Ecological divergence in the yellow-bellied kingsnake (*Lampropeltis calligaster*) at two North American biodiversity hotspots

A.D. McKelvy^{1,2} and F.T.Burbrink^{1,2,3}

¹Department of Biology, The Graduate School and Center, City University of New York, New York, NY 10016, USA

²Department of Biology, 6S-143, College of Staten Island, 2800 Victory Boulevard, Staten Island, NY 10314, USA

³Department of Herpetology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA

Corresponding author: Alexander D. McKelvy email: Smoochy@SnakeEvolution.org

Fax: (718) 982-3852

Key Words: Ecological speciation, statistical phylogeography, biodiversity, environmental niche modeling, snakes, conservation

Abbreviations:

SEUS – Southeastern United States; ESS – Effective sample size; GSI - genealogical sorting index

Abstract

Several biogeographic barriers in the Eastern Nearctic appear to reduce gene flow among populations of many species in predictable ways, however these patterns used to infer process of divergence may be deceiving if alternative modes of diversification are not considered. By using a multilocus statistical phylogeographic approach to examine diversity within a North American snake, *Lampropeltis calligaster*, we find that mode and timing of speciation near the Mississippi River embayment and Peninsular Florida, two main biodiversity hotspots in eastern North America, challenge previously held notions of strict vicariance as the causal factor behind patterns of divergence seen among taxa at these locations. We found three species inhabiting distinct ecological niches with divergences dating to the mid- and early-Pleistocene with subsequently stable or increasing effective population sizes, further supporting the idea that the Pleistocene was an important driver of diversification in North America. Our results lead to a revised hypothesis that ecological divergence has occurred in this group across environments associated with the Mississippi River and at the Florida peninsula. Importantly, in their western distributions, we show that species divergence is associated with the ecological transition from distinct forested habitats to grasslands, rather than the nearby Mississippi River, a barrier often implicated for many other organisms. Additionally, we stress the importance of examining each delimited lineage with respect to conservation, since ecological niche models suggest that by the end of the century changes in climate may negatively alter habitat suitability and, barring adaptation, substantially reduce the suitable range of two of the three species we identified.

1. Introduction

Geographic barriers to gene flow that isolate populations and generate species have been characterized using phylogeographic analyses (Avice, 2000), which play a strong role in understanding how species form with regard to timing, degree of isolation (allopatric or parapatric) and location, while also identifying previously undescribed or unrecognized diversity (Pyron and Burbrink, 2010). In areas such as the Nearctic, physical and ecological barriers have been examined across many taxa and appear to isolate populations, reducing gene flow across organisms in similar ways (Avice, 2000; Pyron and Burbrink, 2010). For wide-ranging taxa, one or more of these barriers is often found to influence population structure predictably. Causes for these repeated patterns observed among taxa may be deceiving in that they appear to have similar superficial effects but vary in key characteristics such as fine-scale location, timing or mode of divergence. Examining these key characteristics associated with biogeographic barriers enhances our understanding of regional processes related to speciation. Certainly, patterns may appear repeated among taxa, but underlying processes may vary greatly for each individual species.

In eastern North America, there are several major areas where terrestrial taxa show strong overlapping patterns of divergence. First, the Mississippi River region is particularly important for isolating populations and generating species (Pyron and Burbrink, 2010; Robison, 1986). During periods of climatic change in the Pleistocene, the volume of glacial meltwater and the river bed itself was subject to fluctuation in size and location (Royall et al., 1991; Smith, 1996). The current prevailing hypothesis is that the Mississippi River was instrumental in generating biodiversity in eastern North America (Burbrink et al., 2000; Fontanella et al., 2008; Gamble et

al., 2008; Guiher and Burbrink, 2008; Howes et al., 2006; Jackson and Austin, 2010; Leaché and Reeder, 2002; Lemmon et al., 2007; Pyron and Burbrink, 2009; Walker and Avise, 1998). Congruently, a longitudinal ecological transition zone from the central prairies to the eastern Nearctic forests suggests the possibility that ecological divergence in these different habitats drove speciation. This transition zone along the upper portions of the river was affected by glacial events during the Pleistocene and may not represent a long term, stable barrier (Robison, 1986; Royall et al., 1991), therefore the processes that isolated populations on separate sides of the Mississippi River (Burbrink et al., 2000; Fontanella et al., 2008; Guiher and Burbrink, 2008; Pyron and Burbrink, 2009; Ruane et al., 2014) may be due to ecological transition and not historical vicariance associated with the river itself. Both processes may be at work and the shared phylogeographic structure observed in codistributed taxa may only be pseudocongruent, indicating that spatial distributions of lineages are broadly similar across the Mississippi River, but may differ in mode, timing or actual location of speciation (Pyron and Burbrink, 2010; Soltis et al., 2006).

Another major region important for isolating populations in the Eastern Nearctic is located where the Florida peninsula meets the continental US (Fontanella et al., 2008; Guiher and Burbrink, 2008). The patterns of divergence there are typically explained by refugial isolation on islands in northern and central peninsular Florida during periods of elevation in sea level during the Oligocene, Pliocene and early Pleistocene (James, 1961; Soltis et al., 2006). Furthermore, this area is home to multiple rivers that may have been responsible for limiting secondary contact (Soltis et al., 2006). Alternatively, the Florida peninsula features an area of fairly sharp ecological turnover where major climate zones and ecoregional transitions occur (Omernik, 1987). This ecological transition from the southern coastal plain to the Piedmont may be a causal factor

driving regional ecological speciation, particularly in groups with key characteristics that are not well explained by traditional island or riverine vicariance.

Although there is disagreement about the role of the Pleistocene in generating modern diversity (Klicka and Zink, 1997), recent work has shown that changes in climate during the Pleistocene may have significantly driven diversification, especially within the common snakes of the genus *Lampropeltis* (Ruane et al., 2014). Estimating timing of diversification while including post-divergence demographics also helps determine how populations have responded to climatic cycles. Testing these hypotheses to assess the importance of the Pliocene and Pleistocene in shaping current biodiversity requires credible estimates of speciation times within taxa to accurately reflect evolutionary history.

To elucidate processes of diversification associated with these barriers at the Mississippi River region and in the Southeastern United States (SEUS), it is important to first examine organisms with distributions that cross these transition zones and respond to landscape features. Snakes are exceptional candidates for studies on the effect of ecology and landscape isolation; they are known to respond both to vicariant barriers and ecological transitions (Brandley et al., 2010; Burbrink et al., 2011). The yellow-bellied kingsnake, *Lampropeltis calligaster* (Harlan, 1827), is found across much of the southern United States and three subspecies are currently recognized, distinguished by color pattern and average number of infralabial scales (Blanchard, 1921; Price, 1987). This morphological variation also corresponds to areas on either side of ecological transitions associated with the Mississippi River, and near the Florida peninsula, offering an opportunity to test associated patterns and processes of speciation at these two biodiversity hotspots. The eastern subspecies of *L. calligaster*, the Mole Kingsnake, *Lampropeltis c. rhombomaculata* (Holbrook, 1836) is generally found east of the Mississippi River and on the

Piedmont and eastern seaboard, while *Lampropeltis c. calligaster* (Harlan, 1827), the Prairie Kingsnake, is primarily found in the eastern Great and Central Plains from Texas, Oklahoma and Kansas east to Indiana (Fig. 1). The two subspecies meet east of the Mississippi River, where morphological evidence suggests ranges overlap and hybridization occurs (Cook, 1945). The eastern and western subspecies were both originally described as unique species (Harlan, 1827; Holbrook, 1836) but were reduced to subspecies by Cook (1945) based on intergradation and parapatric distribution in the eastern portion of the State of Mississippi. A third and much rarer form, *Lampropeltis c. occipitolineata* (Price, 1987), the South Florida Mole Kingsnake, is poorly represented in museum collections and is found in only a few counties in South Florida (Krysko et al., 2000; Price, 1987). Importantly, the former two taxa were considered distinct species occupying distinctly different niches (Wright and Wright, 1957) prior to being arbitrarily delimited to subspecies (Cook 1945).

While statistical species delimitation is becoming common (Fujita et al., 2012), and it is clear that even in North America where surveys of species diversity have been conducted for more than 200 years, phylogeographic studies are still demonstrating that a substantial amount of biological diversity remains undiscovered in terms of uncovering numerous cryptic lineages and populations (Burbrink and Guiher, 2015; McKelvy et al., 2016; Myers et al., 2013a; Pyron and Burbrink, 2009; Ruane et al., 2014). However, the potential fate with regard to conservation of newly recognized species is rarely tested. An often-overlooked consequence of uncovering cryptic diversity is that newly erected taxa are subsets of the total range of the species complex from which they were delimited. Distinct scenarios for each lineage regarding the influence of climate change and urbanization are likely. Conservation efforts to facilitate the survival of these species must reflect the best available data (Malaney and Cook, 2013). Recent discoveries of cryptic

lineages with confined ranges are not rare, and small range size is understood as a factor correlated with extinction (Bielby et al., 2008; Niemiller et al., 2013); if researchers and stakeholders are expected to incorporate newly uncovered taxa into management plans, available information, including niche trajectories modeled from available climate data, should be examined.

In this study, we examine the phylogeographic history of the *L. calligaster* complex to first determine if distinct lineages occur at known geographical barriers. We then test specific causes of lineage diversification from among several potential candidate hypotheses. To do this, we investigate divergence times to assess if diversification corresponds with major Pleistocene events. We use redundancy analysis and Mantel tests to determine if the pattern of divergence we observe at the Mississippi River region is best explained by a model of vicariance at the river or by divergence across an ecological gradient. To test mode of speciation, we explore the niche identity of each species in a biogeographic context via environmental niche modeling and assess if gene flow is asymmetric between delimitable species in adjacent niches. We then use historical demographic analyses to determine if populations from different species respond to climatic fluctuations in similar ways. We generate an estimate of future suitable climatic space across North America, and discuss conservation implications due to changes in potential range size for the rare *occipitolineata* form of *L. calligaster*. Our results lead to a better understanding of what causes diversification at major biogeographic transitions in eastern North America.

2. Materials and Methods

2.1 Sampling, Sequencing and Gene Trees

We obtained tissue samples from 60 individuals from across the known distribution of *Lampropeltis calligaster* (Fig. 1; Supplementary Material 1); no protected or endangered species were involved in the work. Using ~4.5mm tissue subsample from each, we extracted DNA using Qiagen DNeasy kits following the provided protocol for animal tissues. We amplified one mitochondrial gene (Cytochrome b) and four nuclear loci (CLLAT, NT3, PRLR and VIM56) via polymerase chain reaction (PCR) using temperatures and primer sequences shown in Table 1. PCR products were cleaned using 1.5 µl of ExoSap-IT (USB Corporation). We used 2 µl of primer, 5 µl of water and 2 µl of Beckman Coulter DTCS template DNA on a Beckman Coulter CEQ 8000 Separation Module, System ID 490164 to sequence. We supplemented these data with sequences obtained in a previous study on *Lampropeltis* (Ruane et al., 2014) (Supplementary Material 2), and used the program Geneious v4.8.5 (Kearse et al., 2012) with the Geneious alignment algorithm to align all sequences. After conservatively trimming any sequences with missing data, we used PHASE v2.1.1 (Stephens and Donnelly, 2003) to determine the most probable pairs of alleles in the nuclear loci. We transformed files from FASTA alignments to Phase input files using the online converter Seqphase (Flot, 2010) and ran them multiple times to ensure consistency using the default threshold of 90%. We used these data in all genetic analyses. We estimated individual gene trees in BEAST v. 1.7.5 (Drummond et al., 2012). We determined best fit models of nucleotide substitution for each locus using ranked BIC scores in jModeltest 2.1.4 (Darriba et al., 2012).

2.2 Species Delimitation and Divergence Dating

To first find the number of distinct populations, we assigned individuals to genetic clusters without constraining them to species using Structurama (Huelsenbeck et al., 2011). We ran Structurama five times on the phased multilocus dataset using 10,000,000 MCMC cycles and

three chains with prior populations $k=1-4$, as well as five times without mtDNA to ensure that support values did not differ significantly, that the results were not driven exclusively by mtDNA and to reduce the chances that upstream processes influence delimitation (Olave et al., 2014). We sampled every 10,000 generations with a burn-in of one million generations removed.

We estimated divergence time and phylogeny in *BEAST using the groups assigned in Structurama and implemented in the program BEAST v. 1.7.5 (Drummond et al., 2012), ultimately to be used as a guide tree to delimit taxa (see below). All loci were included and we used the outgroups *Cemophora coccinea* and *Lampropeltis extenuata* based on Ruane et al. (2014). We estimated tree shape under a Yule prior, population size under a piecewise linear model with constant root, and assigned each locus the appropriate model estimated by jModelTest (i.e., HKY or HKY+ Γ). We placed a fossil calibration at the node uniting *Lampropeltis* from *Cemophora* as described in detail in Myers et al. (2013a). We estimated species trees in a series of 10 concurrent runs of 200 million generations each and combined with LogCombiner (Drummond et al., 2012), totaling two billion generations sampled every 200,000 generations. Burn in at 20% was removed with LogCombiner (Drummond et al., 2012) and TreeAnnotator (Drummond et al., 2012), and when assessed in Tracer v1.5 (Rambaut and Drummond, 2007), effective sample size (ESS) scores were 1,000 or higher for all parameters. Additionally, we generated admixture proportions for eastern and western lineages ($k=2$) in the program Structure (Pritchard et al., 2000) and used phylo.to.map in the R package phytools (Revell, 2012) to integrate and present admixture information with the phylogenetic tree across the range of the species.

To delimit species groups inferred by Structurama, we used BP&P (Yang and Rannala, 2010), implementing an rjMCMC method to estimate the posterior probability of nodes using the

species tree we generated in *BEAST. We used the species delimitation algorithm and allowed the fine-tuning parameters to automatically adjust so that swapping rates were between 0.30 and 0.70. We used three combinations of priors for ancestral population size and root age: large ancestral population with recent divergence (1,10|2,2000), large ancestral population with ancient divergence (1,10|1,10), and small population size with recent divergence (2,2000|2,2000). We ran both nuclear and nuclear + mitochondrial datasets five times to ensure that support values did not differ significantly and randomized lineage designations to ensure node support was not based solely on priors. We also ran BP&P without our guide tree to provide a posterior probability for number of species (Yang and Rannala, 2014).

To verify results across species delimitation methods, we also used the program STACEY (Jones, 2016) implemented in BEAST 2.4.2 (Bouckaert et al., 2014). STACEY improves the efficiency of the DISSECT model (Jones et al., 2015) to delimit species and infer species trees without guide trees. For the STACEY run, we used the HKY model of evolution, all ploidy scalars at 2.0 for nuclear genes, a yule species tree model with collapse height of 0.0001, birth diff rate of 100.0, relative death rate and collapse weight at 0.5 and origin height of 100.0. We assigned lognormal priors to growth and clock rates instead of $1/x$ as described in the tutorial and a relative death-rate prior as uniform between -0.5 and 0.5. We ran the MCMC chain for 100 million generations, storing the state to disk every 10 million generations and logging trees every 50,000 generations.

Additionally, we used the genealogical sorting index (GSI) test (Cummings et al., 2008) to measure degree of exclusive ancestry of each delimited taxon on a scale of 0 (non-exclusive) to 1 (monophyletic) to determine how often each gene tree agrees with coalescent species

delimitation.

2.3 Historical Demography and Migration

Most diversification within the genus *Lampropeltis* occurred in the Pleistocene (Ruane et al., 2014), with subsequent population size stability or expansion (Burbrink et al., 2011; Myers et al., 2013a). To understand historical demography and confirm that this is also true for *L. calligaster*, we investigated changes in effective population size over time (N_e) in delimited taxa using Extended Bayesian Skyline Plots (EBSP) in the program BEAST v.1.7.5 (Drummond et al., 2012). Each locus was modeled using results from jModeltest and the parameters specified in the tutorial for EBSP. We scaled time by using an estimated mitochondrial substitution rate of 1×10^{-8} per site per generation (Eo and DeWoody, 2010). The dataset was run for 500 million generations, and all ESS values analyzed in Tracer v1.5 (Rambaut and Drummond, 2007) were greater than 200.

To test gene flow between lineages occurring in distinctly different niches, we determined which direction migration may be occurring between adjacent eastern and western lineages by employing the program Migrate 3.6 (Beerli and Palczewski, 2010). Migrate tests four models of migration: panmictic population, bilateral migration and two models of asymmetric migration (Beerli and Palczewski, 2010). We set inheritance scalars for mitochondrial and nuclear genes to reflect that our mitochondrial gene has 1/4 the effective population size of our nuclear genes. Four heated chains were used for MCMC analysis, with values of 1, 3, 10 and 100,000. We summarized results over 10 runs and compared them to ensure stationarity, and checked resulting

output files to ensure posterior distributions and marginal likelihood scores indicated burn-in was achieved.

Additionally, to better understand migration given isolation with ancestry, we explored 24 nested parameter models (Supplementary Material 3) of migration and divergence in the program IMa2p (Sethuraman and Hey, 2016) and calculated AIC scores to determine which parameter estimates were most likely given the molecular data. We provided a scaled mutation rate estimate on the mitochondrial gene of 1×10^{-5} , set inheritance scalars of 0.25 for the mitochondrial gene and 1 for nuclear genes, assigned the HKY models of molecular evolution from jModeltest, a generation time of three years (Ernst and Ernst, 2003) and ran the program in M mode 5 times with a random seed, using four CPU cores and ten chains per core, assigning the recommended heating parameters for medium-sized datasets. We used log-likelihood ratio tests on our 200,000 generated genealogies to determine which of the 24 nested models were best supported and repeated this process four times to ensure consistency.

2.4 Ecological Niche Modeling

To model the niche of *L. calligaster* and test ecological associations for each lineage, we used all known locations of *L. calligaster* collected from records available through the HerpNet data portal with 35 bioclimate variables from the CliMond Database (Kriticos et al., 2012) to create ecological niche models (ENMs). We assessed correlation of variables using ENMTOOLS (Warren et al., 2010), and all strongly (>0.89) correlated variables were removed, leaving 15 bioclim variables on which to base our models. Individuals were assigned to populations by range, consistent with subspecies designation. We downsampled potentially erroneous records and duplicates, as well as points in the same cell from the dataset, leaving 155 georeferenced

localities for the eastern group, 340 localities for the western group and 17 localities for the South Florida lineage. We then used the R package *spThin* (Aiello-Lammens et al., 2015) to thin our dataset spatially to remove sampling bias yet retain as many records and variation as possible. To recover evenly-spaced datasets for our eastern and western groups, we used a 30k threshold in the algorithm, and used a 10k threshold for our rare South Florida localities. For all groups, we repeated our dataset generation 100 times and evaluated them using the function *plotThin* to ensure that the number of records had plateaued. This left us with 82 records for the western group, 39 records for the eastern group, and 8 for the South Florida lineage, reported in Supplementary Material 3. We trimmed our area of extent to include only the regions where *Lampropeltis calligaster* is distributed in eastern North America, created datasets of 10,000 pseudoabsences around our datasets using the R command *BIOMOD_FormatingData* in the package *Biomod2* (Thuiller et al., 2013), and used *Maxent* v3.3.3 (Phillips and Dudík, 2008) to construct initial ENMs as it has been shown to be robust in creating species distribution models from presence only data and from small datasets (Hernandez et al., 2006; Wisz et al., 2008). We replicated our runs 500 times for each group, reserved 30% of samples as a training dataset, created response curves and jackknifed our data to measure variable importance. To ensure that our results were robust, we then used an ensemble modeling approach with the function *BIOMOD_EnsembleForecasting* from the package *Biomod2* (Thuiller et al., 2013), which combines a generalized linear model, generalized additive model, generalized boosting model, classification tree analysis, artificial neural network, surface range envelope, flexible discriminant analysis, multiple adaptive regression splines, maximum entropy and a random forest model into a single output to create the final projections we used to evaluate species distributions. We used 0.7 as our ROC score quality threshold evaluation metric.

We then used ENMTOOLS (Warren et al., 2010) to test whether lineages occupy identical or distinct environmental niches. The niche identity test examines whether niches drawn randomly from the combined samples are significantly different from those drawn from the dataset. The range break test draws random lines through a pooled sample range and estimates degree of similarity to test if populations are linked to either side of a sharp environmental transition (Glor and Warren, 2011). Relative rank as well as summary statistics I and Schoener's D were generated for each repetition and the observed overlap in niche was compared to this distribution (Glor and Warren, 2011). If observed overlap values fell outside of the 95% confidence interval of the null distribution, the null hypothesis was rejected.

To estimate how the environment and the predicted niche of these organisms may change through time as a result of global climate change, we projected our ensemble models onto bioclimate data from the estimated IPCC fourth assessment report models of climate change for the year 2100 (Susan, 2007). We estimated change in size of suitable habitat by quantifying change in number of 30 second cells from the ENM outputs with extracted score 0.4 or higher (the value at which the known range of the species is best represented) in both the current and future niche predictions, as well as the intersection between the two, by summing the counts of the floating point rasters using the Int Spatial Analyst Tool in ArcGIS. Clamping mask estimates revealed few variables outside their training ranges were unlikely to have a large effect on predicted suitability scores.

To determine which fine scale habitat types are associated with the presence of *L. calligaster* for each detected lineage, we used ArcGIS (ESRI, 2011) and all locality records to extract specific habitat information from GAP land-cover data, a dataset provided by the USGS that focuses on habitat identification (Homer et al., 2004). GAP land-cover data are likely highly

dependent on the climate variables recorded in the WorldClim database, but differ from that of strict environmental climate data in that they identify objectively defined ecosystem types based on remote sensing and survey data of flora. This information provided names of the ecosystems and fine-scale habitats in which specimens are found. Habitat results were aggregated by type and tallied. Localities falling on currently developed or otherwise impermeable surfaces were discarded.

2.5 Ecological Speciation Versus Riverine Vicariance

To determine if patterns of divergence at the Mississippi River Embayment are best explained by ecological divergence or vicariance associated with the Mississippi River itself, we performed both a traditional Mantel test and a redundancy (RDA) analysis, which allowed us to test how much of our variance is explained by each of the competing hypotheses using matrices of environmental, geographic and genetic data (Legendre and Fortin, 2010). We extracted environmental and geographic data for each individual from the layer files used for niche modeling. We performed a principle component analysis (PCA) on the extracted bioclim data and used the distance of the first axis (92% of the variance), in our environmental matrices. After importing our datasets into R (Team, 2000), we created distance matrices using the dist function in the R package ade4 (Thioulouse et al., 1997). These tests determine correlation between variables; significant correlation in a Mantel test ($p < 0.05$) and large values for constrained results in RDA indicate strong relationships between matrices. Results showing values with high correlation between spatial distribution and genetic distance matrices would indicate divergence resulting from riverine vicariance as the primary driver of speciation (or isolation by distance), while correlation between environmental niche and genetic distance matrices would support ecological speciation. We used the functions Mantel.rtest in the package ade4 with 9999

replicates and function `rda` in the package `vegan` (Oksanen et al., 2007) to test which matrices (ecological distance or spatial distribution) best describe the variation in the data. Using the same variables we also examined the influence of ecology, geographic distance and vicariance using both standard and partial RDA to isolate the effects of variables and tested significance with the `anova` function.

3. Results

3.1 Sampling, Sequencing and Gene Trees

We sequenced 2,694 base pairs (bp) of DNA from five unlinked loci and estimated the best-fit models of substitution and gene trees for each locus (Table 1). Of 60 individuals, one or more loci could not be recovered from 29 samples (Supplementary Material 1), resulting in 12% missing data. The gene *CLLAT* was difficult to amplify across all individuals and therefore this gene comprised 58% of missing data, but was retained for all subsequent analyses.

3.2 Species Delimitation and Divergence Dating

All runs of *Structurama* for both mtDNA and nDNA revealed that most individuals (98%) are classified into western, eastern and South Florida lineages (Fig. 1). Support for three populations was highest at 65% posterior probability, then four populations at 31%.

Divergence dating and species tree estimation in **BEAST* supported a recent split of the South Florida clade from the eastern clade at 0.88 MYA (95% HPD=0.003–1.72 MYA) and an older split of the western group from the eastern + South Florida lineages at 1.56 MYA (95% HPD=0.64–2.92 MYA; Fig. 1). The species tree showed a sister relationship between two eastern and a single western lineage (100% posterior probability), corresponding to subspecies *L. c.*

rhombomaculata, *L. c. occipitolineata* and *L. c. calligaster*, respectively. Placement of the taxon sister to *L. c. rhombomaculata*, the Florida subspecies *L. c. occipitolineata*, was more poorly supported (79.7% posterior probability). Results from the program Structure and phylogeographic relationships suggest admixture east of the Mississippi River (Fig. 2). We provide a traditional Structure plot in Supplementary Material 4.

All separate runs of BP&P indicated that species were delimited with a probability of 1.0 for the eastern, western and South Florida lineages, corresponding to the recognized subspecies *L. c. rhombomaculata*, *L. c. calligaster* and *L. c. occipitolineata* respectively. Support for the delimitation of either clade approached zero when individuals within lineages were randomized. Without using a guide tree, BP&P returned 100% posterior probability for delimitation of three species. Results from STACEY agreed with the other delimitation and assignment methods, with highest support for three distinct clusters (east, west, Florida), recovering them 43% of the time, as opposed to 27% (two clusters, east+Florida and west), 26% (one cluster, east+west+Florida), 3% (two clusters, east+west and Florida) and .002% (two clusters, east and west+Florida). The GSI test (Cummings et al., 2008) revealed discordance among genes, with nuclear gene trees sorting (0.5 or higher) with respect to individual species tree designations 60% of the time (Table 2; see Supplementary Material 1 for individual gene trees).

3.3 Historical Demography and Migration

Extended Bayesian Skyline Plots revealed increasing N_e in eastern and western lineages over the last 100,000 years (Supplementary Material 5). One population size change was sampled most frequently for the parameter demographic.populationSizeChanges, indicating the population

has expanded slightly over the last 100,000 years. Sample size and minimal sequence divergence prevented reliable estimates for the South Florida population.

Results from Migrate 3.6 (Beerli and Palczewski, 2010) indicate that equal migration between eastern and western populations is most likely (Table 3). A model of east to west and west to east were supported by similar likelihood scores, and taking the product of M and Θ (Beerli and Palczewski, 2010) we estimated migrants per generation between lineages as: west to east 0.390 (95% HPD=0.341-0.438), east to west 0.104 (95% HPD=0.091-0.117; (Supplementary Material 3)).

Results from IMA2 were consistent with those of Migrate. Under the isolation with migration model, we found highest support for nested models of migration in both directions with relatively equal population sizes throughout time, and we recovered rates of migration per generation of 0.4-0.9 in each direction (Supplementary Material 3).

3.4 Ecological Niche Modeling

The identity test showed observed values for I and D were significantly different from the null distribution (t-test, $p < 0.05$), indicating that eastern and western lineages as well as eastern and southern lineages occupy distinct ecological niches in North America. The observed values for the range break test also fell outside the 95% distribution of the null, indicating that there is a transition in ecological niche between eastern and western lineages. The variables that contributed most to the projections (Supplementary Material 7) for *Lampropeltis calligaster* were: highest weekly radiation, mean temperature of warmest quarter and isothermality. For *L. rhombomaculata*: lowest weekly moisture index, mean temperature of driest quarter and highest

weekly radiation. For *L. occipitolineata*: precipitation of warmest quarter, annual mean temperature and highest weekly radiation.

Tests for fine scale habitat occupancy using GAP data revealed (Fig. 3) that western *L. calligaster* were directly sampled primarily in prairie, pasture and shrublands (61.4%, n=86) and secondarily (37.9%, n=53) in mesic forests. *Lampropeltis rhombomaculata* was sampled primarily in temperate oak and pine forests (64.1%, n=41), and secondarily in pasture (32.8%, n=21). *Lampropeltis occipitolineata* was sampled primarily in freshwater wet meadow (37.5%, n=3) and pasture (37.5%, n=3) and secondarily in southeastern warm temperate forest (25%, n=2).

Future climate models predicted that core habitat for *L. calligaster* will remain stable over 99.0% of its range, with overall suitable climatic space increasing 69.1% in distribution northward and eastward across the range currently occupied by *L. rhombomaculata* (Fig. 4). Climate models predict that suitable niche for *L. rhombomaculata* and *L. occipitolineata* will be reduced drastically, as no cells in future climate models for *L. occipitolineata* contained suitability scores higher than 0.4 (100% reduction) and only a few cells contained suitable scores for *L. rhombomaculata* (98.5% reduction). Core habitat will also be reduced, with only 7.0% of current niche space remaining suitable under projected conditions. Number of cells are reported in Table 4.

3.5 Ecological Speciation Versus Riverine Vicariance

The results from both the Mantel test and the RDA suggested that the variation seen between *L. calligaster* and *L. rhombomaculata* is best described by divergence in environment rather than riverine vicariance or isolation by distance. Mantel tests indicated that only ecological

distance and genetic distance matrices were significant ($p=0.04$, geographic distance $p=0.49$). Our RDA analysis also returned similar results, with ecological and genetic distance explaining 86.7% of the data ($p=0.054$). Results for the influence of the Mississippi River show that riverine divergence (position east or west of the Mississippi River) only explained 2.8% of the results ($p=0.886$), indicating that the current position of the river itself likely did not play a strong role in isolating these species (Supplementary Material 8).

4. Discussion

4.1 Historical Biogeography and Phylogeography

Speciation in the Eastern Nearctic is complex and processes generating observed patterns may not be obvious. Our results provide a better understanding of the historical role of the biogeographic barriers near the Mississippi River and at the Florida peninsula with respect to generating ecologically distinct species. Strict vicariance may not be sufficient to describe processes generating distinct lineages for codistributed taxa in regions where habitat transitions are tied closely with a clear isolating barrier. Large barriers such as the Mississippi River are frequently cited as driving regional divergence, but specific tests should be used to account for pseudocongruent effects of underlying processes (Soltis et al., 2006). Results for *L. calligaster* suggest that ecological diversification, rather than the prevalent hypothesis of riverine vicariance, drove speciation. Therefore it is possible, particularly in the Eastern Nearctic, that many regionally co-distributed species with similar distributional patterns of lineages may have been

generated from potentially different processes giving rise to pseudocongruent patterns (Pyrone and Burbrink, 2010).

Here, we determined that speciation has occurred at two major biodiversity hotspots in North America in *Lampropeltis calligaster*, with each of these lineages currently occupying distinct niches. Coalescent delimitation as well as ecological and morphological (Blanchard, 1921; Price, 1987) data indicate that *Lampropeltis calligaster* is composed of three closely related species, historically considered distinct taxa (Blanchard, 1921; Holbrook, 1836; Schmidt and Davis, 1941): *L. calligaster* in the western prairies, *L. rhombomaculata* in the eastern forests, and *L. occipitolineata* in South Florida wet prairies.

While it is possible that the Mississippi River was the initial cause of speciation between *L. calligaster* and *L. rhombomaculata*, ENMs indicate that each species currently occupies a distinct ecological niche in Eastern North America. GAP land-cover data indicates that prairies and pastureland are preferred habitats for *L. calligaster* in the west, and dry oak and pine forests are preferred for *L. rhombomaculata* in the east. Signatures of migration suggest that speciation in these taxa was either accompanied by gene flow where niches overlap or has resumed in secondary contact. This separation in niche coupled with climatic shifts may have driven speciation. Redundancy analyses revealed that genetic divergence is best described by environmental factors and only poorly by riverine vicariance. Interestingly, other species pairs reveal divergence near the Mississippi River but diversification was likely driven by habitat changes; Burbrink and Guher (2015) found a transition in lineages of copperhead snakes (*Agkistrodon contortrix*) found east of the Mississippi River that roughly corresponds to the same area of divergence estimated in this study between *Lampropeltis calligaster* and *L. rhombomaculata*. Divergence times between copperhead lineages were reported overlapping

those of *L. calligaster* at 0.96 MYA (0.46–1.56 MYA), indicating that both *L. calligaster* and *Agkistrodon contortrix* lineages may have diverged at a similar time and area during the Pleistocene, suggesting a congruent patterns of diversification. It is also therefore likely that the Mississippi River itself may not have been the driver of speciation, as indicated in other taxa (Burbrink and Guiher, 2015; Burbrink et al., 2000; Guiher and Burbrink, 2008), but rather the effect of ecological divergence into unique habitats found on each side of the river. Studies revealing vicariance at or near this barrier will benefit from including alternative models of ecological speciation in hypothesis testing.

We demonstrate roughly equal rates of historical migration between eastern and western species. This information corroborates Cook (1945), who found morphological evidence that a small contact zone composed of intermediate phenotypes (considered hybrids) between *L. calligaster* and *L. rhombomaculata* exists east of the Mississippi River. The frequency of this migration, ~0.39 individuals/generation (Migrate) and 0.4-0.9 (Ima2), corresponds to that of other similarly distributed species in the genus *Lampropeltis* (Ruane et al., 2014) and is lower than the threshold of one to ten individuals per generation required for panmixia (Mills and Allendorf, 1996). Existence of gene flow is corroborated by our admixture results, (Fig 2), with admixed individuals present on the leading edge of the contact zone of these two species. Low levels of migration may slow rates of ecological speciation and be detrimental if species accumulate non-adaptive alleles in their respective ecoregions, or beneficial, in that migration allows introduced alleles to be acted upon by selection (Schluter, 2001).

The oldest divergence, that of *L. calligaster*, corresponds with the Pleistocene or late Pliocene at 1.55 MYA (95% HPD=0.635–2.921 MYA) and corroborates the hypothesis that the species diversity currently found within the genus *Lampropeltis* was likely generated during the

Pleistocene (Ruane et al., 2014). This finding adds further evidence that Pleistocene climatic cycles may have been instrumental in generating species diversity, a controversial subject once described as a “failed paradigm” (Klicka and Zink, 1997).

Our post-divergence demographic analyses indicate a stable or increasing population size over the last 100,000 years. This result contradicts some biogeographic explanations of constrained population sizes in small refugia (Hewitt, 2000). Other species in North America show similar patterns where signatures of population expansion are observed throughout the Pleistocene (Feldman and Spicer, 2006; Myers et al., 2013a; Myers et al., 2013b), questioning the general assumption that organisms responded to the Pleistocene in the same way (Hewitt, 2000).

The divergence time between *Lampropeltis rhombomaculata* and *L. occipitolineata* at 0.88 MYA (95% HPD=0.003–1.723 MYA) also occurs during periods of glaciation in North America and both lineages occupy distinct environmental niches. This information adds to the list of species that show a similar pattern in the SEUS but challenges the role that island isolation has played in influencing local biodiversity. It is difficult to conclusively determine the mode and location of speciation with range-limited taxa such as *L. occipitolineata*, which is restricted to south Florida and separated from their sister taxon by > 500km. Ecological divergence suggests that the role differing habitats play may be a driver of regional speciation. Historically, the divergence in other taxa with a south-central Florida distribution is attributed to elevated sea levels during the last interglacial period isolating lineages from adjacent populations distributed on the North American continent (James, 1961; Krysko et al., 2016; Soltis et al., 2006). Divergence dates for other taxa at the Florida peninsula range from two to seven million years (Guiher and Burbrink, 2008; Kubatko et al., 2011), suggesting that the region may be

instrumental as a “species pump” by producing a high number of species periodically over a large time scale through vicariant and ecological processes.

4.2 Conservation Implications

Urbanization of the Florida landscape since 1960 has been faster than anywhere else in the US, with human population doubling every 20 years (Reynolds, 2001). Many species are threatened due to urbanization, and coupling this process with climate change imperils many additional species in Florida (Reece et al., 2013). While current and future niche projections of *L. calligaster* indicate stable core habitat and increasing overall range, estimates of future climate reveal a troubling prediction for *Lampropeltis rhombomaculata* and *Lampropeltis occipitolineata*: areas once found suitable may change drastically over the next century. Assuming no adaptation, climate models indicate that there will be minimal preferred niche space remaining for these two species in 2100. Taxa such as *L. occipitolineata* locked in a rapidly developing peninsula may be unable to respond quickly enough to handle the effects of climate change and urbanization (Parmesan, 2006). Niche evolution is likely far outpaced by climate change (Quintero and Wiens, 2013), uncovering the potential range shifts in these species highlights the downstream implications of species delimitation.

4.3 Taxonomic Conclusions

Species delimitation reveals that *L. calligaster* is composed of three taxa that correspond with previously described subspecies. These three lineages also occupy unique niches in North America and display quantifiable differences in morphology (Blanchard, 1921; Price, 1987). We thus elevate *Lampropeltis calligaster*, *Lampropeltis rhombomaculata* and *Lampropeltis occipitolineata* as species.

Low sample sizes have been shown to be sometimes problematic when dealing with species delimitation (Rittmeyer and Austin, 2012). Due to sampling here, recognition of *Lampropeltis occipitolineata* may be considered by some as premature, however, in addition to these coalescent methods identifying and supporting *Lampropeltis occipitolineata* as a distinct species across five genes, the species shows a 500 km separation in range from *L. rhombomaculata* and exists in a distinctly different niche, displaying striking morphological differences from its sister lineage (Price, 1987). We therefore recommend that this taxon be elevated to the species level.

4.4 Species Synonymy and Distribution

Lampropeltis calligaster (Harlan 1827)

Lampropeltis calligaster (Harlan 1827). Prairie Kingsnake. Type locality: "Missouri", restricted to "vicinity of St. Louis" by Schmidt (1953). No holotype designated. Three syntypes (*Ophibolus evansii*) have been designated (Blaney, 1979); USNM 1593, an adult male, an adult and a juvenile female, collected in central Illinois by R. Kennicott, in 1855.

Coluber calligaster Harlan, 1827:359.

Coluber eximus (part): Holbrook, 1842:69.

Ablabes triangulum, var. *calligaster*: Hallowell, 1856:244.

Ophibolus evansii Kennicott, 1859:99. Type-locality, "Central Illinois."

Lampropeltis calligaster: Cope, 1860:255.

Coronella evansii: Jan, 1863:237.

530 *Ophibolus calligaster*: Cope, 1875:37.

531 *Ophibolus triangulus calligaster*: Garman, 1883:155.

532 *Coronella calligaster*: Boulenger, 1894:198.

533 *Lampropeltis calligaster calligaster*: Cook, 1945:48.

534 *Distribution: Lampropeltis calligaster* is composed of all populations west of the Mississippi
 535 River in the plains states and crosses the Mississippi River embayment east to Illinois, Indiana,
 536 Kentucky, Western Tennessee and northern Mississippi, favoring prairie and pastureland habitat.
 537 This population is considered distinct based on a two-part coalescent analysis of five unlinked
 538 loci. Morphological description follows (Blanchard, 1921): 46-78 (60) dorsal blotches; ventral
 539 plates 196 to 215, (205); caudals 38-57, males 44-57 (51); females 38-52 (46); supralabials 7,
 540 sometimes 8; infralabials usually 9 often 10 occasionally 11; oculars 1 and 2; temporals normally
 541 2+3+4; body slender and uniform; head scarcely distinct from the neck; snout longer and more
 542 pointed than in *L. rhombomaculata*; tail short and tapers rapidly. *Lampropeltis calligaster*
 543 contacts and may hybridize with *Lampropeltis rhombomaculata* in Northern Mississippi (Cook,
 544 1945).

545 ***Lampropeltis rhombomaculata* (Holbrook 1840)**

546 *Coluber rhombomaculatus* Holbrook, 1840:103. Mole Kingsnake.

547 Type-locality, "Georgia and Alabama," restricted to "vicinity of Atlanta, Georgia" by Schmidt
 548 (1953). No holotype designated.

549 *Coronella rhombo-maculata*: Holbrook, 1842:103.

550 *Ophibolus rhombomaculatus*: Baird and Girard, 1853:86.

551 *Lampropeltis rhombomaculata*: Cope, 1860:255.

552 *Ophibolus triangulus rhombomaculatus*: Garman, 1883:156.

553 *Lampropeltis rhombomaculatus*: Garman, 1892:9.

554 *Lampropeltis calligaster rhombomaculata*: Cook, 1945:48.

555 Distribution: *Lampropeltis rhombomaculata* is composed of eastern and southern populations in
 556 the state of Mississippi, favoring eastern dry pine and oak forests south to Louisiana, east to
 557 eastern Tennessee and the panhandle of Florida as well as on the Piedmont northwards along the
 558 eastern seaboard. This population is considered distinct based on this two-part coalescent-based
 559 analysis of five unlinked loci. Morphological description follows (Blanchard, 1921): 48-64 (55)
 560 dorsal blotches; ventral plates 191-213; caudals 31-55, males 37-55 (45) females 31-44 (41);
 561 supralabials 7; infralabials 8 often 9; oculars 1 and 2; temporals 2+3+4; body shape and pattern
 562 similar to that of *L. calligaster*. *Lampropeltis rhombomaculata* contacts and may hybridize with
 563 *Lampropeltis calligaster* in Northern Mississippi (Cook, 1945).

564 ***Lampropeltis occipitolineata* (Price 1987)**

565 *Lampropeltis calligaster occipitolineata* Price 1987. South Florida Mole Kingsnake.

566 Type locality: South Florida. Holotype: FMNH 48265. Adult male from Okeechobee,
 567 Okeechobee County, Florida. Collected 31 May 1942 by Reid Paulk.

568 Distribution: *Lampropeltis occipitolineata* is composed of all populations in South Florida and
 569 has been recorded from Brevard (Layne et al., 1986; Means, 1992), Desoto (Krysko and Hurt,
 570 1998), Glades, Hendry (Krysko et al., 2000), Okeechobee (Layne et al., 1986; Means, 1992) and
 571 Osceola (Rushton, 2013) counties. This population is considered distinct based on a two-part

572 coalescent analysis of five unlinked loci. Morphological description follows Price 1987:
 573 *Lampropeltis occipitolineata* is characterized by at least 78 distinct dorsal blotches, a complex
 574 network of dark lines on the back of the head, retention of the juvenile color pattern by adults, an
 575 obtuse canthus rostralis, a widening of the head in the temporal region and no more than 21 rows
 576 of dorsal scales.

577 **Acknowledgments**

578 We thank P. Beerli, X. Chen, E. Myers, L. Revell, S. Ruane and A. Sethuraman for discussions
 579 and troubleshooting during sequencing and analysis, A. Ozleski-McKelvy for help with
 580 troubleshooting in ArcGIS, A. Figueroa for assistance in field collecting, J. Ahern, C. Austin, D.
 581 Ditman, R. McDiarmid, N. Gilmore, and R. Melius for help locating or loaning specimens, B.
 582 Boone, J. Briggler, T. Colston, M. Coone, W. Van Devender, T. Huff, J. Holbrook, M.
 583 Kenderdine, J. Kleopfer, K. Irwin, S. Ruane, D. Shepard, K. Wray, A. Wynn and numerous others
 584 for contributing tissue samples to the project and A. Stein for help georeferencing. E. Myers
 585 provided comments that improved the manuscript. Photo credits go to K. Enge, K. Krysko and D.
 586 Shepard. Funding was provided by a US National Science Foundation Grant (DEB 1257926) to
 587 FTB.

588 **Author Contributions**

589 A.D.M and F.T.B conceived the study; A.D.M generated sequence data and performed analysis.
 590 Both authors contributed to writing the paper.

591 **Data Accessibility**

592 DNA sequences: Genbank accession numbers in Supplementary Material 2, Data available on
 593 Dryad at XXXXX

594 GPS Coordinates used for ENM: Supplementary Material 6

595 SEUS GAP Land-cover data: <http://gapanalysis.usgs.gov/data/>

596 Bioclimatic Data: <https://www.climond.org/>

597 Literature Cited

598 Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B., Anderson, R.P., 2015. spThin: an R package
 599 for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* 38, 541-
 600 545.

601 Avise, J.C., 2000. *Phylogeography: the history and formation of species*. Harvard University Press.

602 Beerli, P., Palczewski, M., 2010. Unified Framework to Evaluate Panmixia and Migration Direction Among
 603 Multiple Sampling Locations. *Genetics* 185, 313-326.

604 Bielby, J., Cooper, N., Cunningham, A., Garner, T., Purvis, A., 2008. Predicting susceptibility to future
 605 declines in the world's frogs. *Conservation Letters* 1, 82-90.

606 Blanchard, F.N., 1921. A revision of the king snakes: genus *Lampropeltis*. Govt. print. off.

607 Blaney, R., 1979. *Lampropeltis calligaster*. *Cat. Amer. Amphib. Reptiles*, 229.221-229.222.

608 Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A.,
 609 Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput*
 610 *Biol* 10, e1003537.

611 Brandley, M.C., Guiher, T.J., Pyron, R.A., Winne, C.T., Burbrink, F.T., 2010. Does dispersal across an
 612 aquatic geographic barrier obscure phylogeographic structure in the diamond-backed watersnake
 613 (*Nerodia rhombifer*)? *Molecular Phylogenetics and Evolution* 57, 552-560.

614 Burbrink, F.T., Guiher, T.J., 2015. Considering gene flow when using coalescent methods to delimit
 615 lineages of North American pitvipers of the genus *Agkistrodon*. *Zoological Journal of the Linnean Society*
 616 173, 505-526.

617 Burbrink, F.T., Lawson, R., Slowinski, J.B., 2000. Mitochondrial DNA phylogeography of the polytypic
 618 North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept. *Evolution* 54, 2107-
 619 2118.

620 Burbrink, F.T., Yao, H., Ingrasci, M., Bryson Jr, R.W., Guiher, T.J., Ruane, S., 2011. Speciation at the
 621 Mogollon Rim in the Arizona Mountain Kingsnake *Lampropeltis pyromelana*. *Molecular Phylogenetics and*
 622 *Evolution* 60, 445-454.

623 Conant, R., Collins, J.T., 1998. *Field guide to reptiles and amphibians*. Download iTunes eBook.

624 Cook, F.A., 1945. Intergradation of *Lampropeltis calligaster* and *L. rhombomaculata* in Mississippi. *Copeia*
 625 1945, 47-48.

626 Cummings, M.P., Neel, M.C., Shaw, K.L., 2008. A genealogical approach to quantifying lineage divergence.
 627 *Evolution* 62, 2411-2422.

- 628 Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and
629 parallel computing. *Nature Methods* 9, 772.
- 630 Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the
631 BEAST 1.7. *Molecular biology and evolution* 29, 1969-1973.
- 632 Eo, S.H., DeWoody, J.A., 2010. Evolutionary rates of mitochondrial genomes correspond to diversification
633 rates and to contemporary species richness in birds and reptiles. *Proceedings of the Royal Society B: Biological Sciences* 277, 3587-3592.
- 634 Ernst, C.H., Ernst, E.M., 2003. *Snakes of the United States and Canada*. Smithsonian Institution.
- 635 ESRI, 2011. *ArcGIS Desktop: Release 10*. Environmental Systems Research Institute, Redlands, CA.
- 636 Feldman, C.R., Spicer, G.S., 2006. Comparative phylogeography of woodland reptiles in California:
637 repeated patterns of cladogenesis and population expansion. *Mol Ecol* 15, 2201-2222.
- 638 Flot, J.F., 2010. SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence
639 alignments. *Molecular Ecology Resources* 10, 162-166.
- 640 Fontanella, F.M., Feldman, C.R., Siddall, M.E., Burbrink, F.T., 2008. Phylogeography of *Diadophis*
641 *punctatus*: extensive lineage diversity and repeated patterns of historical demography in a trans-
642 continental snake. *Molecular Phylogenetics and Evolution* 46, 1049-1070.
- 643 Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A., Moritz, C., 2012. Coalescent-based species
644 delimitation in an integrative taxonomy. *Trends in ecology & evolution* 27, 480-488.
- 645 Gamble, T., Berendzen, P.B., Bradley Shaffer, H., Starkey, D.E., Simons, A.M., 2008. Species limits and
646 phylogeography of North American cricket frogs (*Acris*: Hylidae). *Molecular phylogenetics and evolution*
647 48, 112-125.
- 648 Glor, R.E., Warren, D., 2011. Testing ecological explanations for biogeographic boundaries. *Evolution* 65,
649 673-683.
- 650 Guiher, T.J., Burbrink, F.T., 2008. Demographic and phylogeographic histories of two venomous North
651 American snakes of the genus *Agkistrodon*. *Molecular Phylogenetics and Evolution* 48, 543-553.
- 652 Harlan, R., 1827. *Genera of North American Reptilia, and synopsis of the species*. J. Acad. Nat. Sci.
653 Philadelphia 5, 317-372.
- 654 Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L., 2006. The effect of sample size and species
655 characteristics on performance of different species distribution modeling methods. *Ecography* 29, 773-
656 785.
- 657 Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405, 907-913.
- 658 Holbrook, J.E., 1836. *North American Herpetology*. J. Dobson, Philadelphia :.
- 659 Homer, C., Huang, C., Yang, L., Wylie, B.K., Coan, M., 2004. Development of a 2001 national land-cover
660 database for the United States.
- 661 Howes, B.J., Lindsay, B., Lougheed, S.C., 2006. Range-wide phylogeography of a temperate lizard, the five-
662 lined skink *Eumeces fasciatus*. *Molecular Phylogenetics and Evolution* 40, 183-194.
- 663 Huelsenbeck, J.P., Andolfatto, P., Huelsenbeck, E.T., 2011. Structurama: Bayesian inference of population
664 structure. *Evolutionary bioinformatics online* 7, 55.
- 665 Jackson, N.D., Austin, C.C., 2010. The combined effects of rivers and refugia generate extreme cryptic
666 fragmentation within the common ground skink (*Scincella lateralis*). *Evolution* 64, 409-428.
- 667 James, C.W., 1961. Endemism in Florida. *Brittonia* 13, 225-244.
- 668 Jones, G., 2016. Algorithmic improvements to species delimitation and phylogeny estimation under the
669 multispecies coalescent. *Journal of Mathematical Biology*, 1-21.
- 670 Jones, G., Aydin, Z., Oxelman, B., 2015. DISSECT: an assignment-free Bayesian discovery method for
671 species delimitation under the multispecies coalescent. *Bioinformatics* 31, 991-998.
- 672 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A.,
673 Markowitz, S., Duran, C., 2012. Geneious Basic: an integrated and extendable desktop software platform
674 for the organization and analysis of sequence data. *Bioinformatics* 28, 1647-1649.
- 675

- 676 Klicka, J., Zink, R.M., 1997. The Importance of Recent Ice Ages in Speciation: A Failed Paradigm. *Science*
677 277, 1666-1669.
- 678 Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J., Scott, J.K., 2012. CliMond: global
679 high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in*
680 *Ecology and Evolution* 3, 53-64.
- 681 Krysko, K.L., Hurt, C., 1998. *Lampropeltis calligaster occipitolineata* (South Florida Mole
682 Kingsnake). *Geographic Distribution. Herpetological Review* 29, 177.
- 683 Krysko, K.L., Krysko, L.E., Hurt, C., 2000. Reproduction and distribution of the South Florida mole
684 kingsnake (*Lampropeltis calligaster occipitolineata*) from central peninsular Florida. *J. Elisha Mitchell Sci.*
685 *Soc* 116, 344-347.
- 686 Krysko, K.L., Nuñez, L.P., Lippi, C.A., Smith, D.J., Granatosky, M.C., 2016. Pliocene–Pleistocene lineage
687 diversifications in the Eastern Indigo Snake (*Drymarchon couperi*) in the Southeastern United States.
688 *Molecular phylogenetics and evolution* 98, 111-122.
- 689 Kubatko, L.S., Gibbs, H.L., Bloomquist, E.W., 2011. Inferring Species-Level Phylogenies and Taxonomic
690 Distinctiveness Using Multilocus Data in *Sistrurus* Rattlesnakes. *Systematic Biology* 60, 393-409.
- 691 Layne, J.N., T. J. Walsh, Meylan, P., 1986. New Records For The Mole Snake, *Lampropeltis calligaster* in
692 Peninsular Florida. *Florida Scientist* 49, 171.
- 693 Leaché, A.D., Reeder, T.W., 2002. Molecular systematics of the eastern fence lizard (*Sceloporus*
694 *undulatus*): a comparison of parsimony, likelihood, and Bayesian approaches. *Systematic Biology* 51, 44-
695 68.
- 696 Legendre, P., Fortin, M.-J., 2010. Comparison of the Mantel test and alternative approaches for detecting
697 complex multivariate relationships in the spatial analysis of genetic data. *Molecular Ecology Resources*
698 10, 831-844.
- 699 Lemmon, E.M., Lemmon, A.R., Cannatella, D.C., 2007. Geological and climatic forces driving speciation in
700 the continentally distributed trilling chorus frogs (*Pseudacris*). *Evolution* 61, 2086-2103.
- 701 Malaney, J.L., Cook, J.A., 2013. Using biogeographical history to inform conservation: the case of Preble's
702 meadow jumping mouse. *Mol Ecol* 22, 6000-6017.
- 703 McKelvy, A.D., Ozelski-McKelvy, A., Figueroa, A., 2016. A new non-coastal record for the Pine Woods
704 Littersnake, *Rhadinaea flavilata* Cope, 1871 (Squamata: Colubridae), in Russell County, Alabama, USA.
705 *Check List* 12, 1913.
- 706 Means, D.B., 1992. Rare: Mole Snake, *Lampropeltis calligaster rhombomaculata*. In: Moler, P. (Ed.), *Rare*
707 *and Endangered Biota of Florida*, Vol. III. Univ. Press Fl, Gainesville, pp. 227-231.
- 708 Mills, L.S., Allendorf, F.W., 1996. The one-migrant-per-generation rule in conservation and management.
709 *Conservation Biology* 10, 1509-1518.
- 710 Myers, E.A., Rodríguez-Robles, J.A., DeNardo, D.F., Staub, R.E., Stropoli, A., Ruane, S., Burbrink, F.T.,
711 2013a. Speciation in the California Mountain Kingsnake (*Lampropeltis zonata*) at a North American
712 Biodiversity Hotspot. *Mol Ecol* 22, 5418–5429.
- 713 Myers, E.A., Weaver, R.E., Alamillo, H., 2013b. Population Stability of the Northern Desert Nightsnake
714 (*Hypsiglena chlorophaea deserticola*) during the Pleistocene. *Journal of Herpetology* 47, 432-439.
- 715 Niemiller, M.L., Graening, G.O., Fenolio, D.B., Godwin, J.C., Cooley, J.R., Pearson, W.D., Fitzpatrick, B.M.,
716 Near, T.J., 2013. Doomed before they are described? The need for conservation assessments of cryptic
717 species complexes using an amblyopsid cavefish (Amblyopsidae: Typhlichthys) as a case study. *Biodivers*
718 *Conserv* 22, 1799-1820.
- 719 Noonan, B.P., Chippindale, P.T., 2006. Vicariant origin of Malagasy reptiles supports Late Cretaceous
720 Antarctic land bridge. *The American Naturalist* 168, 730-741.
- 721 Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., Suggests, M., 2007. The
722 vegan package. *Community ecology package*.

- 723 Olave, M., Solà, E., Knowles, L.L., 2014. Upstream analyses create problems with DNA-based species
724 delimitation. *Systematic biology* 63, 263-271.
- 725 Omernik, J.M., 1987. Ecoregions of the conterminous United States. *Annals of the Association of*
726 *American geographers* 77, 118-125.
- 727 Parmesan, C., 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of*
728 *Ecology, Evolution, and Systematics*, 637-669.
- 729 Phillips, S.J., Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a
730 comprehensive evaluation. *Ecography* 31, 161-175.
- 731 Price, R., 1987. Disjunct occurrence of mole snakes in peninsular Florida, and the description of a new
732 subspecies of *Lampropeltis calligaster*
- 733 *Bulletin of the Chicago Herpetological Society* 22, 148.
- 734 Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus
735 genotype data. *Genetics* 155, 945-959.
- 736 Pyron, R.A., Burbrink, F.T., 2009. Systematics of the Common Kingsnake (*Lampropeltis getula*; Serpentes:
737 *Colubridae*) and the burden of heritage in taxonomy. *Zootaxa* 2241, 22-32.
- 738 Pyron, R.A., Burbrink, F.T., 2010. Hard and soft allopatry: physically and ecologically mediated modes of
739 geographic speciation. *Journal of Biogeography* 37, 2005-2015.
- 740 Quintero, I., Wiens, J.J., 2013. Rates of projected climate change dramatically exceed past rates of
741 climatic niche evolution among vertebrate species. *Ecology letters* 16, 1095-1103.
- 742 Rambaut, A., Drummond, A., 2007. Tracer v1.4. Accessible at <http://tree.bio.ed.ac.uk/software/tracer/>
- 743 Reece, J.S., Noss, R.F., Oetting, J., Hctor, T., Volk, M., 2013. A Vulnerability Assessment of 300 Species in
744 Florida: Threats from Sea Level Rise, Land Use, and Climate Change. *PloS one* 8, e80658.
- 745 Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things).
746 *Methods in Ecology and Evolution* 3, 217-223.
- 747 Reynolds, J.E., 2001. Urbanization and land use change in Florida and the South. *Current Issues*
748 *Associated with Land Values and Land Use Planning Proceedings of a Regional Workshop*, p. 28.
- 749 Rittmeyer, E.N., Austin, C.C., 2012. The effects of sampling on delimiting species from multi-locus
750 sequence data. *Molecular Phylogenetics and Evolution* 65, 451-463.
- 751 Robison, H.W., 1986. Zoogeographic implications of the Mississippi River basin. In: Hocutt, C.H., Wiley,
752 E.O. (Eds.), *The zoogeography of North American freshwater fishes*. John Wiley and Sons, New York, NY,
753 pp. 267-285.
- 754 Royall, P.D., Delcourt, P.A., Delcourt, H.R., 1991. Late Quaternary paleoecology and paleoenvironments of
755 the central Mississippi alluvial valley. *Geological Society of America Bulletin* 103, 157-170.
- 756 Ruane, S., Bryson, R.W., Pyron, R.A., Burbrink, F.T., 2014. Coalescent Species Delimitation in Milksnakes
757 (genus *Lampropeltis*) and Impacts on Phylogenetic Comparative Analyses. *Systematic Biology*.
- 758 Rushton, E., 2013. *Lampropeltis calligaster occipitolineata* (South Florida Mole Kingsnake). *Geographic*
759 *Distribution*. *Herpetological Review* 44, 476.
- 760 Schluter, D., 2001. Ecology and the origin of species. *Trends in Ecology & Evolution* 16, 372-380.
- 761 Schmidt, K.P., Davis, D.D., 1941. *Field book of snakes of the United States and Canada*.
- 762 Sethuraman, A., Hey, J., 2016. IMa2p - parallel MCMC and inference of ancient demography under the
763 Isolation with migration (IM) model. *Molecular Ecology Resources* 16, 206-215.
- 764 Smith, L.M., 1996. Fluvial geomorphic features of the Lower Mississippi alluvial valley. *Engineering*
765 *geology* 45, 139-165.
- 766 Soltis, D.E., Morris, A.B., McLachlan, J.S., Manos, P.S., Soltis, P.S., 2006. Comparative phylogeography of
767 unglaciated eastern North America. *Molecular Ecology* 15, 4261-4293.
- 768 Stephens, M., Donnelly, P., 2003. A comparison of bayesian methods for haplotype reconstruction from
769 population genotype data. *The American Journal of Human Genetics* 73, 1162-1169.

- Susan, S., 2007. Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC. Cambridge University Press.
- Team, R.C., 2000. R Language Definition. Accessible at <http://cran.r-project.org>
- Thioulouse, J., Chessel, D., Dole, S., Olivier, J.-M., 1997. ADE-4: a multivariate analysis and graphical display software. *Statistics and computing* 7, 75-83.
- Thuiller, W., Georges, D., Engler, R., 2013. biomod2: Ensemble platform for species distribution modeling. R package version 2, r560.
- Townsend, T.M., Alegre, R.E., Kelley, S.T., Wiens, J.J., Reeder, T.W., 2008. Rapid development of multiple nuclear loci for phylogenetic analysis using genomic resources: an example from squamate reptiles. *Molecular phylogenetics and evolution* 47, 129-142.
- Walker, D., Avise, J.C., 1998. Principles of phylogeography as illustrated by freshwater and terrestrial turtles in the southeastern United States. *Annual Review of Ecology and Systematics*, 23-58.
- Warren, D.L., Glor, R.E., Turelli, M., 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* 33, 607-611.
- Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H., Guisan, A., Group, N.P.S.D.W., 2008. Effects of sample size on the performance of species distribution models. *Diversity and Distributions* 14, 763-773.
- Yang, Z., Rannala, B., 2010. Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Sciences* 107, 9264-9269.
- Yang, Z., Rannala, B., 2014. Unguided Species Delimitation Using DNA Sequence Data from Multiple Loci. *Molecular Biology and Evolution* 31, 3125-3135.
- Zehner, Z.E., Paterson, B.M., 1983. Characterization of the chicken vimentin gene: single copy gene producing multiple mRNAs. *Proceedings of the National Academy of Sciences of the United States of America* 80, 911.

Figure Legends

Fig. 1 – (A) Approximate range of each species of Yellowbellied Kingsnake and sampling distribution for the Prairie Kingsnake, *Lampropeltis calligaster* (Brown, B), Mole Kingsnake, *Lampropeltis rhombomaculata*, (Green, C) and South Florida Mole Kingsnake *Lampropeltis occipitolineata* (Purple, D) and dated species tree. Ranges are estimated based on Ernst and Ernst (2003) and Conant and Collins (1998) as well as environmental niche models produced in this study. South Florida species *L. occipitolineata* diverged from the Eastern clade *L. rhombomaculata* ~0.88 MYA, posterior probability 0.797. An older split between the western group *L. calligaster* from the two eastern groups occurred ~1.56 MYA, . Posterior probability values are displayed to the left of each node. Photograph B courtesy of Donald Shepard, C Kenneth Krysko and D Kevin Enge.

Fig. 2 – Admixture map and phylogeographic relationships among individuals. Individual gene trees are reported in Supplementary Material 1.

Fig. 3 – Proportion of habitat type recovered. GAP land-cover analysis revealed that western *L. calligaster* were sampled primarily in prairie, pasture and shrublands (61.4%, n=86) and secondarily (37.9%, n=53) in mesic forests. *Lampropeltis rhombomaculata* was sampled primarily in temperate oak and pine forests (64.1%, n=41), and secondarily in pasture land (32.8%, n=21). *L. occipitolineata* was sampled primarily in freshwater wet meadow (37.5%, n=3) and pastureland (37.5%, n=3; combined) and secondarily in southeastern warm temperate forest (25%, n=2).

Fig. 4 – Current and future range predictions based on year 2100 niche models for each recovered lineage within the *L. calligaster* complex showing increase in potential niche range for *L.*

829 *calligaster* and decrease in range for *L. rhombomaculata* and *L. occipitolineata*. Warm colors
830 indicate suitable areas for each species, while unsuitable areas are marked with cool colors.
831 Localities used for each environmental niche model are marked on the map by dots, with colors
832 corresponding to those used in Fig. 1.

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848 Tables

849 **Table 1** - Parameters for each locus used in the analysis. Gene length is reported in base pairs
 850 (bp). Number of variable sites, recommended annealing temperatures and best fit model of
 851 molecular evolution are reported.

Gene	bp	Variable Sites	Annealing Temperature	Substitution Model	Reference
Cytb	1045	78	43°C	HKY + G	(Burbrink et al., 2000; Ruane et al., 2014)
CLLAT	569	4	45.5°C	HKY	(Ruane et al., 2014)
NT3	470	7	49°C	HKY	(Noonan and Chippindale, 2006)
PRLR	209	7	54°C	HKY + G	(Townsend et al., 2008)
VIM56	401	2	49°C	HKY	(Zehner and Paterson, 1983)

852
 853 **Table 2** - Genealogical Sorting Index (GSI) and supporting p-values (in parentheses) for detected
 854 lineages. Individual gene trees were compared to the species tree.

	East	West	Florida
Cytb	1 (0.0001)	1 (0.0001)	1 (0.0139)
CLLAT	0.8056 (0.0001)	0.5625 (0.0001)	0.1135 (0.2522)
NT3	0.6378 (0.0001)	0.5422 (0.0001)	0.1878 (0.0519)
PRLR	0.4443 (0.0001)	0.8709 (0.0001)	0.702 (.0001)
VIM56	0.2495 (0.0001)	0.2198 (0.0109)	0.2525 (0.0042)

855
 856 **Table 3** - Model selection results from the program Migrate. Bezier approximation scores, natural
 857 log Bayes factors (LBF), model probability and the rank of each model are reported. A model of
 858 equal migration between both populations had the most support.

Migration Model	Bezier	LBF	Model Probability	Rank
Equal	-4924.32	0	1.0000000	1
West to East	-4960.04	-35.72	0.0000000	2
East to West	-4960.21	-35.89	0.0000000	3
Panmixia	-5145.41	-221.09	0.0000000	4

859

860 **Table 4** – Number of cells with suitable niche space as modeled under current and future climate

861 scenarios.

	Current	Future	Intersection
<i>Lampropeltis calligaster</i>	392	567	388
<i>Lampropeltis rhombomaculata</i>	374	43	26
<i>Lampropeltis occipitolineata</i>	147	0	0

862

863

Figure 1

Figure 2

Figure 3

Figure 4