Berkeley. D.M.P. was funded by teaching assistantships awarded by the Integrative Biology Department of the University of California, Berkeley. G.O.U.W was supported by NSF DEB 1120356. We thank J.A. McGuire, D.C. Blackburn, R.C.K. Bowie, S.M. Rovito, S. Hykin, P. Skipwith, and R. von May for comments on earlier drafts of this manuscript, and H.C. Liedtke for comparisons of unpublished bufonid molecular data.

References

Akın, C., C. Can Bilgin, P. Beerli, R. Westaway, T. Ohst, S.N. Litvinchuk, T. Uzzell, M. Bilgin, H. Hotz, G.-D.Guex, and J. Plötner. 2010. Phylogeographic patterns of genetic diversity in eastern Mediterranean water frogs were determined by geological processes and climate change in the Late Cenozoic. Journal of Biogeography 37:1365–2699.

Aljanabi, S. and I. Martinez. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Research 25:4692–4693.

Amer, S.A.M. and Y. Kumazawa. 2005. Mitochondrial DNA sequences of the Afro-Arabian spiny-tailed lizards (genus *Uromastyx*; family Agamidae): phylogenetic analyses and evolution of gene arrangements. Biological Journal of the Linnaean Society 85:247–260.

AmphibiaWeb: Information on amphibian biology and conservation. 2014. Berkeley, California: AmphibiaWeb. Available: http://amphibiaweb.org.

Arnold, E.N. 1987. Zoogeography of the reptiles and amphibians of Arabia. *Proceedings of the Symposium on the Fauna and Zoogeography of the Middle East* (ed. by F. Krupp, W. Schneider and R. Kinzelbach), pp. 245–256. Hubert & Co., Göttingen.

Balletto, E., M.A. Cherchi, and J. Gasperetti. 1985. Amphibians of the Arabian Peninsula. Fauna of Saudi Arabia, vol. 7. Edited by Büttiker W, Krupp F. Basle: Pro Entomologia, Natural History Museum; 318–392.

Bazinet, A.L., and M.P. Cummings. 2011. Computing the Tree of Life — Leveraging the power of desktop and service grids (IPDPSW). Pp. 1896–1902 in 2011 IEEE International Symposium on Parallel and Distributed Processing Workshops and Phd Forum.

Bazinet, A.L., and M.P. Cummings. 2008. The Lattice Project: a grid research and production environment combining multiple grid computing models. Pp. 2–13 in M.H.W. Weber (ed.), Distributed & Grid Computing - Science Made Transparent for Everyone. Principles, Applications and Supporting Communities. Marburg: Rechenkraft.net.

Blackburn, D.C., and G.J. Measey. 2009. Dispersal to or from an African biodiversity hotspot? Molecular Ecology 18:1904–1915.

Bohannon, R.G. 1986. Tectonic configuration of the western Arabian continental margin, southern Red Sea. Tectonics 5:477–499.

Bossuyt, F., R.M. Brown, D.M. Hillis, D.C. Canatella, and M.C. Milinkovitch. 2006. Phylogeny and biogeography of a cosmopolitan frog radiation: Late Cretaceous diversification resulted in continent-scale endemism in the family Ranidae. Systematic Biology 55:579–594.

Bosworth, W., P. Huchon, and K. McClay. 2005. The Red Sea and Gulf of Aden Basins. Journal of African Earth Sciences 43:334–378.

Bowie, R.C.K., J. Fjeldså, S.J. Hackett, and T.M.Crowe. 2004. Molecular evolution in space and time: mtDNA phylogeography of the Olive Sunbird (*Nectarinia olivacea*/*obscura*) throughout continental Africa. Molecular Phylogenetics and Evolution 33:56–74.

Bowie, R.C.K., J. Fjeldså, S.J. Hackett, J.M. Bates, and T.M. Crowe. 2006. Coalescent models reveal the relative roles of ancestral polymorphism, vicariance, and

dispersal in shaping phylogeographical structure of an African montane forest robin. Molecular Phylogenetics and Evolution 38:171–188.
Braconnot, P., B. Otto-Bliesner, S. Harrison, S. Jousannaume, J.-Y. Peterchmitt, A. Abe-Ouchi, M. Crucifix, E. Driesschaert, T. Fishefet, C.D. Hewitt, M. Kageyama, A. Kitoh, A. Laine, M.-F. Loutre, O. Marti, U. Merkel, G. Ramstein, P. Valdes, S.L. Weber, Y. Yu, and Y. Zhao. 2007. Results of PMIP2 coupled simulations of the Mid-Holocene and Last Glacial Maximum-Part 1: experiments and large-scale features. Climate of the Past 3:261–277.
deMenocal, P.B. 1995. Plio-Pleistocene African Climate. Science 270:53–59.
deMenocal, P.B. 2004. African climate change and faunal evolution during the Pliocene-Pleistocene. Earth and Planetary Science Letters 220:3–24.
Drummond, A.J., M.A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29:1969–1973.
Fernandes, C.A., E.J. Rohling, and M. Siddall. 2006. Absence of post-Miocene Red Sea land bridges: biogeographic implications. Journal of Biogeography 33:961–966.
Fisher, M., S.A. Ghazanfar, S.A. Chaudhary, P.J. Seddon, E.F. Robertson, S. Omar, J.A. Abbas, and B. Böer. 1998. Diversity and conservation. Pp. 265–302 in S.A. Ghazanfar, F.M. Dordrecht (eds.), Vegetation of the Arabian Peninsula. Kluwer Academic Publishers.
Fisher, M., and D.A. Membery. 1998. Climate. Pp. 5–38 in S.A. Ghazanfar, F.M. Dordrecht (eds.), Vegetation of the Arabian Peninsula. Kluwer Academic Publishers.
Fjeldså, J. and J.C. Lovett. 1997. Geographical patterns of old and young species in African forest biota: the significance of specific montane areas as evolutionary centres. Biodiversity and Conservation 6:325–346.
Fjeldså, J., U.S. Johansson, L.G.S. Lokugalappati, and R.C.K. Bowie. 2007. Diversification of African greenbuls in space and time: linking ecological and historical processes. Journal of Ornithology 148 S2:S359–S367.
Fjeldså, J., and R.C.K. Bowie. 2008. New perspectives on the origin and diversification of Africa's forest avifauna. African Journal of Ecology 46:235–247.
Frost, D.R. 2014. Amphibian Species of the World: an Online Reference. Version 6.0 (January 2014). Electronic Database accessible at http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York, USA.
Frost, D.R., T. Grant, J. Faivovich, R.H. Bain, A. Haas, C.F.B. Haddad, A. Channing, M. Wilkinson, S.C. Donnellan, C.J. Raxworthy, J.A. Campbell, B.L. Blotto, P. Moler, R.C. Drewes, R.A. Nussbaum, J.D. Lynch, D.M. Green, and W.C. Wheeler. 2006. The amphibian tree of life. Bulletin of the American Museum of Natural History 297:1–370.
Gasperetti, J. 1988. Snakes of Arabia. Pp. 169–450 in W. Büttiker, K.F. Basle (eds.), Fauna of Saudi Arabia, Vol. 9. Pro Entomologia, Natural History Museum.
Glennie, K.W., and A.K. Singhvi. 2002. Event stratigraphy, paleoenvironment and chronology of SE Arabian deserts. Quaternary Science Reviews 21:853–869.

Guba, I., and K. Glennie. 1998. Geology and geomorphology. Pp. 39–62 in S.A. Ghazanfar, F.M. Dordrecht (eds.), Vegetation of the Arabian Peninsula. Kluwer Academic Publishers.

Gvoždík, V., J. Moravec, C. Klüsch, and P. Kotlík. 2010. Phylogeography of the Middle Eastern tree frogs (*Hyla*, Hylidae, Amphibia) as inferred from nuclear and mitochondrial DNA variation, with a description of a new species. Molecular Phylogenetics and Evolution 55:1146–1166.

Henrici, A.C. and A.M. Báez. 2001. First occurrence of *Xenopus* (Anura: Pipidae) on the Arabian Peninsula: A new species from the Upper Oligocene of Yemen. Journal of Paleontology 75:870–882.

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

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

Holman, J.A. 2003. Pleistocene amphibians and reptiles in Britain and Europe. Oxford University Press, Bloomington and Indianapolis.

Huelsenbeck, J.P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Jetz, W., C. Rahbek, and R.K. Colwell. 2004. The coincidence of rarity and richness and the potential signature of history in centres of endemism. Ecology Letters 7:1180–1191.

Joger, U. 1987. An interpretation of reptile zoogeography in Arabia, with special reference to Arabian herpetofaunal relations with Africa. Pp. 257–271 in F. Krupp, W. Schneider, K. Göttingen (eds.), Proceedings of the Symposium on the Fauna and Zoogeography of the Middle East. R. Hubert & Co.

Jung, S.J.A., G.R. Davies, G.M. Ganssen, and D. Kroon. 2004. Synchronous Holocene sea surface temperature and rainfall variations in the Asian monsoon system. Quaternary Science Reviews 23:2207–2218.

Kosuch, J., M. Vences, A. Dubois, A. Ohler, and W. Böhme. 2001. Mitochondrial DNA evidence for an Oriental origin of Tiger Frogs, genus *Hoplobatrachus*. Molecular Phylogenetics and Evolution 21:398–407.

Krupp, F. 1983. Freshwater fishes of Saudi Arabia and adjacent regions of the Arabian Peninsula. Pp. 568–636 in H.C.W. Wittmer and W. Büttiker (eds.), Fauna of Saudi Arabia, Vol. 5. Pro Entomologia, Natural History Museum, Basle.

Kürschner, H. 1998. Biogeography and introduction to vegetation. Pp. 63–98 in S.A. Ghazanfar, F.M. Dordrecht (eds.), Vegetation of the Arabian Peninsula. Kluwer Academic Publishers.

Lambeck, K. 1996. Shoreline reconstructions for the Persian Gulf since the last glacial maximum. Earth and Planetary Science Letters 142:43–57.

Lanfear, R., B. Calcott, S.Y.W. Ho, and S. Guindon. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution 29:1695–1701.

Lawson, L.P. 2010. The discordance of diversification: evolution in the tropical-montane frogs of the Eastern Arc Mountains of Tanzania. Molecular Ecology 19:4046–4060.

Lézine, A.-M., J.-F. Saliège, C. Robert, F. Wertz, and M.-L. Inizan. 1998. Holocene lakes from Ramlat as-Sab'atayn (Yemen) illustrate the impact of monsoon activity in Southern Arabia. Quaternary Research 50:290–299.

Lückge, A., H. Doose-Rolinski, A.A. Khan, H. Scuhlz, and U. von Rad. 2001. Monsoonal variability in the northeastern Arabian Sea during the past 5000 years: geochemical evidence from laminated sediments. Paleogeography, Paleoclimatology, Paleoecology 167:273–286.

Mallon, D.P. 2011. Global hotspots in the Arabian Peninsula. Zoology in the Middle East 54: Supplementum 3, 13–20.

Meister, J., M.A. Hubaishan, N. Kilian, and C. Oberprieler. 2005. Chloroplast DNA variation in the shrub *Justicia areysiana* (Acanthaceae) endemic to the monsoon affected coastal mountains of the southern Arabian Peninsula. Botanical Journal of the Linnean Society 148:437–444.

Meister, J., M.A. Hubaishan, N. Kilian, and C. Oberprieler. 2006. Temporal and spatial diversification of the shrub *Justicia areysiana* Deflers (Acanthaceae) endemic to the monsoon affected coastal mountains of the southern Arabian Peninsula. Plant Systematics and Evolution 262:153–171.

Metallinou, M., N.E. Arnold, P.A. Crochet, P. Geniez, J.C. Brito, P. Lymberakis, S.B. El Din, R. Sindaco, M. Robinson, and S. Carranza. 2012. Conquering the Sahara and Arabian deserts: Systematics and biogeography of *Stenodactylus* geckos (Reptilia: Gekkonidae). BMC Evolutionary Biology 12:258.

Miller, A.G., and T.A. Cope. 1996. Flora of the Arabian Peninsula and Socotra, vol. 1. Edinburgh: Edinburgh University Press.

Mittermeier, R.A., N. Myers, P.R. Gil and C.G. Mittermeier. 1999. Hotspots: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions. Monterrey, Mexico: Cemex, Conservation International and Agrupacion Sierra Madre.

Moriarty Lemmon, E., A.R. Lemmon, and D.C. Cannatella. 2007. Geological and climatic forces driving speciation in the continentally distributed trilling chorus frogs (*Pseudacris*). Evolution, 61:2086–2103.

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

Oberprieler, C., J. Meister, C. Schneider, and N. Kilian. 2009. Genetic structure of *Anogeissus dhofarica* (Combretacea) populations endemic to the monsoonal fog oases of the southern Arabian Peninsula. Biological Journal of the Linnaean Society 97:40–51.

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

Pook, C.E., U. Joger, N. Stümpel, and W. Wüster. 2009. When continents collide: phylogeny, historical biogeography and systematics of the medically important viper genus *Echis* (Squamata: Serpentes: Viperidae). Molecular Phylogenetics and Evolution 53:792– 807.

Portik, D.M., and T.J. Papenfuss. 2012. Monitors cross the Red Sea: the biogeographic history of *Varanus yemenensis*. Molecular Phylogenetics and Evolution 62:561–565.
Portik, D.M. and T.J. Papenfuss. 2015. Historical biogeography resolves the origins of endemic Arabian toad lineages (Anura: Bufonidae): Evidence for ancient vicariance and dispersal events with the Horn of Africa and South Asia. BMC Evolutionary Biology 15:152.
Pramuk, J.B., T. Robertson, J.W. Sites, and B.P. Noonan. 2008. Around the world in 10 million years: biogeography of the nearly cosmopolitan true toads (Anura: Bufonidae). Global Ecology and Biogeography 17:72–83.
Rambaut, A., and A.J. Drummond. 2009. Tracer v1.5.0 and higher [http://beast.bio.ed.ac.uk]
Roelants, K., J. Jiang, and F. Bossuyt. 2004. Endemic ranid (Amphibia: Anura) genera in southern mountain ranges of the Indian subcontinent represent ancient frog lineages: evidence from molecular data. Molecular Phylogenetics and Evolution, 31:730–740.
Rögl, F. 1998. Paleogeographic considerations for Meditteranean and Paratethys seaways (Oligocene to Miocene). Annalen des Naturhistorischen Museums in Wien 99A:279–310.
Rögl, F. 1999. Mediterranean and Paratethys. Facts and hypotheses of an Oligocene to Miocene paleogeography (short overview). Geologica Carpathia 50:339–349.
Rohling, E.J., M. Fenton, F.J. Jorissen, P. Bertrand, G. Ganssen, and J.P. Caulet. 1998. Magnitudes of sea-level lowstands of the past 500,000 years. Nature 394:162–165.
Ronquist, F., and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.
Sanchíz, B. 1998. Salientia. Pp. 1–275 in W.P. Munich (ed.), Handbuch der paläoherpetologie Volume Part 4. Verlag Dr. Friedrich Pfeil.
Schätti, B., and A. Desvoignes. 1999. The herpetofauna of southern Yemen and the Socotran Archipelago. Instrumenta Biodiversitatis 4:1–178.
Sirocko, F. 1996. The evolution of the monsoon climate over the Arabian Sea during the last 24,000 years. Paleoecology of Africa 24:53–69.
Šmíd, J., S. Carranza, L. Kratochvíl, V. Gvoždík, A.K. Nasher, and J. Moravec. 2013. Out of Arabia: A complex biogeographic history of multiple vicariance and dispersal events in the Gecko genus *Hemidactylus* (Reptilia: Gekkonidae). PLoS ONE 8(5):e64018.
Smith, S.A., A.N.M. de Oca, T.W. Reeder, and J.J. Wiens. 2007. A phylogenetic perspective on elevational species richness patterns in Middle American treefrogs: why so few species in lowland tropical rainforests? Evolution 61:1188–1207.
Stöck, M., C. Moritz, M. Hickerson, D. Frynta, T. Dujsebayeva, V. Eremchenko, R.J. Macey, T.J. Papenfuss, and D.B. Wake. 2006. Evolution of mitochondrial relationships and biogeography of Palearctic green toads (*Bufo viridis* subgroup) with insights in their genomic plasticity. Molecular Phylogenetics and Evolution 41:663–689.

Stuart, S.N., J.S. Chanson, N.A. Cox, B.E. Young, A.S.L. Rodrigues, D.L. Fischman, and R.W. Waller. 2004. Status and trends of amphibian declines and extinctions worldwide. Science 306:1783–1786.

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

Uchupi, E., S.A. Swift, and D.A. Ross. 1999. Late Quaternary stratigraphy, Paleoclimate and neotectonism of the Persian (Arabian) Gulf region. Marine Geology 160:1–23.

Yu, X.-Q., M. Maki, B.T. Drew, A.J. Paton, H.-W. Li, J.-L. Zhao, J.G. Conran, and J. Li. 2014. Phylogeny and historical biogeography of *Isodon* (Lamiaceae): Rapid radiation in south-west China and Miocene overland dispersal into Africa. Molecular Phylogenetics and Evolution 77:183–194.

Van Bocxlaer, I., S.D. Biju, S.P. Loader, and F. Bossuyt. 2009. Toad radiation reveals into-India dispersal as a source of endemism in the Western Ghats-Sri Lanka biodiversity hotspot. BMC Evolutionary Biology, 9, 131.

Van Bocxlaer, I., S.P. Loader, K. Roelants, S.D. Biju, M. Menegon, and F. Bossuyt. 2010. Gradual adaptation toward a range-expansion phenotype initiated the global radiation of toads. Science 327:679–682.

Vesey-Fitzgerald, D.F. 1955. Vegetation of the Red Sea coast south of Jedda, Saudi Arabia. Journal of Ecology 43:477–489.

Vesey-Fitzgerald, D.F. 1957. The vegetation of central and eastern Arabia. Journal of Ecology 45:779–798.

Voelker, G., R.K. Outlaw, and R.C.K. Bowie. 2010. Pliocene forest dynamics as a primary driver of African bird speciation. Global Ecology and Biogeography 19:111–121.

Voelker, G., R.C.K. Bowie, B. Wilson, and C. Anderson. 2012. Phylogenetic relationships and speciation patterns in an African savanna dwelling bird genus (Myrmecocichla). Biological Journal of the Linnaean Society 106:180–190.

Wiens, J.J., J. Sukumaran, A.R. Pyron, and R.M. Brown. 2009. Evolutionary and biogeographic origins of high tropical diversity in old world frogs. Evolution 63:1217–1231.

Wong, R.A., J.J. Fong, and T.J. Papenfuss. 2010. Phylogeography of the African helmeted terrapin, *Pelomedusa subrufa*: genetic structure, dispersal, and human introduction. Proceedings of the California Academy of Sciences 61:575–585.

Zinner, D., L.F. Groeneveld, C. Keller, and C. Roos. 2009. Mitochondrial phylogeography of baboons (*Papio* spp.) – indication for introgressive hybridization? BMC Evolutionary Biology 9:1–15.

Zwickl, D.J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD thesis, The University of Texas at Austin, Austin.

Figure 28. The Arabian Peninsula. The Arabian Hotspot Area consists of the combined area of the disjunct extensions of the Horn of Africa Hotspot (medium grey) and the Eastern Afromontane Hotspot (dark grey). Northern Oman Highlands are identified by dotted outline.

Figure 28

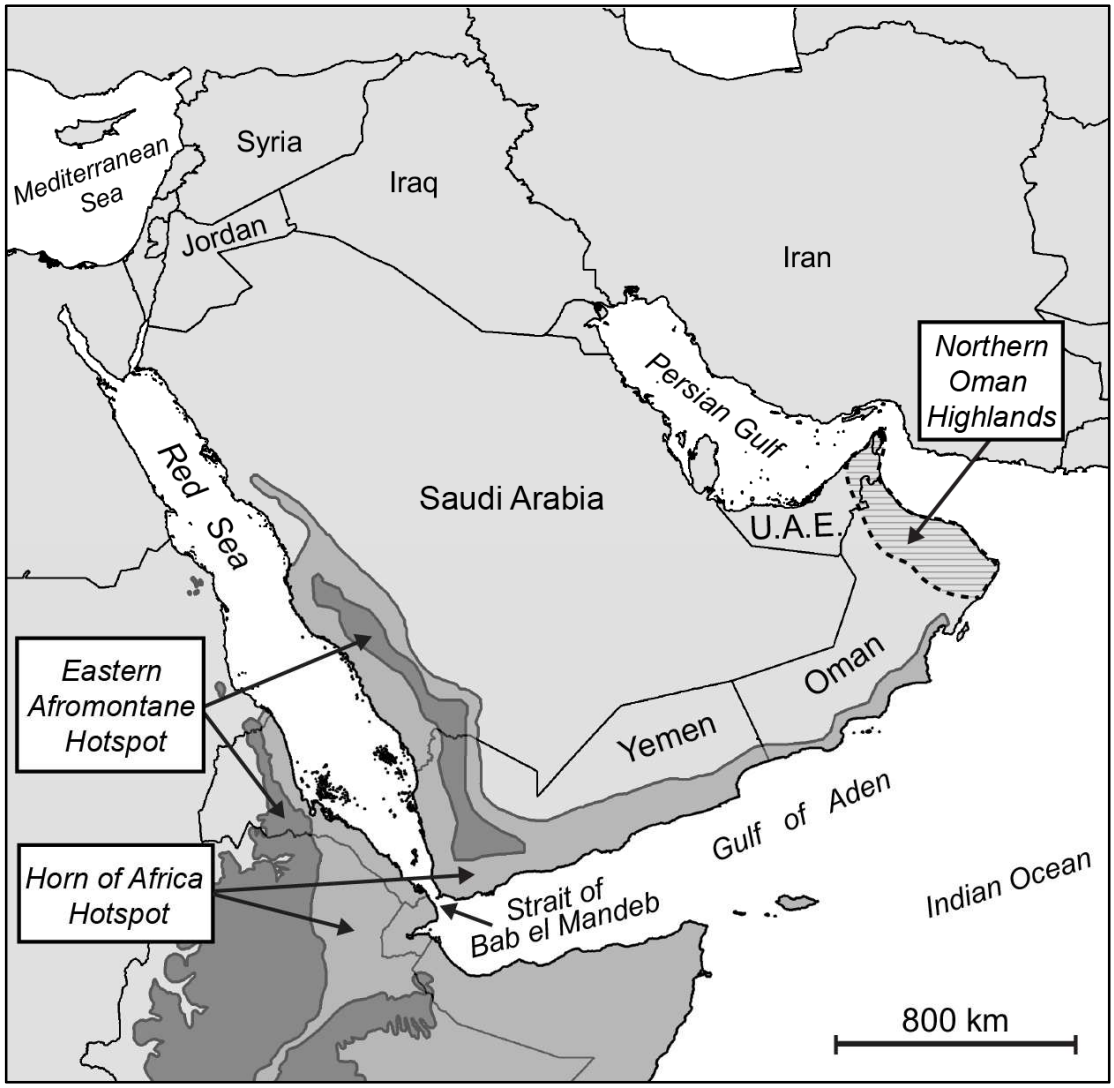

Figure 29. Summary of major geologic events driving vicariance and dispersal events: (a) African-Arabian Plate prior to major tectonic activity, 30 Ma, (b) formation of southern Red Sea began at 27 Ma, with extension northwards and separation of Arabian Plate by 23 Ma, (c) collision with the Anatolian plate approximately 18–16 Ma established the temporary *Gomphotherium* bridge, (d) collision with Eurasia established a permanent land connection by 15 Ma, (e) the closing of the Strait of Bab el Mandeb established a southern land bridge between Horn of Africa and Arabia, 10 Ma, and (f) reopening of Strait of Bab el Mandeb by 5 Ma. Hashed fill of Red Sea in (d) and (e) indicates increased salinity levels resulting from reduction in marine flow. Geologic reconstructions are based on Rögl (1998, 1999) and Bosworth *et al*. (2005).

Figure 29

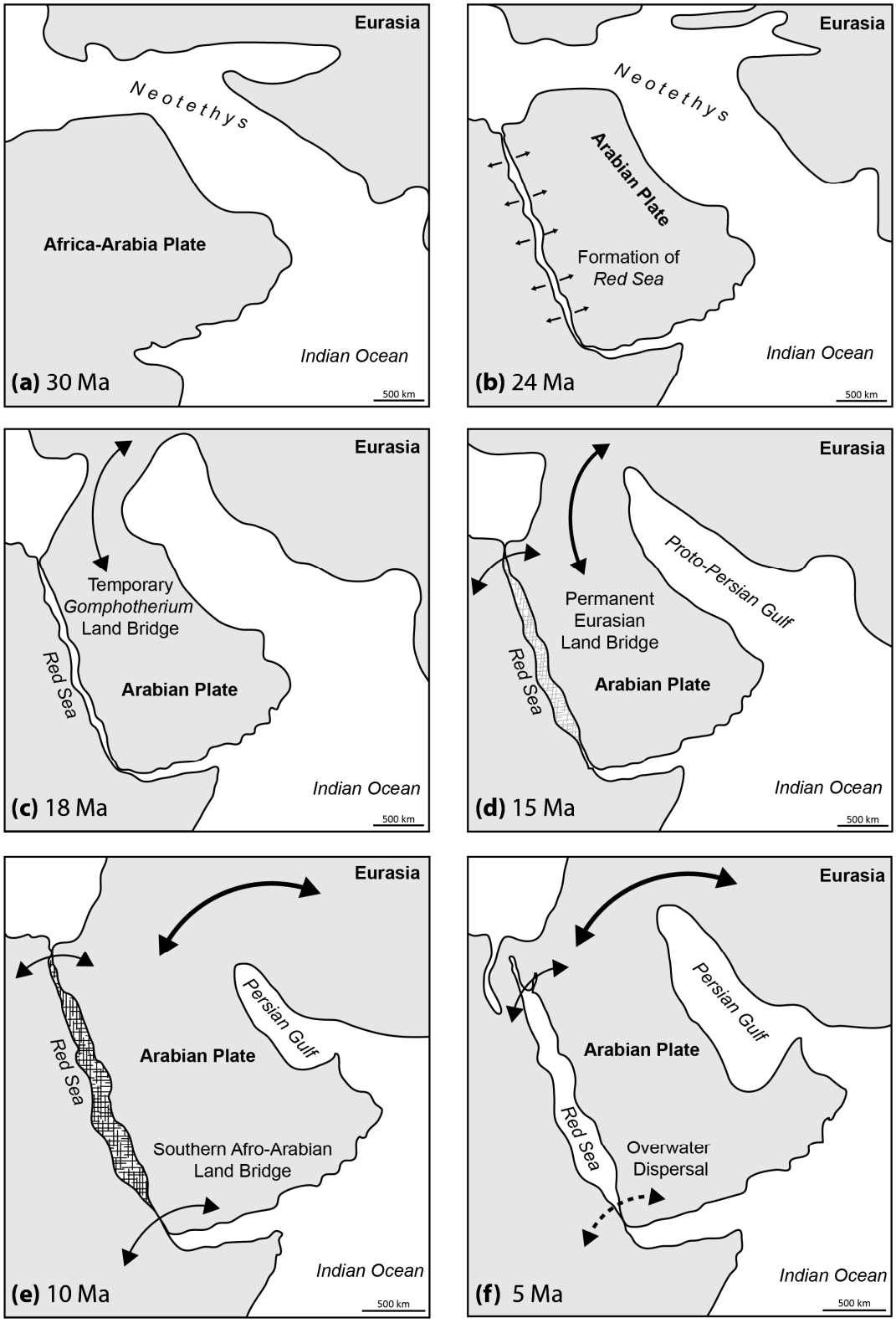
Eurasia
Neotethys
Africa-Arabia Plate
Indian Ocean
(a) 30 Ma
500 km
Eurasia
Neotethys
Arabian Plate
Formation of
Red Sea
Indian Ocean
(b) 24 Ma
500 km
Eurasia
Temporary
Gomphotherium
Land Bridge
Red Sea
Arabian Plate
Indian Ocean
(c) 18 Ma
500 km
Eurasia
Proto-Persian Gulf
Permanent
Eurasian
Land Bridge
Red Sea
Arabian Plate
Indian Ocean
(d) 15 Ma
500 km
Eurasia
Persian Gulf
Red Sea
Arabian Plate
Southern Afro-Arabian
Land Bridge
Indian Ocean
(e) 10 Ma
500 km
Eurasia
Persian Gulf
Arabian Plate
Red Sea
Overwater
Dispersal
Indian Ocean
(f) 5 Ma
500 km

Figure 30. Elevational ranges for species of Arabian anurans.

Figure 30

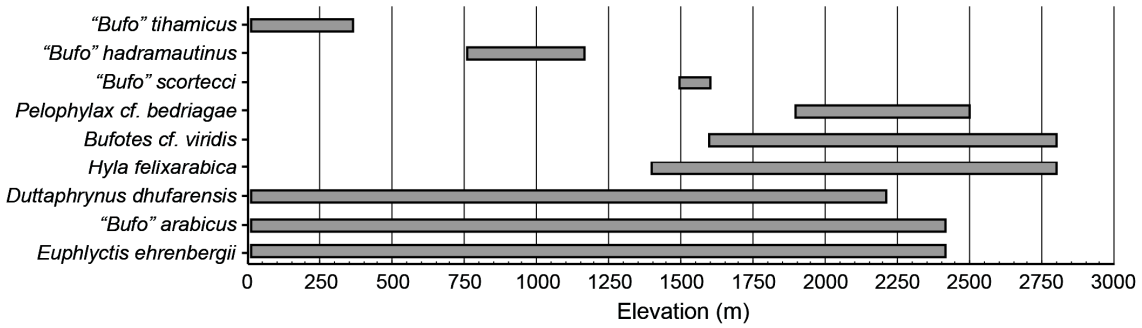

Figure 31. Bayesian maximum clade credibility chronogram of the (a) dicroglossids, and (b) hylids, inferred in BEAST with mean divergence times, 95% HPD of dates, with timing of relevant geologic events and ranges of Arabian taxa and sister groups. Filled circles on nodes represent high support (PPB > 0.95 and ML bootstrap > 70%), open circles represent support values below either threshold.

Figure 31

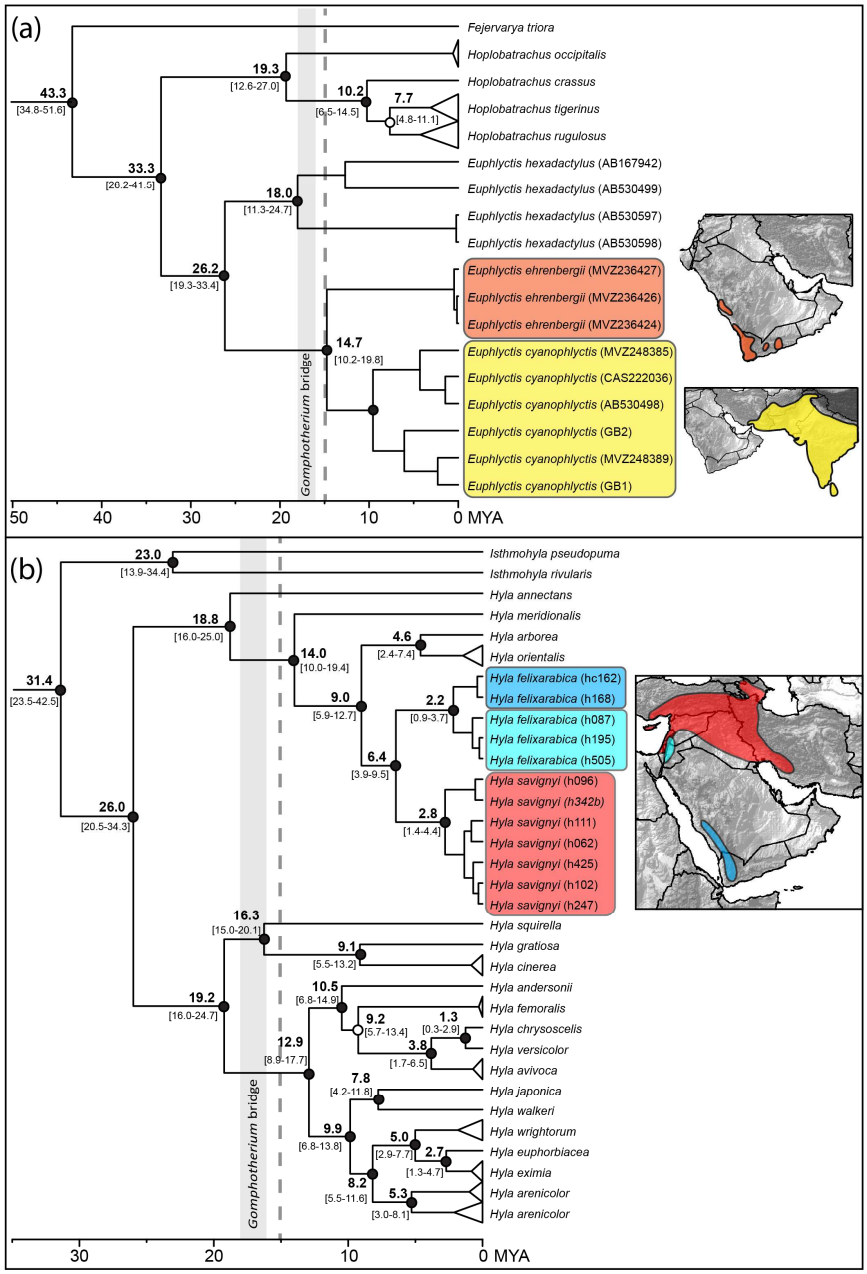

Figure 32. Stability surfaces representing regions of persistent suitable habitat across LGM and current climate regimes that are in common among (a) highland amphibian species and (b) low to mid-elevation amphibian species.

Figure 32

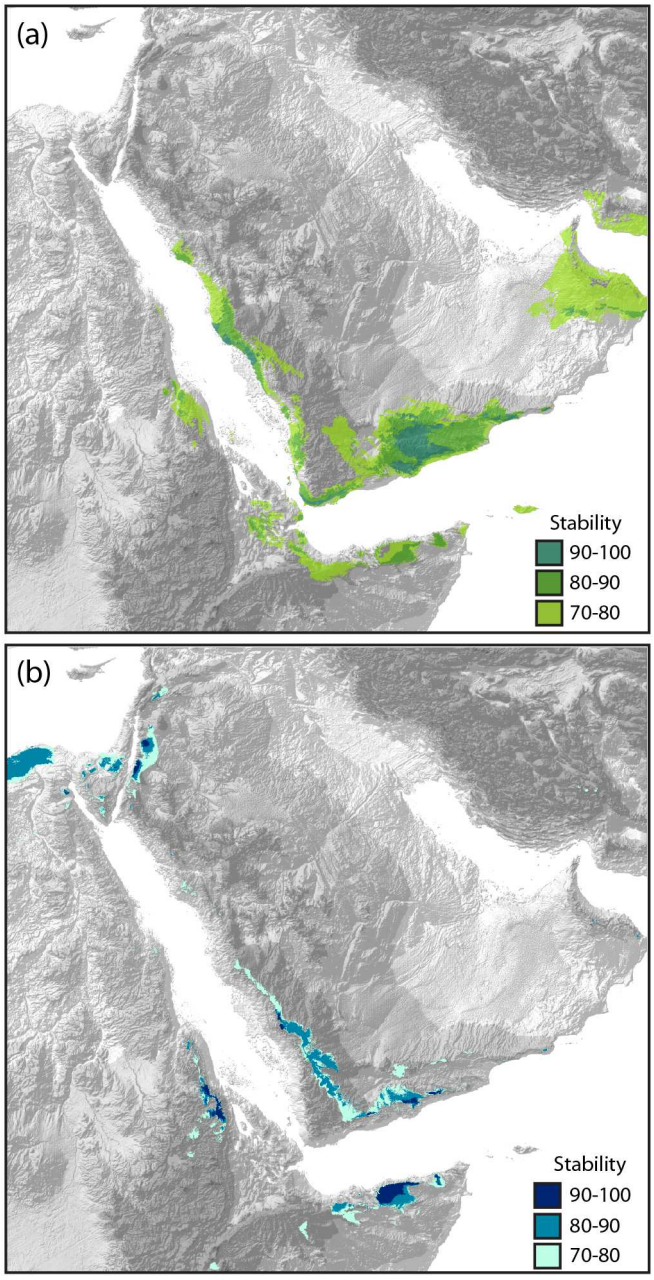

Figure 33. Summary of divergence date estimates from phylogenetic studies of Arabian amphibians, reptiles, mammals, and plants. Data are divided into two categories: continental exchanges (below), and events occurring on the Arabian Peninsula (top).

Figure 33

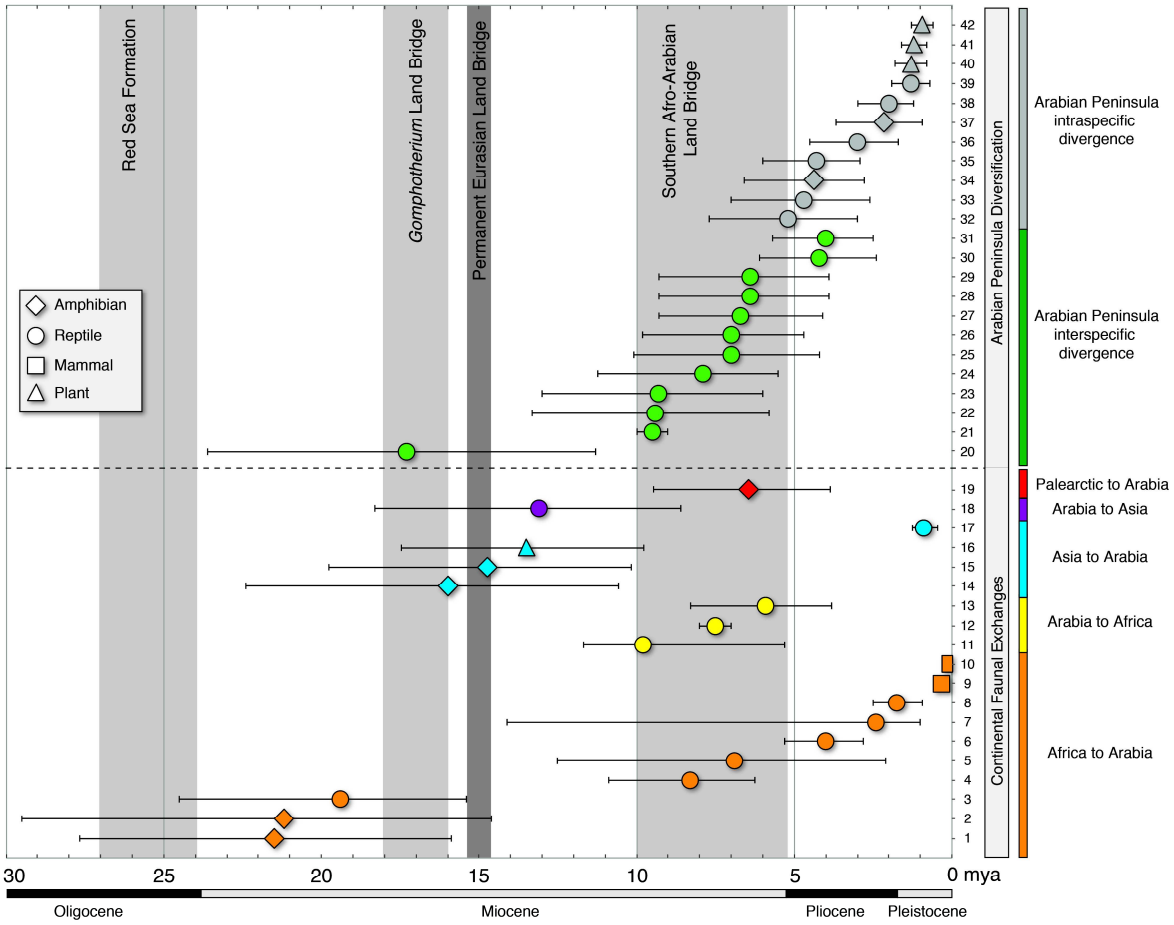

Table 20. Genbank information for Dicroglossid sampling.

Table 20

Species	16S	Rhodopsin	Tyrosinase
Euphylctis cyanophlyctis	AF249053	AF249111	AF249174
Euphylctis cyanophlyctis	AB488901	AB489038	AB489017
Euphylctis cyanophlyctis	AB530498		
Euphylctis hexadactylus	AB530499		
Euphylctis hexadactylus	AB530598		
Euphylctis hexadactylus	AB530597		
Euphylctis hexadactylus	AB167942		
Fejervarya triora	AB488883	AB489022	AB489003
Hoplobatrachus crassus	AF249109	AF249172	AF249044
Hoplobatrachus occipitalis	EU979934	EU980025	EU979845
Hoplobatrachus occipitalis	EU979935	EU980026	EU979846
Hoplobatrachus rugulosus	AY322289	AY322221	AY322360
Hoplobatrachus rugulosus	DQ458257	DQ458272	DQ458250
Hoplobatrachus rugulosus	DQ458258	DQ458273	DQ458251
Hoplobatrachus rugulosus	EU979933	EU980024	EU979844
Hoplobatrachus tigerinus	AB489039	AB277358	AB488902
Ingerana tenasserimensis	AY322302	AY322236	AY322344
Limnonectes finchi	AY322295	AY322230	AY322355
Nanorana parkeri	AY322283	AY322219	AY322350
Nanorana pleskei	AY322282	AY322235	AY322339
Occidozyga laevis	AY322300	AY322227	AY322342
Paa yunnanensis	AY322298	AY322229	AY322361
Quasipaa boulengeri	AY322280	AY322240	AY322349

Table 21. Genbank information for non-Arabian Hylid sampling.

Table 21

Species	Voucher	16S	12S	Rhodopsin	Tyrosinase
Hyla andersonii		AY843598	AY843598	AY844572	AY844044
Hyla annectans		AY843600	AY843600	AY844574	AY844045
Hyla arborea		AY843601	AY843601	AY844575	AY844046
Hyla arenicolor		AY843603		AY364401	AY844048
Hyla arenicolor	KEK HAR11	GU989068	GU989068	GU944698	
Hyla arenicolor	KEK HAR1	GU989087	GU989087	GU944717	
Hyla arenicolor	KEK HAR3468	GU989080	GU989080	GU944710	
Hyla arenicolor	KEK HAR49	GU989065	GU989065	GU944695	
Hyla arenicolor	KEK HAR50	GU989084	GU989084	GU944714	
Hyla arenicolor	KEK HAR51	GU989075	GU989075	GU944705	
Hyla arenicolor	KEK HAR52	GU989067	GU989067	GU944697	
Hyla arenicolor	KEK HAR53	GU989086	GU989086	GU944716	
Hyla arenicolor	KEK HAR63	GU989091	GU989091	GU944721	
Hyla arenicolor	KEK HAR65	GU989092	GU989092	GU944722	
Hyla arenicolor	KEK HAR8	GU989077	GU989077	GU944707	
Hyla arenicolor	KEK HARA20	GU989072	GU989072	GU944702	
Hyla arenicolor	KEK HARB28	GU989070	GU989070	GU944700	
Hyla arenicolor	KEK HARB33	GU989071	GU989071	GU944701	
Hyla arenicolor	KEK HARB43	GU989066	GU989066	GU944696	
Hyla arenicolor	KEK HARB9a	GU989076	GU989076	GU944706	
Hyla avivoca		AY843605	AY843605	AY844581	AY844051
Hyla avivoca	KEK HAV	GU989090	GU989090	GU944720	
Hyla chrysoscelis		GU989082	GU989082	GU944712	
Hyla cinerea		AY330892	AY680271	AY844597	AY844063
Hyla cinerea		AY819498	AY819366	AY323749	
Hyla cinerea	LSUMZ 48181	DQ830810	DQ830810		
Hyla cinerea	MVZ 1453854	AY549327	AY549327		
Hyla cinerea	TNHC 61054	AY680271			
Hyla cinerea		X86238	X86238		
Hyla cinerea		X86272	X86272		
Hyla cinerea		X86306	X86306		
Hyla euphorbiacea		AY843625	AY843625	AY844606	AY844072
Hyla euphorbiacea	KEK HEX31	GU989088	GU989088	GU944718	
Hyla euphorbiacea	KEK HEX32	GU989074	GU989074	GU944704	
Hyla euphorbiacea	KEK HEX369	GU989078	GU989078	GU944708	
Hyla euphorbiacea	KEK HEX370	GU989083	GU989083	GU944713	AY844073
Hyla femoralis		AY843627	AY843627	AY844609	AY844074

Hyla femoralis	KEK HFE3858	GU989073	GU989073	GU944703	
Hyla gratiosa		AY843630	AY843630	AY844611	AY844076
Hyla japonica		AY843633	AY843633	AY844615	AY844078
Hyla meridionalis		GQ916810	GQ916753	GQ916820	GQ916722
Hyla squirella		AY843678	AY843678	AY844670	AY844119
Hyla versicolor		AY843682	AY843682	AY844675	AY844124
Hyla walkeri		AY843684	AY843684	AY844677	AY844125
Hyla wrightorum	KEK HWR33	GU989079	GU989079	GU944709	
Hyla wrightorum	KEK HWR34	GU989081	GU989081	GU944711	
Hyla wrightorum	KEK HWR58	GU989085	GU989085	GU944715	
Hyla wrightorum	KEK HWR59	GU989069	GU989069	GU944699	
Hyla wrightorum	KEK HWR60	GU989089	GU989089	GU944719	
Isthmohyla pseudopum		AY843656	AY843656	AY844643	AY844101
Isthmohyla rivularis		AY843659	AY843659	AY844649	
Tlalocohyla picta		AY843654	AY843654	AY844640	AY844099
Tlalocohyla smithii		AY843668	AY843668	AY844659	AY844111

Table 22. Genbank sampling for Arabian Hylid species.

Table 22

Species	haplo code	12S	16S	Rhodopsin h1	Rhodopsin h2	Tyrosinase h1	Tyrosinase h2
Hyla orientalis	h420	GQ916749	GQ916798	GQ916814		GQ916713	GQ916716
Hyla orientalis	h251	GQ916749	GQ916799	GQ916815	GQ916816	GQ916713	
Hyla orientalis	h160a	GQ916745	GQ916792	GQ916815		GQ916715	GQ916717
Hyla orientalis	h164a	GQ916746	GQ916795	GQ916815		GQ916715	
Hyla felixarabica	h505	GQ916744	GQ916787	GQ916814		GQ916708	GQ916711
Hyla felixarabica	h087	GQ916739	GQ916782	GQ916814		GQ916708	
Hyla felixarabica	h195	GQ916743	GQ916787	GQ916814		GQ916708	GQ916710
Hyla felixarabica	hc162	GQ916741	GQ916785	GQ916814		GQ916706	GQ916707
Hyla felixarabica	h168	GQ916742	GQ916786	GQ916814		GQ916706	
Hyla savignyi	h062	GQ916723	GQ916758	GQ916811		GQ916694	
Hyla savignyi	h425	GQ916735	GQ916774	GQ916811		GQ916694	GQ916704
Hyla savignyi	h247	GQ916727	GQ916767	GQ916811		GQ916703	GQ916703
Hyla savignyi	h342b	GQ916724	GQ916762	GQ916811	GQ916812	GQ916691	
Hyla savignyi	h102	GQ916727	GQ916767	GQ916811		GQ916691	
Hyla savignyi	h096	GQ916724	GQ916765	GQ916812		GQ916691	
Hyla savignyi	h111	GQ916731	GQ916768	GQ916811	GQ916812	GQ916699	GQ916699

Table 23. Summary of divergence events and time estimates for given Arabian taxa.

Table 23

Divergence event	Number	Genetic Data	Analysis	Mean	Lower bound	Upper bound	Study
Amietophrynus arabicus vs. *xeros*	1	mtDNA, nucDNA	BEAST	20.5	15.91	26.3	Portik & Papenfuss (2015)
Amietophrynus tihamicus vs. *pentoni*	2	mtDNA, nucDNA	BEAST	17.9	9.0	26.6	Portik & Papenfuss (2015)
Echis coloratus vs. *pyramidum*	3	mtDNA	BEAST	19.4	15.4	24.5	Pook *et al*. (2009)
Echis pyramidum intraspecific	4	mtDNA	BEAST	8.3	6.3	10.9	Pook *et al*. (2009)
Varanus yemenensis vs. *albigularis*	5	mtDNA, nucDNA	BEAST	6.9	2.1	12.5	Portik & Papenfuss (2012)
Bitis arietans disjunct population	6	mtDNA	BEAST	4.0	2.8	5.3	Pook *et al*. (2009)
Pelomedusa subrufa disjunct population	7	mtDNA	BEAST	2.4	1.0	14.1	Wong *et al*. (2010)
Naja haje disjunct population	8	mtDNA	BEAST	1.8	0.9	2.5	Pook *et al*. (2009)
Papio hamadryas disjunct population	9	mtDNA	BEAST	0.32	0.16	0.51	Zinner *et al*. (2009)
Ichneumia albicauda disjunct population	10	mtDNA	BEAST	0.039	0.014	0.068	Fernandes (2011)
Hemidactylus (*barodanus, macropholis, granchii*) vs. (*jumailiae, montanus, yerburii*)	11	mtDNA, nucDNA	BEAST	9.8	5.3	11.7	Šmíd *et al*. (2013)
Uromastyx ocellata vs. *ornata*	12	mtDNA	Molecular clock	7.5	7.0	8.0	Amer & Kumazawa (2005)
Hemidactylus sp. 11 vs. (sp. 5, 6)	13	mtDNA, nucDNA	BEAST	5.9	3.8	8.3	Šmíd *et al*. (2013)
Duttaphrynus dhufarensis vs. (*stomaticus*, *olivaceus*)	14	mtDNA, nucDNA	BEAST	16.0	10.6	22.4	present
Euphlyctis ehrenbergii vs. *cyanophlyctis*	15	mtDNA, nucDNA	BEAST	14.7	10.2	19.8	present
Isodon (*schimperi*, *ramossimus*) vs. South Asian *Isodon*	16	cpDNA	BEAST	13.6	9.8	18.5	Yu *et al.* (2014)
Echis carinatus disjunct population	17	mtDNA	BEAST	0.9	0.5	1.3	Pook *et al*. (2009)
Hemidactylus persicus vs. (*luqueorum*, *hajarensis*)	18	mtDNA, nucDNA	BEAST	13.1	8.6	18.3	Šmíd *et al*. (2013)

Hyla felixarabica vs. *savignyi*	19	mtDNA, nucDNA	BEAST	6.5	3.9	9.5	present, Gvoždík *et al*. (2010)
Stenodactylus pulcher vs. (*arabicus*, *sharqiyahensis*)	20	mtDNA, nucDNA	BEAST	17.3	11.3	23.6	Metallinou *et al*. (2012)
Uromastyx benti vs. (*ocellata*, *ornata*)	21	mtDNA	Molecular clock	9.5	9.0	10.0	Amer & Kumazawa (2005)
Hemidactylus luqueorum vs. *hajarensis*	22	mtDNA, nucDNA	BEAST	9.4	5.8	13.3	Šmíd *et al*. (2013)
Stenodactylus slevini vs. (*grandiceps*, *affinis*)	23	mtDNA, nucDNA	BEAST	9.3	6.0	13.0	Metallinou *et al*. (2012)
Echis coloratus vs. *omanensis*	24	mtDNA	BEAST	7.9	5.5	11.3	Pook *et al*. (2009)
Stenodactylus leptocosymbotes vs. *doriae*	25	mtDNA, nucDNA	BEAST	7.0	4.2	10.1	Metallinou *et al*. (2012)
Hemidactylus jumailiae vs. (*montanus*, *yerburii*)	26	mtDNA, nucDNA	BEAST	7.0	4.7	9.8	Šmíd *et al*. (2013)
Stenodactylus grandiceps vs. *affinis*	27	mtDNA, nucDNA	BEAST	6.7	4.1	9.3	Metallinou *et al*. (2012)
Stenodactylus arabicus vs. *sharqiyahensis*	28	mtDNA, nucDNA	BEAST	6.4	3.9	9.3	Metallinou *et al*. (2012)
Hemidactylus shihraensis vs. *festivus*	29	mtDNA, nucDNA	BEAST	6.4	3.9	9.3	Šmíd *et al*. (2013)
Hemidactylus sp. 5 vs. 6	30	mtDNA, nucDNA	BEAST	4.2	2.4	6.1	Šmíd *et al*. (2013)
Hemidactylus montanus vs. *yerburii*	31	mtDNA, nucDNA	BEAST	4.0	2.5	5.7	Šmíd *et al*. (2013)
Hemidactylus alkiyumii 1 vs. (*alkiyumii* 2, 3)	32	mtDNA, nucDNA	BEAST	5.2	3.0	7.7	Šmíd *et al*. (2013)
Hemidactylus hajarensis intraspecific	33	mtDNA, nucDNA	BEAST	4.7	2.6	7.0	Šmíd *et al*. (2013)
Amietophrynus arabicus intraspecific	34	mtDNA, nucDNA	BEAST	4.4	2.8	6.6	present
Echis coloratus intraspecific	35	mtDNA	BEAST	4.3	2.9	6.0	Pook *et al*. (2009)
Hemidactylus alkiyumii 2 vs. *alkiyumii* 3	36	mtDNA, nucDNA	BEAST	3.0	1.7	4.5	Šmíd *et al*. (2013)
Hyla felixarabica intraspecific	37	mtDNA, nucDNA	BEAST	2.2	0.9	3.7	present, Gvoždík *et al*. (2010)
Hemidactylus festivus intraspecific	38	mtDNA, nucDNA	BEAST	2.0	1.2	3.0	Šmíd *et al*. (2013)
Hemidactylus yerburii intraspecific	39	mtDNA, nucDNA	BEAST	1.3	0.7	1.9	Šmíd *et al*. (2013)
Justicia areysiana intraspecific	40	AFLP, nrDNA ITS	Molecular clock	1.3	0.8	1.8	Meister *et al*. (2006)
Justicia areysiana intraspecific	41	AFLP, nrDNA	Molecular	1.2	0.8	1.6	Meister *et al*. (2006)

		ITS	clock				
Justicia areysiana intraspecific	42	AFLP, nrDNA ITS	Molecular clock	1.0	0.6	1.3	Meister *et al*. (2006)