Exploring Snake Evolution Through Time and Space

Leonard N. Jones II

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2020

Reading Committee:

Adam Leaché, Chair

Joe Felsenstein

Richard Olmstead

Program Authorized to Offer Degree:

Biology

University of Washington

Abstract

Exploring snake evolution through time and space

Leonard N. Jones II

Chair of the Supervisory Committee:
Adam D. Leaché
Biology

Lineages diversify as they disperse through time and space, which provides ever increasing complexity to be considered in investigations at different taxonomic levels. Population genetic and phylogenetic signal sometimes correspond or are driven by spatial features or historic vicariant events, from which one might reasonably conclude that similarly distributed species might exhibit similar patterns. In chapter 1, I address shared and divergent population patterns in the first comparative population genetic study of the garter snake (*Thamnophis*) assemblage from the San Juan Archipelago, which has likely only been recolonized post-Pleistocene glacial retreat in the last 10,000 years. In chapter 2, I characterized the rangewide phylogeography of the widest ranging of these three species featured on the San Juan Archipelago, *T. sirtalis*, which is distributed throughout North America. To characterize the influence of long established and nascent geographic barriers that characterize the North American continent, I conduct ecological niche modeling on the inferred genetic lineages, and comment on the tempo and mode of a species that has rapidly spread across an entire continent. In Chapter 3, I expand the coalescent methodology to trans-continental evolutionary histories, where I reevaluate the evolutionary relationships of the cobras and closely related elapids (family Elapidae) from targeted sequence capture of UCEs and SNPs derived from ddRADseq data. I assess the stability of divergence times from prior mtDNA and concatenation-based studies by estimating clade ages under the multispecies coalescent model. Returning to the shallow depths of chapters 1 and 2, I evaluate the age and topology of the Asiatic *Naja* radiation and the *Naja melanoleuca* species complex, and weigh the two markers' ability to resolve shallow radiations under the coalescent model.

ACKNOWLEDGMENTS

I wish to acknowledge the San Juan County community who granted me access to their properties for specimen collection. Chapter 1 benefited from the collaboration of the San Juan County Land Bank, San Juan Preservation Trust, and SeaDoc Society for access to preserved habitat, and the Springer family and Fred Richardson for granting me housing and permission to sample their properties on Shaw and Waldron Islands. Sharon Birks and Kevin Epperly processed specimen deposition and sample preparation at the Burke Museum. Joe Felsenstein and Richard Olmstead provided comments that greatly improved the manuscript. This work was supported by funding from the National Science Foundation (DEB 1600844).

I thank all the collections managers and assistants in museums who prepared so many specimens for these three projects including but not limited to: Sharon Birks and Kevin Epperly (UWBM), Luke Welton (KU), Jens Vindum (CAS), Mary Sollows (NBM), David Kizirian (AMNH)**,** Eli Greenbaum (UTEP)**,** Travis LaDuc (TNHC), Donna Dittmann (LSUMZ), Gregory Watkins-Colwell (YPM). The writing guidance of Joe Felsenstein and Richard Olmstead provided comments that drastically reshape these texts for publication to the greater research community. Chapters 1 and 2 were supported by grants from the National Science Foundation (DEB #1600844) and the American Museum of Natural History (Theodore Roosevelt Award). Kevin Epperly's wet lab expertise drove successful data collection throughout my graduate research career. Finally, I thank the Leaché lab and all visiting scholars who have passed through it for engaging discussion, collaboration, and perspective.

DEDICATION

For Halston N. Jones

Chapter 1: A recently recolonized archipelago harbors discordant population structure between sympatric garter snake species (Serpentes: *Thamnophis*)

Leonard N. Jones II[1] & Adam D. Leaché[1]

[1]Department of Biology & Burke Museum of Natural History and Culture, University of Washington, Seattle, WA, USA 98195

INTRODUCTION

Founder populations in recently established habitats offer narrow spatial and temporal windows in which one can evaluate the relative influence of contemporary processes and biogeographic history on the observed genetic structure. In sympatric founder populations of closely related species, comparison of population genetic structure patterns between them can shed light on how ecological similarities/differences between them influence how they establish in new environments. On one hand, one might expect the populations of sympatric species to share population structure patterns due to similar ecological profiles that drive sympatry (e.g. shared habitat and prey sources) (Manier & Arnold 2005). Alternatively, populations of co-distributed species might exhibit population structure differences because of genetic signatures of distinct phylogeographic histories (Weese et al., 2013), habitat use differences (King & Lawson 2001), or competition (Edgehouse et al., 2014). Therefore, to understand fully what causes population structure in new species assemblages also requires an understanding of the phylogenetic history and ecology underlying each.

The San Juan Archipelago is located between Vancouver Island and Washington State in Pacific Northwest America (**Fig 1**). The product of Paleozoic collision and uplift between continental and coastal volcanic plates, the island system and the surrounding Puget Sound have most recently been dramatically carved by multiple glaciation, recession, erosion and sea level fluctuation events culminating in the complete recession of the Cordilleran Ice sheet from the Puget Sound ~15,000 years ago (Easterbrook 1969, Mosher & Hewitt 2004). Recent work has characterized the post-glacial chronology of recolonization by plants and then large mammals (Leopold et al, 2016), that, along with Holocene warming, facilitated the rise of the mixed habitat features that characterize the archipelago today. Terrestrial ectotherms that inhabit the archipelago likely followed the Holocene-associated habitat formation but the timing of these species' arrivals and population genetic structures within the San Juan Archipelago have not been extensively studied. The frequent glaciation cycles in the region almost certainly precluded Pleistocene habitation by vertebrate ectotherms and the Holocene-arrival (<11 KYA) likely provided a limited time for secondary contact between previously isolated populations. While some populations are certainly products of human introduction (*Rana catsebiana*), determining

the ages of contemporary endemic ectotherm populations add new lines of evidence to the existing body of work investigating the ecological history of the region.

Three species of the garter snake genus *Thamnophis*--the common *T. sirtalis*, the Northwestern *T. ordinoides*, and Western terrestrial *T. elegans*, cohabitate the San Juan Islands and present a unique model for comparing population genetic structure patterns between congeners with drastically different ranges and by proxy biogeographic histories. The common garter snake *T. sirtalis* is named for its cosmopolitan distribution—closely associated with water, it can be found throughout North America only to the exclusion of the Great Plains and Western deserts, and has a wide thermal tolerance, making it the northernmost occurring reptile species in the Western hemisphere (Larsen et al, 1993; Rossman, 1996). While similarly associated with water, the western terrestrial garter snake *T* elegans has a near inverse distribution to *T. sirtalis* throughout the xeric regions of western North America, reaching the western Great Plains on the easternmost part of its range and occurring in fragmented but stable populations throughout the Great Basin, major western United States river drainages, and the majority of the California coast. The Northwestern garter snake *T. ordinoides* has the smallest range of the three species, occurring only in the coastal Pacific Northwest with the Cascade range representing its easternmost extent and is most often associated with forest edges and meadows compared to its more explicitly riparian *T. elegans* and *T. sirtalis* (Rossman 1996). Gregory et al found consistency in these habitat affiliations in island *Thamnophis* assemblages (Gregory 1984), and additionally observed disproportionate abundances between *T. ordinoides* and *T. elegans* at a series of Vancouver study sites, suggesting a either a degree of competitive exclusion of *T. elegans* by *T. ordinoides* or predation of *T. ordinoides* by the larger *T. elegans* (Gregory, 1978, 1984). The populations inhabiting the neighboring San Juan Archipelago have not been studied but are assumed to exhibit similar characteristics. The fragmented landscape of the archipelago however adds an additional variable in whether connectivity or lack thereof between the islands is shaping dispersal patterns and in turn, population structure.

Here we assess phylogenetic history and population structure across the central region of the San Juan Archipelago to ask 1) Are populations structured by islands? 2) How congruent are population structure and phylogeographic histories? 3) To what degree do ecological differences

between species shape population structure? In doing so, we specifically address whether species with marked lower fidelities to water (*T. ordinoides*) exhibit population structure associated with island identity compared to those known to be semi-aquatic in most of their distributions and therefore more likely to be prone to move between islands (*T. sirtalis* and *T. elegans*). Alternatively, competitive exclusion of one species (*T. elegans*) in significant portions of the archipelago might correspond to isolated highly differentiated populations, or one small panmictic population of sparsely distributed individuals. Furthermore, structure could represent staggered founder events from the mainland in each species, which might be reflected in few highly divergent haplotypes that correspond to no spatial pattern shared by all individuals.

MATERIALS & METHODS

2.1 Sampling and data collection

We sampled 6 islands across the San Juan Archipelago and neighboring Island County (Fig 1, Table 1) for *Thamnophis* and isolated gDNA from liver tissues using a NaCl extraction. Prior to library preparation we qualified gDNA via gel electrophoresis and the Qubit dsDNA BR quantification assay (Life Technologies Inc.). Echoing the observations of the Gregory et al., very few *T. elegans* individuals were observed in any of the collections sites in comparison to *T. ordinoides*, leading to the discrepancy in species sample sizes between the two species (N = 59 and 15, for *T. ordinoides* and *T. elegans*, respectively). (Gregory 1984). We detected no immediate spatial pattern underlying *T. elegans* that might suggest regional isolation during field collection.

We generated ddRADseq libraries as outlined by Peterson et al., 2012. We first double-digested 50 -1000 ng DNA with two restriction enzymes (SbfI and MspI, New England Biolabs Inc # R3642, R0106) and ligated sample-specific barcodes to individual samples. Samples ligated to unique barcodes were then pooled together in groups of up to 8 individuals, and pools were size selected for fragments in the range of 415-515 bp with a Pippin Prep system (Sage Science Inc). The resulting size-selected products were amplified using a Phusion High Fidelity Taq polymerase kit (New England Biolabs Inc # M0530) with Illumina primers that introduce unique multiplexing indices to each pool upon amplification. Pools were amplified for 30 cycles at 51

degrees annealing temperature. Final libraries were purified using SeraPure Sera-Mag Speedbeads and fragment length distribution and molarity were calculated with an Agilent 2200 Tapestation. Libraries were sequenced on an Illumina HiSeq 4000 (50-bp single end reads) at UC Berkeley's QB3 facility.

For a subset of individuals representing different island populations we sequenced a fragment of the mitochondrial locus cytochrome B (*cytB*) using primers from previous studies (L14910: 5'-GAC CTG TGA TMT GAA AAA CCA YCG TTG T-3'; H16064: 5'-CTT TGG TTT ACA AGA ACA ATG CTT TA-3', (de Queiroz et al., 2002, (Burbrink et al., 2000). We supplemented these data with previously published *cytB* sequences of additional specimens and other congeners and Natricine species to contextualize the history of this population within *Thamnophis* (Bronikowski, 2000, Janzen et al., 2002, Alfaro & Arnold, 2001, de Queiroz et al., 2002; Table 1). We edited and aligned the resulting dataset with Geneious (Geneious 5.1.7, http://www.geneious.com) under global alignment default parameters (Cost matrix: 65 percent similarity; gap open penalty: 12; gap extension penalty: 3).

2.1 Bioinformatics

We processed Illumina reads with stacks v2.5 (Rochette et al., 2019). Reads were first de-multiplexed and filtered with the 'process radtags' function. We discarded reads with a quality score limit =< 20 or with differences in adapter and barcode sequences. We generated loci and SNP datasets in separate assemblies for each species by aligning reads to the *Thamnophis sirtalis* genome (Accession: GCA001077635.2, Perry et al., 2018) with stacks' internal 'ref align' pipeline. For the downstream analyses we generated datasets for each species under the 'r80' parameter flag-- this setting restricts the final locus catalog to those present in at least 80% of all individuals as per the best stacks practices outlined by Paris et al., 2017. The 'ordered-export' flag was included to ensure the absence of duplicate SNPs resulting from overlapping loci. Pairwise-FST values for island-defined populations in each species assembly were calculated using stacks' 'populations' module.

2.3 Population structure estimation

We estimated population genetic relationships across the San Juan Archipelago using two methods. We first inferred the number of populations, *K*, under Hardy Weinberg equilibrium assumptions using the model-based clustering method STRUCTURE v2.3.4 (Pritchard et al., 2000). We ran these analyses under the 'correlated allele frequencies' model, which assumes reasonably close relationships between the populations. While this approach introduces a risk of overestimating population genetic structure, we find the correlated allele frequency model appropriate here to test whether any distinct structure can be inferred by data from such young populations. We also recognize that distinct populations might not correspond to their island assignments, particularly given the highly variable geographic distance between them (e.g. the relative geographic isolation of Waldron versus the adjacency of Henry to San Juan). We executed 10 analyses across K values from 1 to 10 under the admixture model for 500,000 steps with a burn-in of 50,000. We selected the optimal K applying the Evanno method (Evanno et al., 2005) executed in STRUCTURE HARVESTER (Earl and vonHoldt 2012). We visualized results using CLUMPP (Jakobsson & Rosenberg 2007) and distruct (Rosenberg 2004).

We also estimated population structure using non-parametric principal components analysis (PCA) and the sequential *K*-means clustering method discriminant analysis of principal components (DAPC) implemented in adegenet 2.1.1 (Jombart 2008). For the latter, we ran successive *K*-means over K values 1:10 using adegenet's *find.clusters* function and selected the optimal K using BIC.

2.4 Directional migration rate estimation between island populations

We estimated recent migration rates between island populations using a BA3-SNPs v1.1 (Mussmann et al. 2019) . This method builds off of the Bayesian migration rate estimator BayesAss (Wilson and Rannala 2003) to allow analysis of SNP data under the coalescent model. For each species we chose mixing parameters from a series of short BA3-SNPs preliminary runs using the BA3-SNPs-autotune program. We then ran four replicate BA3-SNPs analyses for 10,000,0000 generations, sampling every 1000 and discarding the first 1,000,000 as burnin. We then verified convergence of the replicate analyses and effective ESS values for each estimate in Tracer 1.7 (Rambaut et al. 2018).

2.5 Species tree estimation

A limitation of our aforementioned migration rate estimation approach is BA3-SNPS' emphasis on identifying *recent* migrants in relatively well differentiated populations, a scenario that might not reflect the San Juan Archipelago (herein, SJA) *Thamnophis* assemblage. To assess deeper phylogenetic history of the species, we estimated phylogenetic relationships between the island populations for each species using the Bayesian coalescent program SNAPP (Bryant et al. 2012), implemented in Beast 2.6 (Bouckaert et al. 2014). For each dataset we sampled backward and forward mutation rates *u* and *v* with initial values = 1 and employed an uninformative 1/x distribution prior for the species tree. We applied a gamma distribution with a mean alpha/beta = 0.004. Five replicate analyses were conducted with random starting seeds for 1,500,000 generations while sampling every 100 to verify convergence between runs. These runs were then combined with the first 25% of each analysis discarded as burnin. Finally, we estimated divergence times by scaling the tree height to time units by dividing all the branch lengths by the human nuclear mutation rate 1.1x10-8 per site per generation, which is shown to be similar to that of *T. sirtalis* in recent estimates from the whole genome (Perry et al. 2018).

RESULTS

3.1 Genomic data

An overview of the summary statistics for each species assembly can be found in Table 1. For the ddRAD data, the most loci were recovered for *T. sirtalis* (24,066) which could be a function of alignment to a *T. sirtalis* reference genome, while similar numbers were recovered for *T. elegans* and *T. ordinoides* (16,483 and 14,156, respectively). However, variant sites occurred in similar proportions in each assembly, accounting for ~0.4% of the sampled sites. The final Thamnophiine cytB dataset consisted of a maximum 1117 bp alignment of 114 individuals with 63.8% identical sites (sample-specific sequence lengths found in Table 1).

3.2 Population structure

Inbreeding coefficient values for all sampled SJA populations are shown in **Table 1**. Estimates are consistent between islands within populations, with levels deviating substantially in *T. elegans.* Comparative population structure estimation results are visualized in **Figure 2**. STRUCTURE inferred K = 2 genetic clusters for both *T. sirtalis* and *T. ordinoides*, and K = 3 for *T. elegans*. Like the pairwise-FST values, these clusters do not conform to any immediately discernable spatial pattern (Fig 2A). BIC ranking of DAPC analyses supported 3 clusters for *T. sirtalis*, and 2 clusters for *T. elegans* and *T. ordinoides*, with single individuals determined to be significantly admixed in STRUCTURE analyses, instead representing distinct clusters in each species. PCA results show a similar pattern in *T. ordinoides* and *T. sirtalis* with all island populations forming a loose cluster to the exclusion of a few divergent individuals. PCA clusters in *T. elegans* supported slightly more pronounced differentiation between Shaw and San Juan populations compared to the other species.

3.3 Species tree estimation

Maximum clade compatibility trees for *Thamnophis elegans*, *T. ordinoides*, and *T. sirtalis* are shown in **Figure 3**. The *T. elegans* analysis was the only one to reach convergence, with support for a sister relationship between the Shaw and San Juan populations to the exclusion of the Orcas population with strong support (Posterior Probability = 1.0). Support values for all nodes within the *T. ordinoides* and *T. sirtalis* species trees are consistently low, as are ESS values for all parameter estimates. The phylogenetic structuring of the *T. elegans* population could reflect either an artifact of sampling bias/size or the high levels of genetic diversity and introgression observed in the species that have led to taxonomic instability in prior studies (Uyeda 2012; Queiroz and Lawson 1994).

4. Discussion

This study represents the first comparative population genetic study of the Pacific Northwest *Thamnophis* species assemblage. We find temporally and spatially contrasting phylogeographic histories within each species but conclude that these histories are not reflected in the contemporary population structure patterns within the San Juan Archipelago. We stress the

importance of weighing model and non-model approaches to determine consensus population structure and call for more robust sampling from both the archipelago system and the mainland around it to more finely characterize both the structure therein and the pre-colonization history of the species that might be dictating the patterns seen in contemporary estimates.

Population structure and geography

Using model-based and non-parametric methods, we detect 2-3 genetic clusters in each species that are not confined to any immediately discernible geographic barriers. Coupled with the similar and low pairwise-FST values observed between the island populations of each species, we can conclude that either the water does not represent a barrier to gene flow in any of the species regardless of affinity for water, or the populations of each species are so young that lineages have not sorted to reflect island isolation. In accounting for the single individuals to clusters in each species, population structure estimates were consistent between methods and species despite the sampling discrepancies between the latter. However, the divergence observed in single individuals within each species suggests that expanded sampling of the archipelago could uncover greater levels of genetic diversity within the region. Furthermore, by focusing our efforts on whether structure exists between islands of the San Juan Archipelago and not the mainland, we cannot comment on the directionality of colonization.

Species tree estimation

Coalescent species tree analyses of SNP data from the SJA populations were largely inconclusive due to the computational intractability of analyzing the *T. sirtalis* and *T. ordinoides* datasets. Analysis of the smaller *T. elegans* SNP dataset however produced a fully resolved species tree supporting a sister relationship between the San Juan and Shaw Island populations to the exclusion of Orcas Island. Further sampling of the disparate *T. elegans* populations west of The Cascades will illuminate the patterns underlying the genetic diversity within the species. Downsampling of the *T. sirtalis* and *T. ordinoides* could further validate that the phylogenetic pattern in *T. elegans* reflects a species specific pattern rather than sampling bias.

Species interactions

Like prior studies of the Vancouver *Thamnophis* species assemblage, *Thamnophis elegans* individuals were found less frequently despite substantial ideal habitat overlap with *T. ordinoides* and *T. sirtalis*, providing evidence for an exclusionary dynamic between *T. elegans* and *T. ordinoides*. An open question we cannot answer with the data presented in this study is whether the San Juan Archipelago *T. elegans* populations are sparsely distributed throughout the archipelago or if there are isolated populations in which they are dominant compared to *T. ordinoides*. In our sampling of the largest islands of the central region, *T. ordinoides* was the most common of the three species, but future investigations focused on the entire archipelago and mainland could characterize the degree of panmixia in *T. elegans* across the archipelago or isolation of sympatric populations due to competitive exclusion.

ACKNOWLEDGMENTS

We are grateful to those of the San Juan County community who granted us access to their properties for specimen collection. We are particularly thankful for the collaboration of the San Juan County Land Bank, San Juan Preservation Trust, and SeaDoc Society for access to preserved habitat, and the Springer family and Fred Richardson for granting us housing and permission to sample their property on Shaw and Waldron Islands. Sharon Birks and Kevin Epperly processed specimen deposition and sample preparation at the Burke Museum. Joe Felsenstein and Richard Olmstead provided comments that greatly improved the manuscript. This work was supported by funding from the National Science Foundation (DEB 1600844).

REFERENCES

Alexander Pyron R., Burbrink F.T. 2009. Lineage diversification in a widespread species: roles for niche divergence and conservatism in the common kingsnake, *Lampropeltis getula*. Mol Ecol. 18:3443–3457.

Alfaro M.E., Arnold S.J. 2001. Molecular Systematics and Evolution of Regina and the Thamnophiine Snakes. Molecular Phylogenetics and Evolution. 21:408–423.

Andrews K.R., Good J.M., Miller M.R., Luikart G., Hohenlohe P.A. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics. 17:81–92.

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

Avise J.C. 2000. Phylogeography: The History and Formation of Species. Harvard University Press.

Avise J.C., Arnold J., Ball R.M., Bermingham E., Lamb T., Neigel J.E., Reeb C.A., Saunders N.C. 1987. Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annu. Rev. Ecol. Syst. 18:489–522.

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

Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.-H., Xie D., Suchard M.A., Rambaut A., Drummond A.J. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLOS Computational Biology. 10:e1003537.

Boulenger G.A. 1896. Catalogue of the Snakes in the British Museum (Natural History)... order of the Trustees.

Bronikowski A.M. 2000. Experimental Evidence for the Adaptive Evolution of Growth Rate in the Garter Snake Thamnophis Elegans. Evolution. 54:1760–1767.

Brunsfeld S.J., Sullivan J., Soltis D.E., Soltis P.S. 2001. Comparative phylogeography of northwestern North America: a synthesis. Special Publication-British Ecological Society. 14:319–340.

Bryant D., Bouckaert R., Felsenstein J., Rosenberg N.A., RoyChoudhury A. 2012. Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. Molecular Biology and Evolution. 29:1917–1932.

Buckley L.B., Jetz W. 2007. Environmental and historical constraints on global patterns of amphibian richness. Proceedings of the Royal Society B: Biological Sciences. 274:1167–1173.

Burbrink F.T. 2002. Phylogeographic analysis of the cornsnake (Elaphe guttata) complex as inferred from maximum likelihood and Bayesian analyses. Molecular Phylogenetics and Evolution. 25:465–476.

Burbrink F.T., Chan Y.L., Myers E.A., Ruane S., Smith B.T., Hickerson M.J. 2016. Asynchronous demographic responses to Pleistocene climate change in Eastern Nearctic vertebrates. Ecology letters. 19:1457–1467.

Burbrink F.T., Lawson R. 2007. How and when did Old World ratsnakes disperse into the New World? Molecular Phylogenetics and Evolution. 43:173–189.

Burbrink F.T., Lawson R., Slowinski J.B. 2000. Mitochondrial DNA Phylogeography of the Polytypic North American Rat Snake (elaphe Obsoleta): A Critique of the Subspecies Concept. Evolution. 54:2107–2118.

Burbrink F.T., Yao H., Ingrasci M., Bryson R.W., Guiher T.J., Ruane S. 2011. Speciation at the Mogollon Rim in the Arizona Mountain Kingsnake (Lampropeltis pyromelana). Molecular Phylogenetics and Evolution. 60:445–454.

Burnham K.P., Anderson D.R. 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research. 33:261–304.

Calsbeek R., Thompson J.N., Richardson J.E. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. Molecular Ecology. 12:1021–1029.

Castoe T.A., Bronikowski A.M., Brodie E.D., Edwards S.V., Pfrender M.E., Shapiro M.D., Pollock D.D., Warren W.C. 2011. A proposal to sequence the genome of a garter snake (Thamnophis sirtalis). Stand in Genomic Sci. 4:257–270.

Castoe T.A., Smith E.N., Brown R.M., Parkinson C.L. 2007. Higher-level phylogeny of Asian and American coralsnakes, their placement within the Elapidae (Squamata), and the systematic affinities of the enigmatic Asian coralsnake Hemibungarus calligaster (Wiegmann, 1834): RELATIONSHIPS OF CORALSNAKES AND HEMIBUNGARUS. Zoological Journal of the Linnean Society. 151:809–831.

Caye K., Deist T.M., Martins H., Michel O., François O. 2016. TESS3: fast inference of spatial population structure and genome scans for selection. Molecular Ecology Resources. 16:540–548.

Ceríaco L.M.P., Marques M.P., Schmitz A., Bauer A.M. 2017. The “Cobra-preta” of São Tomé Island, Gulf of Guinea, is a new species of Naja Laurenti, 1768 (Squamata: Elapidae). Zootaxa. 4324:121.

Chifman J., Kubatko L. 2014. Quartet Inference from SNP Data Under the Coalescent Model. Bioinformatics. 30:3317–3324.

Darriba D., Taboada G.L., Doallo R., Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature methods. 9:772–772.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution. 24:332–340.

Deutsch C.A., Tewksbury J.J., Huey R.B., Sheldon K.S., Ghalambor C.K., Haak D.C., Martin P.R. 2008. Impacts of climate warming on terrestrial ectotherms across latitude. PNAS. 105:6668–6672.

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

Earl D.A., vonHoldt B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genet Resour. 4:359–361.

Easterbrook D.J. 1969. Pleistocene chronology of the Puget lowland and San Juan islands, Washington. Geological Society of America Bulletin. 80:2273–2286.

Eaton D.A., Overcast I. 2016. ipyrad: Interactive assembly and analysis of RADseq datasets. Bioinformatics.

Edgehouse M., Latta L.C., Brodie E.D., Brodie E.D. 2014. Interspecific Aggression and Habitat Partitioning in Garter Snakes. PLoS One. 9.

Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution. 63:1–19.

Edwards S.V., Potter S., Schmitt C.J., Bragg J.G., Moritz C. 2016. Reticulation, divergence, and the phylogeography–phylogenetics continuum. Proc Natl Acad Sci USA. 113:8025–8032.

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology. 14:2611–2620.

Faircloth B.C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics. 32:786–788.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved Elements Anchor Thousands of Genetic Markers Spanning Multiple Evolutionary Timescales. Syst Biol. 61:717–726.

Felsenstein J. 1978. Cases in which Parsimony or Compatibility Methods will be Positively Misleading. Syst Biol. 27:401–410.

Fick S.E., Hijmans R.J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology. 37:4302–4315.

Fitch H.S. 1941. Geographic Variation in Garter Snakes of the Species Thamnophis sirtalis in the Pacific Coast Region of North America. American Midland Naturalist. 26:570.

Fitch H.S. 1981. Thamnophis sirtalis. Catalogue of American Amphibians and Reptiles (CAAR).

Flouri T., Jiao X., Rannala B., Yang Z. 2018. Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. Mol Biol Evol. 35:2585–2593.

Flouri T., Jiao X., Rannala B., Yang Z. 2019. A Bayesian Implementation of the Multispecies Coalescent Model with Introgression for Phylogenomic Analysis. Mol Biol Evol.

Gang W., Peng G., Ning X., JingZhuo Z., Zhuang S. 2011. Naja kaouthia: a new snake record to Guizhou. Guizhou Agricultural Sciences.:173–176.

Gaston K.J. 2003. The structure and dynamics of geographic ranges. Oxford: Oxford University Press.

GBIFAvailable from https://www.gbif.org/.

Gould S.J., Johnston R.F. 1972. Geographic Variation. Annual review of ecology and systematics. 3:457–498.

Gray L.N., Barley A.J., Poe S., Thomson R.C., Nieto-Montes de Oca A., Wang I.J. 2019. Phylogeography of a widespread lizard complex reflects patterns of both geographic and ecological isolation. Mol Ecol. 28:644–657.

Green R.E., Braun E.L., Armstrong J., Earl D., Nguyen N., Hickey G., Vandewege M.W., St. John J.A., Capella-Gutierrez S., Castoe T.A., Kern C., Fujita M.K., Opazo J.C., Jurka J., Kojima K.K., Caballero J., Hubley R.M., Smit A.F., Platt R.N., Lavoie C.A., Ramakodi M.P., Finger J.W., Suh A., Isberg S.R., Miles L., Chong A.Y., Jaratlerdsiri W., Gongora J., Moran C., Iriarte A., McCormack J., Burgess S.C., Edwards S.V., Lyons E., Williams C., Breen M., Howard J.T., Gresham C.R., Peterson D.G., Schmitz J., Pollock D.D., Haussler D., Triplett E.W., Zhang G., Irie N., Jarvis E.D., Brochu C.A., Schmidt C.J., McCarthy F.M., Faircloth B.C., Hoffmann F.G., Glenn T.C., Gabaldon T., Paten B., Ray D.A. 2014. Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. Science. 346:1254449–1254449.

Gregory P.T. 1978. Feeding habits and diet overlap of three species of garter snakes (Thamnophis) on Vancouver Island. Can. J. Zool. 56:1967–1974.

Gregory P.T. 1984. Habitat, diet, and composition of assemblages of garter snakes (Thamnophis) at eight sites on Vancouver Island. Can. J. Zool. 62:2013–2022.

Guo P., Liu Q., Xu Y., Jiang K., Hou M., Ding L., Alexander Pyron R., Burbrink F.T. 2012. Out of Asia: Natricine snakes support the Cenozoic Beringian Dispersal Hypothesis. Molecular Phylogenetics and Evolution. 63:825–833.

Gutenkunst R.N., Hernandez R.D., Williamson S.H., Bustamante C.D. 2009. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. PLoS Genet. 5.

Harris R.S. 2007. Improved pairwise Alignmnet of genomic DNA. .

Harvey M.G., Smith B.T., Glenn T.C., Faircloth B.C., Brumfield R.T. 2016. Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. Syst Biol. 65:910–924.

Head J.J., Holroyd P.A., Hutchison J.H., Ciochon R.L. 2005. First report of snakes (Serpentes) from the Late Middle Eocene Pondaung Formation, Myanmar. Journal of Vertebrate Paleontology. 25:246–250.

Heath T.A., Zwickl D.J., Kim J., Hillis D.M. 2008. Taxon Sampling Affects Inferences of Macroevolutionary Processes from Phylogenetic Trees. Syst Biol. 57:160–166.

Hewitt G. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society. 58:247–276.

Hickerson M.J., Carstens B.C., Cavender-Bares J., Crandall K.A., Graham C.H., Johnson J.B., Rissler L., Victoriano P.F., Yoder A.D. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Molecular Phylogenetics and Evolution. 54:291–301.

Holman J.A., others. 2000. Fossil snakes of North America: origin, evolution, distribution, paleoecology. Indiana University Press.

Huson D.H., Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol. 23:254–267.

Ineich I. 1995. Etat actuel de nos connaissances sur la classification des serpents venimeux. Bulletin de la Société herpétologique de France.:7–24.

Irwin D.E. 2002. Phylogeographic breaks without geographic barriers to gene flow. Evolution. 56:2383–2394.

Jackson N.D., Carstens B.C., Morales A.E., O'Meara B.C. 2017. Species Delimitation with Gene Flow. Syst Biol. 66:799–812.

Jakobsson M., Rosenberg N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 23:1801–1806.

Janzen F.J., Krenz J.G., Haselkorn T.S., Brodie E.D. 2002. Molecular phylogeography of common garter snakes (Thamnophis sirtalis) in western North America: implications for regional historical forces. Mol Ecol. 11:1739–1751.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 24:1403–1405.

Jouganous J., Long W., Ragsdale A.P., Gravel S. 2017. Inferring the Joint Demographic History of Multiple Populations: Beyond the Diffusion Approximation. Genetics. 206:1549–1567.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution. 30:772–780.

King R.B., Lawson R. 2001. Patterns of Population Subdivision and Gene Flow in Three Sympatric Natricine Snakes. Copeia. 2001:602–614.

Kishida T., Go Y., Tatsumoto S., Tatsumi K., Kuraku S., Toda M. 2019. Loss of olfaction in sea snakes provides new perspectives on the aquatic adaptation of amniotes. Proceedings of the Royal Society B. 286:20191828.

Kubatko L.S., Degnan J.H. 2007. Inconsistency of Phylogenetic Estimates from Concatenated Data under Coalescence. Syst Biol. 56:17–24.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2016. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol Biol Evol.:msw260.

Larsen K.W., Gregory P.T., Antoniak R. 1993. Reproductive Ecology of the Common Garter Snake Thamnophis sirtalis at the Northern Limit of Its Range. The American Midland Naturalist. 129:336–345.

Leaché A.D., Chavez A.S., Jones L.N., Grummer J.A., Gottscho A.D., Linkem C.W. 2015. Phylogenomics of Phrynosomatid Lizards: Conflicting Signals from Sequence Capture versus Restriction Site Associated DNA Sequencing. Genome Biology and Evolution. 7:706–719.

Leaché A.D., Zhu T., Rannala B., Yang Z. 2019. The Spectre of Too Many Species. Syst Biol. 68:168–181.

Lee M.S.Y., Sanders K.L., King B., Palci A. 2016. Diversification rates and phenotypic evolution in venomous snakes (Elapidae). Royal Society Open Science. 3:150277.

Lehner B., Verdin K., Jarvis A. 2008. New Global Hydrography Derived From Spaceborne Elevation Data. Eos, Transactions American Geophysical Union. 89:93–94.

Leopold E.B., Dunwiddie P.W., Whitlock C., Nickmann R., Watts W.A. 2016. Postglacial vegetation history of Orcas Island, northwestern Washington. Quaternary Research. 85:380–390.

Lobeck A.K. 1948. Physiographic Provinces of North America; Physiographic Diagram of North America. Geographical Press, Division of CS Hammond & Company.

Longbottom J., Shearer F.M., Devine M., Alcoba G., Chappuis F., Weiss D.J., Ray S.E., Ray N., Warrell D.A., Castañeda R.R. de, Williams D.J., Hay S.I., Pigott D.M. 2018. Vulnerability to snakebite envenoming: a global mapping of hotspots. The Lancet. 392:673–684.

Maddison W.P. 1997. Gene trees in species trees. Systematic biology. 46:523–536.

Manier M.K., Arnold S.J. 2005. Population genetic analysis identifies source–sink dynamics for two sympatric garter snake species (Thamnophis elegans and Thamnophis sirtalis). Molecular Ecology. 14:3965–3976.

Manthey J.D., Campillo L.C., Burns K.J., Moyle R.G. 2016. Comparison of Target-Capture and Restriction-Site Associated DNA Sequencing for Phylogenomics: A Test in Cardinalid Tanagers (Aves, Genus: Piranga). Syst Biol. 65:640–650.

Marth G.T., Czabarka E., Murvai J., Sherry S.T. 2004. The Allele Frequency Spectrum in Genome-Wide Human Variation Data Reveals Signals of Differential Demographic History in Three Large World Populations. Genetics. 166:351–372.

Mayr E. 1963. Animal species and evolution. Animal species and evolution.

McVay J.D., Flores-Villela O., Carstens B. 2015. Diversification of North American natricine snakes. Biol. J. Linn. Soc. 116:1–12.

Minton S.A. 1986. Origins of poisonous snakes: evidence from plasma and venom proteins. Natural toxins—animal, plant and microbial. Clarendon Press, Oxford.:3–21.

Mooi R.D., Wiens J.P., Casper G.S. 2011. Extreme Color Variation within Populations of the Common Gartersnake, Thamnophis sirtalis, in Central North America, with Implications for Subspecies Status. Copeia. 2011:187–200.

Moritz C., Langham G., Kearney M., Krockenberger A., VanDerWal J., Williams S. 2012. Integrating phylogeography and physiology reveals divergence of thermal traits between central and peripheral lineages of tropical rainforest lizards. Philosophical Transactions of the Royal Society B: Biological Sciences. 367:1680–1687.

Mosher D.C., Hewitt A.T. 2004. Late Quaternary deglaciation and sea-level history of eastern Juan de Fuca Strait, Cascadia. Quaternary International. 121:23–39.

Mussmann S.M., Douglas M.R., Chafin T.K., Douglas M.E. 2019. BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data. Methods in Ecology and Evolution. 10:1808–1813.

Myers E.A., Weaver R.E., Alamillo H. 2013. Population Stability of the Northern Desert Nightsnake (Hypsiglena chlorophaea deserticola) during the Pleistocene. Journal of Herpetology. 47:432–439.

Nabhan A.R., Sarkar I.N. 2012. The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. Brief Bioinform. 13:122–134.

Nichols R. 2001. Gene trees and species trees are not the same. Trends in Ecology & Evolution. 16:358–364.

Paris J.R., Stevens J.R., Catchen J.M. 2017. Lost in parameter space: a road map for stacks. Methods in Ecology and Evolution. 8:1360–1373.

Pasquesi G.I.M., Adams R.H., Card D.C., Schield D.R., Corbin A.B., Perry B.W., Reyes-Velasco J., Ruggiero R.P., Vandewege M.W., Shortt J.A., Castoe T.A. 2018. Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. Nature Communications. 9:2774.

Perry B.W., Card D.C., McGlothlin J.W., Pasquesi G.I.M., Adams R.H., Schield D.R., Hales N.R., Corbin A.B., Demuth J.P., Hoffmann F.G., Vandewege M.W., Schott R.K., Bhattacharyya N., Chang B.S.W., Casewell N.R., Whiteley G., Reyes-Velasco J., Mackessy S.P., Gamble T., Storey K.B., Biggar K.K., Passow C.N., Kuo C.-H., McGaugh S.E., Bronikowski A.M., de Koning A.P.J., Edwards S.V., Pfrender M.E., Minx P., Brodie E.D., Brodie E.D., Warren W.C., Castoe T.A. 2018. Molecular Adaptations for Sensing and Securing Prey and Insight into Amniote Genome Diversity from the Garter Snake Genome. Genome Biology and Evolution. 10:2110–2129.

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S., Hoekstra H.E. 2012a. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLOS ONE. 7:e37135.

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S., Hoekstra H.E. 2012b. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS one. 7.

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

Placyk Jr. J.S., Burghardt G.M., Small R.L., King R.B., Casper G.S., Robinson J.W. 2007. Post-glacial recolonization of the Great Lakes region by the common gartersnake (Thamnophis sirtalis) inferred from mtDNA sequences. Molecular Phylogenetics and Evolution. 43:452–467.

Portik D.M., Leaché A.D., Rivera D., Barej M.F., Burger M., Hirschfeld M., Rödel M.-O., Blackburn D.C., Fujita M.K. 2017. Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. Molecular Ecology. 26:5245–5263.

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics. 155:945–959.

Pyron R.A., Burbrink F.T., Colli G.R., de Oca A.N.M., Vitt L.J., Kuczynski C.A., Wiens J.J. 2011. The phylogeny of advanced snakes (Colubroidea), with discovery of a new subfamily and comparison of support methods for likelihood trees. Molecular Phylogenetics and Evolution. 58:329–342.

Pyron R.A., Burbrink F.T., Wiens J.J. 2013a. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. BMC Evolutionary Biology. 13:93.

Pyron R.A., Kandambi H.K.D., Hendry C.R., Pushpamal V., Burbrink F.T., Somaweera R. 2013b. Genus-level phylogeny of snakes reveals the origins of species richness in Sri Lanka. Molecular Phylogenetics and Evolution. 66:969–978.

de Queiroz A., Lawson R., Lemos-Espinal J.A. 2002. Phylogenetic Relationships of North American Garter Snakes (Thamnophis) Based on Four Mitochondrial Genes: How Much DNA Sequence Is Enough? Molecular Phylogenetics and Evolution. 22:315–329.

Queiroz A., Lawson R. 1994. Phylogenetic relationships of the garter snakes based on DNA sequence and allozyme variation. Biological Journal of the Linnean Society. 53:209–229.

Queiroz A.D., Lawson R. 2008. A peninsula as an island: multiple forms of evidence for overwater colonization of Baja California by the gartersnake Thamnophis validus. BJLS. 95:409–424.

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic biology. 67:901.

Ricketts T.H., Dinerstein E., Olson D.M., Eichbaum W., Loucks C.J., Kavanaugh K., Hedao P., Hurley P., DellaSalla D., Abell R., Carney K., Walters S. 1999. Terrestrial Ecoregions of North America: A Conservation Assessment. Island Press.

Riddle B.R. 2016. Comparative phylogeography clarifies the complexity and problems of continental distribution that drove A. R. Wallace to favor islands. Proc Natl Acad Sci USA. 113:7970–7977.

Rochette N.C., Rivera-Colón A.G., Catchen J.M. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Molecular ecology. 28:4737–4754.

Rolland J., Silvestro D., Schluter D., Guisan A., Broennimann O., Salamin N. 2018. The impact of endothermy on the climatic niche evolution and the distribution of vertebrate diversity. Nature Ecology & Evolution. 2:459–464.

Rosenberg N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular ecology notes. 4:137–138.

Rossman D.A. 1996. The Garter Snakes: Evolution and Ecology. University of Oklahoma Press.

Ruane S., Austin C.C. 2017. Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. Molecular Ecology Resources. 17:1003–1008.

Ruthven A.G. 1908. Variations and genetic relationships of the garter-snakes. Govt. print. off.

Sanders K.L., Lee M.S.Y., Leys R., Foster R., Keogh J.S. 2008. Molecular phylogeny and divergence dates for Australasian elapids and sea snakes (hydrophiinae): evidence from seven genes for rapid evolutionary radiations. Journal of Evolutionary Biology. 21:682–695.

Santra V., Wüster W. 2017. Natural History: Naja kaouthia (Monocoled Cobra). Behavior / spitting. Herpetological Review. 48:455–456.

Schoener T.W. 1968. The Anolis Lizards of Bimini: Resource Partitioning in a Complex Fauna. Ecology. 49:704–726.

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

Shepard D.B., Burbrink F.T. 2009. Phylogeographic and demographic effects of Pleistocene climatic fluctuations in a montane salamander, *Plethodon fourchensis*. Mol Ecol. 18:2243–2262.

Slowinski J.B., Keogh J.S. 2000. Phylogenetic Relationships of Elapid Snakes Based on Cytochrome b mtDNA Sequences. Molecular Phylogenetics and Evolution. 15:157–164.

Smith B.T., Harvey M.G., Faircloth B.C., Glenn T.C., Brumfield R.T. 2014. Target Capture and Massively Parallel Sequencing of Ultraconserved Elements for Comparative Studies at Shallow Evolutionary Time Scales. Syst Biol. 63:83–95.

Soltis D.E., Gitzendanner M.A., Strenge D.D., Soltis P.S. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Pl Syst Evol. 206:353–373.

Soltis D.E., Morris A.B., McLACHLAN J.S., Manos P.S., Soltis P.S. 2006a. Comparative phylogeography of unglaciated eastern North America. Mol Ecol. 15:4261–4293.

Soltis D.E., Morris A.B., McLACHLAN J.S., Manos P.S., Soltis P.S. 2006b. Comparative phylogeography of unglaciated eastern North America. Molecular Ecology. 15:4261–4293.

Song S., Liu L., Edwards S.V., Wu S. 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proceedings of the National Academy of Sciences. 109:14942–14947.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.

Stange M., Sánchez-Villagra M.R., Salzburger W., Matschiner M. 2018. Bayesian Divergence-Time Estimation with Genome-Wide Single-Nucleotide Polymorphism Data of Sea Catfishes (Ariidae) Supports Miocene Closure of the Panamanian Isthmus. Syst Biol. 67:681–699.

Streicher J.W., Schulte J.A., Wiens J.J. 2016. How Should Genes and Taxa be Sampled for Phylogenomic Analyses with Missing Data? An Empirical Study in Iguanian Lizards. Syst Biol. 65:128–145.

Suchard M.A., Lemey P., Baele G., Ayres D.L., Drummond A.J., Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus evolution. 4:vey016.

Sukumaran J., Knowles L.L. 2017. Multispecies coalescent delimits structure, not species. PNAS. 114:1607–1612.

Suryamohan K., Krishnankutty S.P., Guillory J., Jevit M., Schröder M.S., Wu M., Kuriakose B., Mathew O.K., Perumal R.C., Koludarov I., others. 2020. The Indian cobra reference genome and transcriptome enables comprehensive identification of venom toxins. Nature Genetics.:1–12.

Swofford D.L. 2003. Phylogenetic analysis using parsimony (* and other methods). .

Syfert M.M., Smith M.J., Coomes D.A. 2013. The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. PLOS ONE. 8:e55158.

Trape J.-F., Chirio L., Broadley D.G., Wüster W. 2009. Phylogeography and systematic revision of the Egyptian cobra (Serpentes: Elapidae: Naja haje) species complex, with the description of a new species from West Africa. Zootaxa. 2236:1–25.

Uetz P., Cherikh S., Shea G., Ineich I., Campbell P.D., DORONIN I.V., Rosado J., Wynn A., Tighe K.A., McDiarmid R. 2019. A global catalog of primary reptile type specimens. Zootaxa. 4695:438–450.

Ullate-Agote A., Milinkovitch M.C., Tzika A.C. 2015. The genome sequence of the corn snake (Pantherophis guttatus), a valuable resource for EvoDevo studies in squamates. International Journal of Developmental Biology. 58:881–888.

Uyeda J.C. Connecting Microevolutionary Processes to Macroevolutionary Patterns Across Space and Time. :306.

Vonk F.J., Casewell N.R., Henkel C.V., Heimberg A.M., Jansen H.J., McCleary R.J.R., Kerkkamp H.M.E., Vos R.A., Guerreiro I., Calvete J.J., Wüster W., Woods A.E., Logan J.M., Harrison R.A., Castoe T.A., Koning A.P.J. de, Pollock D.D., Yandell M., Calderon D., Renjifo C., Currier R.B., Salgado D., Pla D., Sanz L., Hyder A.S., Ribeiro J.M.C., Arntzen J.W., Thillart G.E.E.J.M. van den, Boetzer M., Pirovano W., Dirks R.P., Spaink H.P., Duboule D., McGlinn E., Kini R.M., Richardson M.K. 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. PNAS. 110:20651–20656.

Voorhies M.R. 1990. Vertebrate biostratigraphy of the Ogallala Group in Nebraska. Geologic Framework and Regional Hydrology: Upper Cenozoic Blackwater Draw and Ogallala Forma- tion, Great Plains.:115–151.

Wallach V., Wüster W., Broadley D.G. 2009. In praise of subgenera: taxonomic status of cobras of the genus Naja Laurenti (Serpentes: Elapidae). Zootaxa. 2236:26–36.

Warren D.L., Glor R.E., Turelli M. 2008. Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. Evolution. 62:2868–2883.

Warren D.L., Glor R.E., Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography. 33:607–611.

Warren D.L., Seifert S.N. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecological Applications. 21:335–342.

Weese D.A., Fujita Y., Santos S.R. 2013. Multiple Colonizations Lead to Cryptic Biodiversity in an Island Ecosystem: Comparative Phylogeography of Anchialine Shrimp Species in the Ryukyu Archipelago, Japan. The Biological Bulletin. 225:24–41.

Wiens J.J. 2004. Speciation and Ecology Revisited: Phylogenetic Niche Conservatism and the Origin of Species. Evolution. 58:193–197.

Wiens J.J., Hutter C.R., Mulcahy D.G., Noonan B.P., Townsend T.M., Sites J.W., Reeder T.W. 2012. Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. Biology Letters. 8:1043–1046.

Wilson G.A., Rannala B. 2003. Bayesian Inference of Recent Migration Rates Using Multilocus Genotypes. Genetics. 163:1177–1191.

Wisz M.S., Pottier J., Kissling W.D., Pellissier L., Lenoir J., Damgaard C.F., Dormann C.F., Forchhammer M.C., Grytnes J.-A., Guisan A., Heikkinen R.K., Høye T.T., Kühn I., Luoto M., Maiorano L., Nilsson M.-C., Normand S., Öckinger E., Schmidt N.M., Termansen M., Timmermann A., Wardle D.A., Aastrup P., Svenning J.-C. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biological Reviews. 88:15–30.

Wood D.A., Vandergast A.G., Lemos Espinal J.A., Fisher R.N., Holycross A.T. 2011. Refugial isolation and divergence in the Narrowheaded Gartersnake species complex (Thamnophis rufipunctatus) as revealed by multilocus DNA sequence data: REFUGIAL ISOLATION IN NARROWHEADED GARTERSNAKES. Mol Ecol. 20:3856–3878.

Wüster W., Broadley D.G. 2007. Get an eyeful of this: a new species of giant spitting cobra from eastern and north-eastern Africa (Squamata: Serpentes: Elapidae: Naja). Zootaxa. 1532:51–68.

Wüster W., Chirio L., Trape J.-F., Ineich I., Jackson K., Greenbaum E., Barron C., Kusamba C., Nagy Z.T., Storey R., Hall C., Wüster C.E., Barlow A., Broadley D.G. 2018. Integration of nuclear and mitochondrial gene sequences and morphology reveals unexpected diversity in the forest cobra (Naja melanoleuca) species complex in Central and West Africa (Serpentes: Elapidae). Zootaxa. 4455:68–98.

Wüster W., Crookes S., Ineich I., Mané Y., Pook C.E., Trape J.-F., Broadley D.G. 2007. The phylogeny of cobras inferred from mitochondrial DNA sequences: Evolution of venom spitting and the phylogeography of the African spitting cobras (Serpentes: Elapidae: Naja nigricollis complex). Molecular Phylogenetics and Evolution. 45:437–453.

Wüster W., Thorpe R.S. 1992. Dentitional Phenomena in Cobras Revisited: Spitting and Fang Structure in the Asiatic Species of Naja (Serpentes: Elapidae). Herpetologica. 48:424–434.

Yap M.K.K., Tan N.H., Fung S.Y. 2011. Biochemical and toxinological characterization of Naja sumatrana (Equatorial spitting cobra) venom. Journal of Venomous Animals and Toxins including Tropical Diseases. 17:451–459.

Young B.A. 2004. The buccal buckle: the functional morphology of venom spitting in cobras. Journal of Experimental Biology. 207:3483–3494.

Zaher H., Murphy R.W., Arredondo J.C., Graboski R., Machado-Filho P.R., Mahlow K., Montingelli G.G., Quadros A.B., Orlov N.L., Wilkinson M., Zhang Y.-P., Grazziotin F.G. 2019. Large-scale molecular phylogeny, morphology, divergence-time estimation, and the fossil record of advanced caenophidian snakes (Squamata: Serpentes). PLOS ONE. 14:e0216148.

Zink R.M. 2002. Methods in Comparative Phylogeography, and Their Application to Studying Evolution in the North American Aridlands. Integr Comp Biol. 42:953–959.

Fig 1. Study system: A) (Top – Bottom) *Thamnophis sirtalis*, *T. ordinoides*, and *T. elegans*. B) The species' respective distributions. C) Sampling scheme across the central region of the San Juan Archipelago for this study.

Fig 2. Comparative population structure estimation between *T. sirtalis, T. ordinoides,* and *T. elegans* using A) STRUCTURE. B) Discriminant analysis of principal components (DAPC), and C) PCA. Inferred clusters largely do not correspond to island identity across. Both STRUCTURE and DAPC analyses grouped single individuals that exhibit admixture in STRUCTURE analyses and high differentiation in PCA analyses into distinct clusters for all three species.

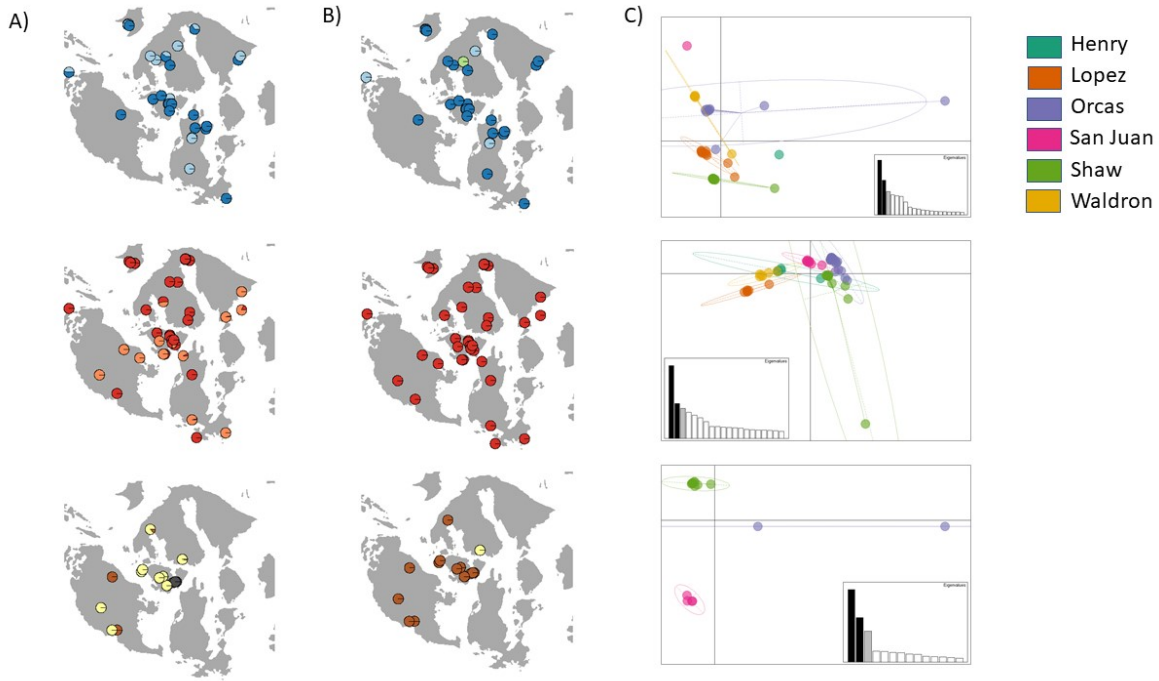


Fig 3. Coalescent estimates of *Thamnophis* island population relationships (A) *T. elegans*, (B) *T. sirtalis* (C) *T. ordinoides*) from SNP data using SNAPP. Node values represent posterior probabilities.

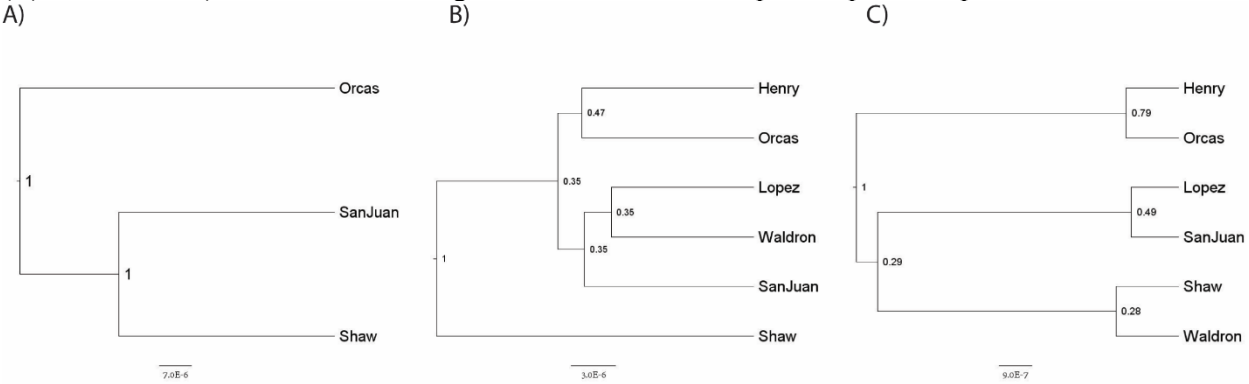


Table 1. Inbreeding coefficient values (parenthetical values = standard error) estimated by BayesAss with values across shared islands highlighted in grey.

	Waldron	**San Juan**	**Shaw**	**Orcas**	**Henry**	**Lopez**
T. sirtalis	0.7576(0.0061)	0.4885(0.2877)	0.7437(0.0068)	0.7016(0.0055)	0.7263(0.0126)	0.7401(0.0056)
T. ordinoides	0.6457(0.0087)	0.6513(0.0082)	0.7503(0.0054)	0.6573(0.0080)	0.6573(0.0080)	0.6132(0.0091)
T. elegans	x	0.3860(0.0173)	0.4800(0.0102)	0.6657(0.0129)	x	x

Table S1. Samples examined in this study.

Institution/Collector ID	Species	Locality	Country	Retained reads (%)	*cytB* length (bp)
LNJ004	*T. sirtalis*	Lopez Island, WA	USA	2437973 (91.29)	
LNJ005	*T. sirtalis*	Lopez Island, WA	USA	2309380 (95.67)	
LNJ006	*T. sirtalis*	Lopez Island, WA	USA	2112839 (79.95)	
LNJ085	*T. ordinoides*	Orcas Island, WA	USA	174881 (85.74)	790
LNJ086	*T. ordinoides*	Orcas Island, WA	USA	767197 (78.15)	
LNJ087	*T. ordinoides*	Orcas Island, WA	USA	987469 (94.23)	
LNJ088	*T. ordinoides*	Orcas Island, WA	USA	48175 (31.94)	
LNJ089	*T. ordinoides*	Orcas Island, WA	USA	2354990 (95.19)	
LNJ090	*T. ordinoides*	Orcas Island, WA	USA	2734938 (94.05)	
LNJ091	*T. sirtalis*	Orcas Island, WA	USA	5184721 (95.97)	
LNJ092	*T. ordinoides*	Shaw Island, WA	USA	3069139 (96.91)	
LNJ093	*T. elegans*	Shaw Island, WA	USA	2535626 (96.27)	
LNJ094	*T. ordinoides*	Orcas Island, WA	USA	173450 (82.4)	
LNJ095	*T. ordinoides*	Orcas Island, WA	USA	4871002 (76.52)	
LNJ096	*T. ordinoides*	Orcas Island, WA	USA	961241 (97.2)	
LNJ097	*T. sirtalis*	Shaw Island, WA	USA	1695307 (96.24)	
LNJ098	*T. elegans*	Shaw Island, WA	USA	3832060 (96.57)	
LNJ099	*T. ordinoides*	Shaw Island, WA	USA	63811 (85.54)	766
LNJ100	*T. sirtalis*	Shaw Island, WA	USA	3780571 (87.41)	
LNJ101	*T. ordinoides*	Shaw Island, WA	USA	7546764 (96.2)	
LNJ102	*T. ordinoides*	Orcas Island, WA	USA	5012076 (90.59)	
LNJ103	*T. ordinoides*	Orcas Island, WA	USA	3635430 (95.11)	
LNJ104	*T. ordinoides*	Orcas Island, WA	USA	1664942 (95.3)	
LNJ105	*T. ordinoides*	Orcas Island, WA	USA	581432 (97.01)	
LNJ106	*T. sirtalis*	Orcas Island, WA	USA	3794236 (81.89)	
LNJ107	*T. ordinoides*	Orcas Island, WA	USA	4987338 (93.91)	
LNJ108	*T. elegans*	Orcas Island, WA	USA	4462196 (95.56)	

LNJ109	*T. ordinoides*	Orcas Island, WA	USA	4288087 (63.35)	
LNJ110	*T. ordinoides*	Orcas Island, WA	USA	11606439 (91.53)	
LNJ111	*T. ordinoides*	Orcas Island, WA	USA	1926328 (95.32)	
LNJ112	*T.ordinoides*	Orcas Island, WA	USA	4900810 (97.27)	
LNJ113	*T. ordinoides*	Shaw Island, WA	USA	3133570 (96.37)	
LNJ114	*T. sirtalis*	Orcas Island, WA	USA	1741201 (96.79)	
LNJ117	*T. sirtalis*	Orcas Island, WA	USA	4383167 (68.1)	780
LNJ118	*T. sirtalis*	Orcas Island, WA	USA	11015692 (96.32)	
LNJ121	*T. elegans*	Orcas Island, WA	USA	1586521 (94.39)	
LNJ122	*T. sirtalis*	Orcas Island, WA	USA	7025170 (92.78)	768
LNJ123	*T. elegans*	San Juan Island, WA	USA	3191287 (95.74)	
LNJ124	*T. ordinoides*	Shaw Island, WA	USA	4489592 (89.36)	769
LNJ125	*T. elegans*	Shaw Island, WA	USA	7363367 (91.62)	
LNJ126	*T. ordinoides*	San Juan Island, WA	USA	3474063 (97.5)	
LNJ127	*T. elegans*	San Juan Island, WA	USA	2803545 (91.55)	
LNJ128	*T. elegans*	San Juan Island, WA	USA	3031970 (73.64)	
UWBM8114	*T. sirtalis*	Henry Island, WA	USA	296904 (92.95)	
UWBM8115	*T. ordinoides*	Henry Island, WA	USA	2968744 (93.53)	
UWBM8701	*T. ordinoides*	Orcas Island, WA	USA	1560505 (92.51)	
UWBM8702	*T. elegans*	San Juan Island, WA	USA	5497315 (95.39)	
UWBM8703	*T. elegans*	Shaw Island, WA	USA	2573064 (93.9)	
UWBM8704	*T. elegans*	Shaw Island, WA	USA	6758239 (84.92)	
UWBM8705	*T. elegans*	Shaw Island, WA	USA	736968 (92.09)	
UWBM8706	*T. elegans*	Shaw Island, WA	USA	5243112 (72.96)	
UWBM8707	*T. elegans*	Shaw Island, WA	USA	693196 (94.99)	
UWBM8708	*T. ordinoides*	Shaw Island, WA	USA	5567003 (94.17)	
UWBM8710	*T. ordinoides*	San Juan Island, WA	USA	1565971 (90.88)	
UWBM8711	*T. ordinoides*	San Juan Island, WA	USA	622648 (84.24)	800
UWBM8713	*T. ordinoides*	Henry Island, WA	USA	410321 (90.52)	790
UWBM8714	*T. ordinoides*	Henry Island, WA	USA	2794229 (79.8)	791
UWBM8715	*T. ordinoides*	Henry Island, WA	USA	1798007 (84.95)	

UWBM8716	*T. ordinoides*	Henry Island, WA	USA	2336489 (91.34)	
UWBM8717	*T. ordinoides*	Henry Island, WA	USA	3382917 (74.38)	
UWBM8718	*T. ordinoides*	Lopez Island, WA	USA	2342313 (93.81)	
UWBM8719	*T. ordinoides*	Lopez Island, WA	USA	3582661 (97.3)	
UWBM8720	*T. ordinoides*	Lopez Island, WA	USA	3707664 (94.71)	
UWBM8721	*T. ordinoides*	Lopez Island, WA	USA	1950005 (93.39)	
UWBM8722	*T. ordinoides*	Lopez Island, WA	USA	2898857 (92.35)	
UWBM8723	*T. ordinoides*	Lopez Island, WA	USA	5571116 (68.23)	785
UWBM8724	*T. ordinoides*	Orcas Island, WA	USA	1595156 (91.36)	
UWBM8725	*T. sirtalis*	Orcas Island, WA	USA	2644495 (93.79)	
UWBM8727	*T. ordinoides*	San Juan Island, WA	USA	1044849 (88.21)	
UWBM8729	*T. ordinoides*	San Juan Island, WA	USA	8313477 (91.97)	
UWBM8731	*T. ordinoides*	San Juan Island, WA	USA	1061519 (82.23)	
UWBM8732	*T. ordinoides*	San Juan Island, WA	USA	3679460 (94.6)	
UWBM8733	*T. ordinoides*	San Juan Island, WA	USA	4376291 (69.49)	803
UWBM8734	*T. ordinoides*	Shaw Island, WA	USA	1959049 (89.67)	
UWBM8735	*T. ordinoides*	Shaw Island, WA	USA	6789647 (93.75)	
UWBM8736	*T. ordinoides*	Shaw Island, WA	USA	1238007 (92.25)	
UWBM8737	*T. ordinoides*	Shaw Island, WA	USA	9704387 (90.12)	
UWBM8738	*T. ordinoides*	Shaw Island, WA	USA	6455497 (88.39)	
UWBM8740	*T. elegans*	Shaw Island, WA	USA	5522533 (92.19)	
UWBM8741	*T. ordinoides*	Shaw Island, WA	USA	145033 (85.17)	
UWBM8742	*T. ordinoides*	Waldron Island, WA	USA	3376943 (81.25)	
UWBM8743	*T. ordinoides*	Waldron Island, WA	USA	4792855 (92.47)	776
UWBM8744	*T. ordinoides*	Waldron Island, WA	USA	2747754 (93.5)	
UWBM8745	*T. ordinoides*	Waldron Island, WA	USA	3541884 (84.26)	
UWBM8746	*T. ordinoides*	Waldron Island, WA	USA	4956635 (97.24)	763
UWBM8747	*T. ordinoides*	Waldron Island, WA	USA	1216262 (91.96)	
UWBM8748	*T. sirtalis*	San Juan Island, WA	USA	1897448 (93.94)	
UWBM8749	*T. sirtalis*	Lopez Island, WA	USA	984016 (74.5)	
UWBM8750	*T. sirtalis*	Lopez Island, WA	USA	466531 (64.32)	

UWBM8751	*T. sirtalis*	Lopez Island, WA	USA	165004 (81.5)	
UWBM8752	*T. sirtalis*	Lopez Island, WA	USA	7005232 (92.19)	
UWBM8753	*T. sirtalis*	Lopez Island, WA	USA	1350382 (72.93)	
UWBM8755	*T. sirtalis*	Lopez Island, WA	USA	1572338 (93.91)	
UWBM8756	*T. sirtalis*	Lopez Island, WA	USA	3325961 (91.32)	818
UWBM8757	*T. sirtalis*	Orcas Island, WA	USA	1477141 (87.95)	
UWBM8758	*T. sirtalis*	Orcas Island, WA	USA	2655068 (94.65)	
UWBM8759	*T. sirtalis*	Shaw Island, WA	USA	3824863 (92.93)	808
UWBM8760	*T. sirtalis*	Shaw Island, WA	USA	720472 (65.14)	
UWBM8761	*T. sirtalis*	Shaw Island, WA	USA	1529381 (91.99)	
UWBM8762	*T. sirtalis*	Shaw Island, WA	USA	145989 (59.44)	
UWBM8763	*T. sirtalis*	Shaw Island, WA	USA	2599072 (96.05)	
UWBM8764	*T. sirtalis*	Shaw Island, WA	USA	1849478 (94.47)	
UWBM8765	*T. sirtalis*	Shaw Island, WA	USA	1872351 (93.5)	
UWBM8766	*T. sirtalis*	Waldron Island, WA	USA	3401729 (92.3)	
UWBM8767	*T. sirtalis*	Waldron Island, WA	USA	2298317 (93.78)	
UWBM8768	*T. sirtalis*	Waldron Island, WA	USA	9262074 (97.31)	
UWBM8769	*T. sirtalis*	Waldron Island, WA	USA	510001 (77.11)	750
UWBM8770	*T. sirtalis*	Waldron Island, WA	USA	8493248 (96.18)	804

Chapter 2: Phylogeography and historical demography of the common gartersnake (*Thamnophis sirtalis*)

Leonard N. Jones II[1], Adam D. Leaché[1], Stevan J. Arnold[2], Frank T. Burbrink[3]

[1]Department of Biology & Burke Museum of Natural History and Culture, University of Washington, Seattle, WA, USA 98195

[2]Department of Integrative Biology, Oregon State University, Corvallis, OR, USA 97331

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

INTRODUCTION

The goal of phylogeography is to characterize how biotic and abiotic landscape processes shape organismal diversity (Mayr 1963; Gould and Johnston 1972; Gaston 2003; Wisz et al. 2013). The complexity of these processes and their influence on contemporary population genetic and phylogenetic patterns can increase with an organism's spatial distribution and age (Avise et al. 1987; Hickerson et al. 2010; Edwards et al. 2016; Riddle 2016). For example, intraspecific divergences caused by ecological adaptation to adjacent biomes, glacial isolation followed by recent secondary contact, or genomic incompatibility might be indistinguishable (Edwards et al. 2016; Riddle 2016). Lineages within transcontinental species may exhibit clinal variation and/or be subject to asynchronously timed geologic and climactic processes (i.e., tectonic events in the Pacific Northwest America versus glacial advance/recession in the Midwest and Northeast). Phylogeographic comparisons between co-distributed taxa can help clarify demographic and evolutionary histories, but pseudocongruence of shared diversification patterns caused by different processes (Irwin 2002; Soltis et al. 2006) can confound identification of dispersal barriers. Therefore, accurate inference of the causes of contemporary diversity in widely distributed species demands incorporation of population structure, phylogeny and divergence time estimation, demography, and ecology.

Subarctic North America comprises multiple ecoregions containing over 150 terrestrial habitat types (Lobeck 1948; Ricketts et al. 1999), and has been subject to numerous phylogeographic investigations due to its variable topography, climate, and glaciation histories (Avise 2000; Brunsfeld et al. 2001; Calsbeek et al. 2003; Shafer et al. 2010; Soltis et al. 1997, 2006; Zink 2002). North American species with widespread distributions often cross multiple longstanding environmental gradients and geologic barriers, resulting in local adaptation and lineage diversification. In the Eastern United States, considerable attention has been given to the divisive effects of the Mississippi and Apalachicola-Chattahoochee-Flint drainage systems and Appalachian Mountains on species distributions (Soltis et al. 2006). In the West and Pacific Northwest, the population-isolating effects of Miocene mountain chain uplifts and multiple Pleistocene glacial cycles have been implicated in demographic shifts leaving genetic signatures in floral and faunal assemblages (Soltis et al. 1997; Shafer et al. 2010). Terrestrial ectotherms are

key systems for investigations of historical demography and phylogeography due to the physiological constraints that keep their distributions tightly linked to their optimal environment (Hewitt 1996; Wiens 2004; Deutsch et al. 2008; Moritz et al. 2012) and generally low dispersal patterns compared to mammalian and avian systems (Buckley and Jetz 2007; Rolland et al. 2018). Wide ranging terrestrial ectotherms to an extent violate this paradigm, and therefore can be particularly interesting for testing hypotheses concerning dispersal, vicariance, and local adaptation due to the biotic diversity in habitats their populations occupy.

The common garter snake *Thamnophis sirtalis* is the most widely distributed reptile in North America and one of the most well-studied snakes in the world (Boulenger 1896; Fitch 1941, 1981; Rossman 1996; Alfaro and Arnold 2001; Castoe et al. 2011). A consummate riparian generalist, its range (**Fig 1**) spans a rich diversity of North American habitats across phylogeographic breaks understood to impact population structure within a variety of disparately related taxa (Avise 2000; Brunsfeld et al. 2001; Soltis et al. 2006). *Thamnophis sirtalis* is perhaps most well-known for its extensive regional phenotypic variation--13 recognized subspecies defined by dorsolateral color patterning are currently recognized throughout its range (Janzen et al. 2002; Mooi et al. 2011; Rossman 1996; Uetz et al. 2019). Life history traits can vary widely between the populations occupying the extreme latitudinal extents of its range, making *T. sirtalis* a candidate system for testing hypotheses linking landscape effects and intraspecific regional adaptation. Attempts to contextualize these population differences in an evolutionary/taxonomic framework are not unlike prior investigations on its parent genus, which "…has long stood in the minds of herpetologists as a synonym for chaos" (Ruthven 1908). Prior mtDNA studies of West and Pacific Northwest populations (Janzen et al. 2002) and in the Great Lakes region (Placyk Jr. et al. 2007) suggest that vicariant history and dispersal patterns from multiple Pleistocene refugia, rather than shared ancestry, explain the observed variation. However, a range-wide study of *T. sirtalis* has not been conducted, leaving critical populations at the extents of its distribution (Eastern and Southeastern U.S., Eastern Canada) unevaluated. Investigations of other terrestrial ectotherms in these regions found cryptic signatures of intraspecific divergence and demographic shifts (Burbrink et al. 2000, 2016; Burbrink 2002; Shepard and Burbrink 2009), but none of these species exhibit the cosmopolitan distribution of *T. sirtalis*.

Here we present the first phylogeography of *T. sirtalis* encompassing its entire distribution. From mtDNA and SNP data we identified four distinct lineages and infer divergence timing, population structure, and degree and directionality of contemporary gene flow between them. By contextualizing the demographic histories of *T. sirtalis* lineages across its entire distribution, we characterize its origin and dispersal across North America. Specifically, we test whether *T. sirtalis* lineages currently occupying historically glaciated regions exhibit the signatures of post-Pleistocene population expansion as observed in other snake systems (Burbrink and Lawson 2007; Burbrink et al. 2011; Myers et al. 2013), but untested in a cold-adapted ectotherm that may have persisted in Northern and Southern refugia. Due to the fidelity of *T. sirtalis* to bodies of water throughout its range, we generate species distribution models to evaluate the effect of drainage systems and ancient rivers that represent dispersal barriers in other taxa on the dispersal of *T. sirtalis* lineages. Finally, we compare phylogenetic and population genetic patterns to those found in phylogeographic investigations of co-distributed ectotherms and comment on implications for taxonomy and conservation.

MATERIALS & METHODS

2.1 Sampling and data collection

We assembled tissues from localities spanning the entire *T. sirtalis* distribution (**Fig 1**) from field and museum collections (**Table 1**) and isolated genomic DNA using NaCl extraction. We assessed extracted DNA quality via gel electrophoresis and quantity with a Qubit dsDNA BR assay (Life Technologies Inc.) prior to library preparation.

We generated ddRADseq libraries for 212 samples as outlined by Peterson et al. 2012. We first digested 500-1000 ng DNA with two restriction enzymes (*SbfI* and *MspI*, New England Biolabs Inc # R3642, R0106) and ligated sample-specific barcodes to individual samples. Samples ligated to unique barcodes were then pooled together in groups of up to 8, and pools were size selected for fragments in the range of 415-515 bp with a Pippin Prep system (Sage Science Inc). The resulting size-selected products were amplified using a Phusion High Fidelity Taq polymerase kit (New England Biolabs Inc # M0530) with Illumina primers that introduce unique multiplexing indices to each pool upon amplification. Pools were amplified for 30 cycles at 51º

C annealing temperature. Final libraries were purified using Ampure XP beads and fragment length distribution and molarity were calculated with an Agilent 2200 Tapestation. Libraries were sequenced across 3 Illumina HiSeq 4000 lanes (50-bp single end reads) at UC Berkeley QB3 facility.

Additionally, for a subset of individuals we sequenced a fragment mitochondrial locus cytochrome B (*cytB*) using primers from previous studies (L14910: 5'-GAC CTG TGA TMT GAA AAA CCA YCG TTG T-3'; H16064: 5'-CTT TGG TTT ACA AGA ACA ATG CTT TA-3', (Burbrink et al. 2000; de Queiroz et al. 2002). We supplemented these data with previously published *cytB* sequences for *T. sirtalis* from additional localities and 18 additional congeners and other natricine species (Alfaro and Arnold 2001; de Queiroz et al. 2002; Queiroz and Lawson 2008) for more precise phylogenetic inference (**Table 2**). We edited and aligned the resulting dataset with Geneious 5.1.7 (Geneious 5.1.7, http://www.geneious.com) under global alignment default parameters (Cost matrix: 65% similarity; gap open penalty: 12; gap extension penalty: 3).

2.2 ddRADseq Bioinformatics

We processed raw Illumina reads using stacks v2.5 (Rochette et al. 2019) 'ref_align' pipeline against a recently annotated *Thamnophis sirtalis* genome (Accession: GCA_001077635.2 (Perry et al. 2018)). We de-multiplexed and filtered reads with the 'process_radtags' function and imposed a strict quality score, discarding reads with a quality score limit < 20 or with a single base pair difference in adapter and barcode sequences. We filtered the remaining reads with the 'populations' function. We retained loci present in 80% of the individuals of a given putative population, following the *r80* method described in prior studies of the impact of bioinformatic practices on stacks analyses (Paris et al. 2017). For the final dataset used for the population structure analyses, we restricted SNP collection to a single SNP per locus and the 'ordered-export' flag was used to ensure the absence of duplicate SNPs resulting from overlapping loci.

2.3 Population structure estimation

We estimated population genetic relationships using four methods. First, we identified the number of roughly detectable clusters directly from the data (i.e., with no consideration of

biological context) using the multivariate methods principal components analysis (PCA) and discriminant analysis of principal components (DAPC), both of which are implemented in adegenet 2.1.1 (Jombart 2008) for analyses of unlinked SNPs. PCA summarizes between-individual variation and requires no *a priori* group assignment. DAPC partitions individuals into clusters that maximize group variation while minimizing within group variation via a K-means clustering search. A sample of strongly genetically differentiated groups might exhibit the same pattern using both methods, but in potentially heavily admixed populations, an exploratory evaluation of intra and inter-group variation might better inform analyses under more biologically relevant models. We conducted the PCA analysis with the *dudi.pca* function, retaining the 3 axes accounting for most of the global variation in the SNP data and plotted the resulting factorial maps with the *s.class* function. For the DAPC, we used the *find.clusters* function to determine from all retained principal components the most suitable number of clusters K up to 10, and selected the optimal value based off of the lowest Bayesian Information Criterion (BIC) score.

Results from the above analyses cannot speak to admixture patterns shaped by biological processes and are limited by subjectivity of the a priori assumptions required in DAPC, so we incorporated two model-based methods to further assess the range-wide structure in a biological context. We inferred the number of populations K under Hardy Weinberg assumptions using the Bayesian clustering method STRUCTURE v2.3.4 (Pritchard et al. 2000). We ran 10 analyses across K values 1-10 under the admixture model for 500,000 steps with a burn-in of 50,000, and selected optimal K using Evanno's ΔK method (Evanno et al. 2005). We generated mean cluster membership coefficient matrices from these replicate analyses for the optimal K values with CLUMPP (Jakobsson and Rosenberg 2007). This method maximizes the similarity between potentially unordered cluster membership coefficient matrices resulting from replicates for a given K value. We tested 1000 input orders of the 10 runs using the "Greedy" search algorithm. To account for effects of spatial distribution of samples, we estimated admixture proportions across the same range of K values with TESS3 (Caye et al. 2016). This approach incorporates geospatial distribution of samples into a least-squares optimization to estimate K. Unlike STRUCTURE, TESS3 does not assume Hardy-Weinberg Equilibrium, making it a suitable complementary analysis.

2.4 Gene tree estimation

To determine the evolutionary relationships between the inferred *T. sirtalis* clusters we estimated intraspecific divergence times in the context of the *Thamnophis* phylogeny from *cytB* in BEAST 1.10 (Suchard et al. 2018). The high evolutionary rate, haploidy, and near-zero recombination rate of mitochondrial loci preserve their value alongside multilocus markers generated from high throughput sequencing methods, particularly for identifying incomplete lineage sorting (ILS) and accurate divergence time estimation in the absence of a known species tree. We time calibrated the tree using the earliest known Thamnophis fossil from the Medial Barstovian (13-14.5 MYA) (Holman 2000) and constrained the clade age of *Thamnophis + Adelophis foxi* to a normal prior distribution with a mean = 14 and SD = 1 as performed in previous phylogenetic investigations of natricine species relationships (Wood et al. 2011; McVay et al. 2015). We selected the GTR+G+I substitution model with the Bayesian Information Criterion in JModelTest2 (Darriba et al. 2012) and ran the analysis with a relaxed lognormal clock for 7,000,000 iterations sampling every 1000 steps. Two identical analyses were run with different random starting seeds and combined with LogCombiner with the first 10% discarded as burn-in. We checked parameter estimates for convergence with Tracer v1.7.1 (Rambaut et al. 2018).

2.5 Species tree estimation

We inferred species trees from SNP data using the coalescent package SNAPP (Bryant et al. 2012), implemented in BEAST 2.6 (Bouckaert et al. 2014). We resampled SNPs from 10 individuals with high read counts and locus coverage that also covered the spatial distribution of the populations detected by the population structure analyses (bolded, Table S1) with no missing data to eliminate its potential impact on phylogenetic estimates. We sampled backward and forward mutation rates *u* and *v* with initial values = 1 and because we have no knowledge of the speciation rate lambda within *T. sirtalis*, employed an illegitimate/infinite distribution prior (1/x) for it. For the population size parameter applied a gamma distribution prior with a mean alpha/beta = 0.0085, which represents the mean number of variants/site found in the total dataset used for the population structure analysis and the reduced set used for species tree estimation. We ran the analysis with a random starting seed for 5,000,000 generations while sampling every 100, and then removed the first 25% as burn-in. We assessed parameter estimate stability and

generated a maximum clade credibility tree using TreeAnnotator. We inferred divergence times by scaling the tree height to time units by dividing all the branch lengths by the human nuclear mutation rate 1.1x10-8 per site per generation. The most recent neutral substitution rate estimate for *Thamnophis sirtalis* was found to be the highest among squamate lineages and similar to that of humans (Perry et al. 2018), justifying the use of the human rate over the substantially lower generalized squamate rate, which is biased primarily by the substantially lower rate of *Python bivittatus* (Green et al. 2014; Pasquesi et al. 2018).

2.6 Detection of introgression between disjunct lineages under the MSC

Our phylogenetic approach using nuclear data does not incorporate gene flow or introgression, a biological reality in many empirical widespread systems that has remained difficult to model in a coalescent framework, particularly with large datasets. We analyzed the reduced dataset under a multispecies coalescent-with introgression model implemented in BPP 4.1 (Flouri et al. 2019). We set a fixed species tree and the species tree model prior to reflect the rooted relationships between the Central, East, and West lineages inferred from SNAPP (see RESULTS), i.e., the A00 analysis. We assume a JC substitution model with *T. sirtalis* and set inverse gamma priors for tau = (3, 0.001) and theta (3,0.002). To directly test the hypothesis of Westward expansion from central North America, we conducted separate analyses to estimate the probability of Central → West and East → West introgression (φ) with a starting value = 0.5 and a uniform beta distribution prior = (1,1). For each introgression scenario we ran 5 replicate analyses which for 200,000 generations, sampling every 2 iterations. We discarded the first 50,000 iterations as burn-in. After assessing stationarity between replicates, we combined the results of these 5 runs to achieve sufficient ESS values for each analysis.

2.7 Demographic analyses

We simulated the joint two-dimensional site frequency spectrum (2D-JSFS) using MOMENTS (Jouganous et al. 2017). MOMENTS uses the diffusion approximation framework of δaδi (Gutenkunst et al. 2009) and employs differential equations to simulate changes in allele frequencies between population pairs under different evolutionary regimes. Here we evaluated the frequency spectra for East-Central, Southeast-Central, Southeast-East, and Central-West groupings, reflecting likelihood of gene flow due to geographic proximity and sister relationships

in both the gene and species trees. For each population pair we maximized the number of segregating sites by projecting the frequency spectra (FS) down to a smaller sample size using MOMENTS' *fs.project* function, a common approach in analysis of incomplete datasets with missing data between populations and/or individuals. This function essentially averages over all possible re-samplings of the larger dataset (Marth et al. 2004). The user then chooses an appropriate projection based on seeking a balance between increasing the number of SNPs for use in the analysis while not substantially d We then evaluated the fit of 20 demographic models (Supplementary data) representing historic population size changes and varying degrees of migration that might reflect range reductions and expansions through the late Miocene to Pleistocene. For every population pair, we optimized each model for 4 rounds of 100 replicates with parameters for each round's highest log-likelihood used as the starting values for the round after. We ranked the models for each population pair from their AIC, ΔAIC, and Akaike weights (Burnham and Anderson 2004), and scaled reference population sizes to fit the intraspecific divergence times inferred by the mtDNA and SNP phylogenies. We fit models to population pairs and plotted all results with modified python scripts developed for δaδi by Dan Portik ((Portik et al. 2017), https://github.com/dportik).

2.8 Species distribution modeling

We constructed species distribution models (SDM) for *T. sirtalis* using MAXENT (Phillips et al. 2006) implemented in ENMTools (Warren et al. 2010). We first assembled occurrence data for all samples used in the genetic analyses (n=217) and research grade *T. sirtalis* occurrence data (n=13,331) from GBIF (GBIF.org, 2019) for maximum range estimation accuracy. We assembled elevation data, 19 bioclimatic variables derived from temperature, seasonality, precipitation records (Fick and Hijmans 2017), and proximity to major water bodies (Lehner et al. 2008) for model construction. Prior to analyses, global raster layers were cropped to fit the extent of *T. sirtalis*' North American distribution (-130 ~ -53 degrees longitude; 25 ~ 60 degrees latitude). To mitigate the effects of collinearity, we removed variables with a Pearson's correlation coefficient of > 0.75, reducing the final dataset to 9 bioclimatic variables related to temperature, diurnality, and precipitation, in addition to elevation and proximity to water. To minimize the risk of overfitting the model to the data (Warren and Seifert 2011), 25% of the occurrence points were withheld to train the model (Syfert et al. 2013) as conservatively

practiced in similar studies (Gray et al. 2019). We ranked Model efficacy with the AUC score (Phillips et al. 2006). We calculated niche overlap between putative *T. sirtalis* lineages using Schoener's D (Schoener 1968) and Warren's I statistic (Warren et al. 2008). We made comparisons between the lineages detected by the population structure and phylogenetic analyses, and then conducted niche equivalency tests to generate a null probability distribution of niche overlap between proximal *T. sirtalis* lineages (see SUPPLEMENT).

RESULTS

3.1 Genomic data

The final thamnophiine *cytB* alignment consisted of 812 bp from 99 individuals, with 530 identical sites. STACKS processed 15747 loci (713935 sites) consisting of 8448 SNPs across 217 samples. For all samples, an average of 86.58% of the total reads were retained. The reduced 10 sample/population assembly for the SNAPP species tree estimation consisted of in 4630 loci with 1327 unlinked SNPs. Additional summary statistics from genomic data processing in STACKS are shown in **Table 3**.

3.2 Population structure

The results for all population structure analyses are summarized in **Fig 2.** Both STRUCTURE and TESS analyses most strongly supported a K = 3 model and inferred distinct genetic clusters corresponding to the West Coast and Pacific Northwest, the Eastern half of North America, and a significantly less admixed population corresponding to the Southeast Atlantic coastal plain (Southeastern Alabama, Georgia, Florida, and coastal South Carolina). Under the K = 4 model STRUCTURE and TESS inferred a 4th lineage largely corresponding to the Midwestern North America (West of the Mississippi River Basin, northward towards Central Canada). PCA analyses recapitulated this pattern, with PC1 summarizing the variation across an East-West cline, and PC2 accounting for the strong differentiation between Southeast and the East.

3.3 Phylogenetic relationships and divergence times

The mtDNA tree topology recapitulates broader, highly supported, *Thamnophis* relationships and divergence times estimated from prior studies (McVay et al. 2015) from which additional samples were sourced (**Supplementary data**). Within *T. sirtalis*, inferred mitochondrial clades

represent geographically structured lineages spanning distributions largely congruent with those inferred by nuclear SNP-based population structure analyses (**Fig 3**). Fossil calibrated divergence time estimates based on *cytB* within *T. sirtalis* support a Pliocene divergence (6.3 MYA, 95% HPD: 3.9-8.9) of a clade restricted to the Atlantic Coastal Plain (herein, the southeast lineage) and the rest of the tree. An eastern lineage occupies the temperate broadleaf forests of the eastern United States and extends into westward to the Ohio River. This border is marked by the transition to the more arid grasslands and floodplains of the central U.S. and Canada, and remaining individuals distributed westward to the throughout the Pacific Northwest and west coast from a strongly supported clade with a split into two weakly supported subclades: a central lineage containing individuals dispersed throughout the Mississippi River Basin east of the Ohio River and isolated individuals in the Southwest along the Rio Grande, and a western lineage of individuals found along the Pacific Coast through CA up to Vancouver, BC.

Contrary to the mtDNA tree, the SNAPP species tree topology shows strong support for sister relationships between western + central lineages (PD = 0.87) and the southeast + east lineages (PD = 0.99) with a substantially shallower late Pleistocene (~27 KYA) root age. All BPP analyses incorporating introgression topologically disagree with both the SNAPP and mtDNA trees. The A01 analysis weakly supports a sister relationship between the east and central lineages (PD=0.61) with a southeast divergence at the root (PD = 1.00). Both paired A00 analyses evaluating central→west and east→west introgression converge upon similar root ages (~23 and ~25 KYA, respectively) as the SNAPP analysis, but the probability of introgression (φ) from the central to west lineage is reasonably higher (φ (central→west)= 0.78) than that from the east (φ (east→west)= 0.61). Generally, node ages (τ) and θ estimates are consistent between the two BPP analyses and sufficient ESS values (>200) were reached each replicate. Detailed results are available in the supplement.

3.4 Demographic analyses

Demographic models accounting for >95% of the cumulative **ω**AIC for all lineage pairs are summarized in **Table 2**. The highest ranked models for all lineage pairs support size changes following sustained periods of isolation with varying degrees of migration between them (**Fig 4**). For the southeast-east pair, the highest ranked model suggests isolation followed by increased

population size changes in both lineages, and asymmetric migration into the southeast upon secondary contact. The central and east lineages follow a similar isolation pattern, but with symmetric migration between them. For all three population pair comparisons. For the central and west lineages, two similarly complex models of asymmetric migration and population size change accounted for the 90% of the cumulative **ω**AIC, with disproportionately greater gene flow from the central lineage to the west, with no size change in either over time.

3.6 Ecological niche modeling/overlap

Predicted species distributions for *T. sirtalis* lineages are shown in **Fig 5**. All species distribution models exhibited high AUC values (>0.9). The four models predict the wide fundamental niche as manifested in its occurrence throughout North America to the relative exclusion of the Great Basin and the southwest. The overlap of the southeast and east lineage SDMs exhibit a spatial pattern to be expected under the highest supported demographic model of asymmetric gene flow into the southeast from the east, albeit with slightly more suitable habitat for the east lineage along the Atlantic side of peninsular Florida than on the Gulf side. Estimated suitable habitat for the central lineage extends the farthest north of the four models and overlaps substantially with the eastern extent of suitable habitat suggested by the west lineage SDM. The southeast lineage SDM exhibits the most restricted distribution of suitable habitat of the four models, with the remaining three showing moderately suitable habitat in parts of the Great Basin and southwestern US.

DISCUSSION

Population structure

Our SNP-based investigation of the population structure within *T. sirtalis* reveals previously undetected hierarchical clustering patterns across North America. The different methods roughly agree on the location and composition of three spatially defined groups, with the southeast individuals forming the most genetically distinct cluster (or the most uniformly admixed under STRUCTURE's K=3 model). The spatial extents of the fourth central North American cluster identified by DAPC and the STRUCTURE and TESS K=4 models vary between analyses, with DAPC producing the least spatially structured results. Taken together with the phylogenetic estimates discussed below, we interpret the admixture patterns under the TESS3 K=3 and K=4

models and the clinal pattern exhibited across east, central, and west individuals along the PC2 axis in the PCA results as the most biologically meaningful. Assuming these patterns best reflect the "true" genetic variation within and between clusters, we note that the east cluster of *T. sirtalis* does not clearly conform to isolation patterns informed by the Mississippi River, Appalachian Mountains, or Atlantic/Gulf Coast discontinuities as observed in other Eastern squamate and amphibian systems (Pyron and Burbrink 2009; Soltis et al. 2006). The southeast cluster exhibits the clearest spatial isolation in its restriction to the South Atlantic coastal plain, but elevated admixture in southern Alabama and non-peninsular Florida individuals complicate objective assignment to the central or southeast lineages.

Phylogenetic relationships and Phylogeography

The results of our phylogenetic analyses are the first to characterize the intraspecific relationships within *Thamnophis sirtalis*. We find continued support for the mid-Miocene (~12 MYA) origin of the clade containing *Thamnophis sirtalis* and the central and central/eastern North American *T. sauritus*+*T. proximus* subclade found in prior studies (McVay et al. 2015), and conclude a likely mid-Miocene origin in the central or southeastern US for the ancestor of the inferred *T. sirtalis* lineages. This period was characterized by the expansion of North American grasslands and associated with an increase in diversification of snake species as evidenced in the fossil record and recent thamnophiine phylogenetic studies (Guo et al. 2012; Holman et al. 2000; McVay et al. 2015). The stark divergence of the mtDNA-defined southeast lineage coupled with the minimal admixture detected in SNP data from the TESS3 and STRUCTURE K=4 models suggest a period of sustained isolation that predates the Pleistocene glacial cycles generally thought to impact contemporary population structure in other species distributed across the east and southeast United States (Kozak et al., 2006). The striking difference in divergence time estimates from the nuclear and mtDNA suggest a substantial amount of gene flow since late Pleistocene glacial recession and Holocene recolonization.

The SNP-based phylogenetic analyses that incorporate introgression corroborate a scenario of recent admixture across the soft boundaries delineating the central and east lineages, as well as a significant degree of shared ancestry between the central and west lineages. This pattern exists in violation of the assumption of no gene flow inherent in many coalescent species tree estimation

methods. We demonstrate the shortcomings of model inadequacy here with our SNAPP results, which produced a topology uniquely incongruent with the results of our species and gene tree estimation methods. Incorporation of natural phenomena like introgression and consideration of model bias in species tree estimation and population structure analyses are critical to future investigations of taxa exhibiting hierarchical population structure and/or substantial gene flow between putative lineages. Whether or not the central lineage is a subpopulation of the east or west lineage is further complicated by the gene tree/species tree discordance in its phylogenetic placement, but we cautiously conclude from the species tree and population structure results that *T. sirtalis* exhibited a rapid westward expansion from the central or southeastern US following late Pleistocene glacial recession or an artifact of the unique genomics of the species, supporting conclusions reached in prior work (Janzen et al. 2002).

Demography and ecology

Our demographic analyses suggest multiple mechanisms underlying the recent dispersal history of *Thamnophis sirtalis*. The predicted asymmetric gene flow observed in all lineage comparisons except the central-southeast comparison operates in concordance with a scenario of movement mediated by local adaptation to environment. Outgoing gene flow is observed in all comparisons between the central lineage and other lineages, which together with its SDM suggest a lineage-specific proclivity towards successful occupation of multiple habitat types. It is well understood that *T. sirtalis* populations vary widely in ecology, and our demographic and ecological characterization of the *T. sirtalis* lineages highlight a possibility of differing dispersal ability corresponding to differing life history traits (i.e., fidelity to overwintering den sites). Future investigations of the processes underlying southeast-east genetic isolation and central-southeast gene flow should involve whole-genome interrogation across the lineages' respective contact zones for loci under potential ecology-associated divergent selection and hybrid vigor or decay.

The putative contact zone between the central and east lineages is the largest between any two lineages, leading to multiple opportunities for restriction to shared refugia and post-Pleistocene contact. This shared history between the two lineages perhaps explains both the recency of their inferred divergence time from SNP data and the ambiguity of the central lineage's west and east geographic boundaries across the population structure analyses. The central lineage's spatial

distribution largely spans the Mississippi River Drainage with the East comprising the entire United States and Canada east of the Mississippi and Apalachicola-Chattahoochee-Flint drainages. The inferred sister phylogenetic relationship between the central and west mtDNA lineages is contrasted by their nuclear divergence as the 2nd most genetically differentiated lineage pair (after the west and the southeast). We note that our demographic analyses do not model founder effect, which cannot be ruled out as a partial explanation for the genetic differentiation of the west lineage. Further sampling of the northwest, southwest (disjunct Transverse Range populations) and central margins of the range of *T. sirtalis* will be required to further characterize the post-Pleistocene dispersal history of the west lineage and the boundaries of the central lineage.

As a general caveat to demographic modeling of populations, we stress that the complexity of the historic and contemporary interactions between the southeast, east, west, and central lineages cannot be fully captured by any single model or in a two-dimensional framework. In the case here, genetic contribution from the unsampled second and third populations is inevitably unaccounted for in all 2D comparisons, potentially impacting measures of variation and consequently, ancestral population size estimates. Furthermore, to our knowledge the statistical impacts of down-projecting the FS to increase SNP counts for analysis have not yet been studied and remains an open question beyond the scope of this study.

Species distribution modeling

We present the first species distribution models for *T. sirtalis* throughout its distribution and between persistent genetic lineages. Our results show a strong correlation between ecological niche divergence and intraspecific variation, but we note that this alone cannot infer causality in either direction. Niche similarity between identified lineages is high even while calculated suitable habitat largely mimics the distribution of each lineage. As in all species distribution analyses, our proxy for "niche" (spatially biased bioclimatic data, altitude, proximity to established bodies of water, and vegetation cover) is limited in capturing the entirety of the environment with which *T. sirtalis* interacts. The intractability of accounting for the impact of variables like variation in population density, resource competition with other natricine snakes, predator threat, and regional geographic history on genetic patterns is a widely acknowledged

limitation of species distribution modeling (Guisan et al., 2005, Sexton et al., 2009) and these factors must be considered as additional mutually nonexclusive drivers of the observed genetic patterns through time and space.

Taxonomy

This study provides new lines of evidence to consider in the taxonomic status of *Thamnophis sirtalis*, from which subspecies have previously only been described based on regional phenotypes. Prior studies with mtDNA and extensive character data have already highlighted the unreliability of these designations as a proxy for intraspecific evolutionary relationships, and further evaluating them is not an aim of this study. The population-level variation in *T. sirtalis* ecology and morphology is perhaps unjustifiably described as a product of plasticity in the absence of 1) understanding of population and phylogenetic history from more genomic data, 2) evidence for the genetic basis of these regional differences and reproductive compatibility in controlled lab conditions. Here we present genomic evidence of four phylogenetic lineages within *Thamnophis sirtalis* and give context to their relationships with demographic and ecological niche modeling. The North American *Thamnophis* clade exhibits the elevated Pliocene diversification pattern observed in other North American terrestrial ectotherms (McVay et al. 2015) and given that *Thamnophis sirtalis* exhibits deeper intraspecific divergence times (when estimated from mtDNA) than what is observed between other sister species of a genus known for high rates of polymorphism, we contend that the southeastern lineage represents a unique candidate for taxonomic reevaluation. In consideration of high genetic divergence and clear restriction to the subtropical coniferous forests and marshland of the southeastern United States, we draw attention to the need for more robust evaluation of the extent of gene flow between the east and southeastern lineages prior to taxonomic revision.

ACKNOWLEDGMENTS

The authors dedicate this study to Douglas A. Rossman and Robin Lawson, whose indelible contributions to our understanding of natricine evolution and ecology provided the foundation for this study. Multiple museum institutions contributed tissue samples used in this study. The authors are especially grateful for Sharon Birks and Kevin Epperly (UWBM), Luke Welton (KU), Jens Vindum (CAS), Mary Sollows (NBM), David Kizirian (AMNH), Eli Greenbaum

(UTEP), Travis LaDuc (TNHC), Donna Dittmann (LSUMZ), Gregory Watkins-Colwell (YPM) for the processing and contribution of catalogued tissue loans used for this study. Joe Felsenstein and Richard Olmstead provided comments that greatly improved the manuscript. This work was supported by grants from the National Science Foundation (DEB #1600844) and the American Museum of Natural History (Theodore Roosevelt Award).

REFERENCES

Alexander Pyron R., Burbrink F.T. 2009. Lineage diversification in a widespread species: roles for niche divergence and conservatism in the common kingsnake, *Lampropeltis getula*. Mol Ecol. 18:3443–3457.

Alfaro M.E., Arnold S.J. 2001. Molecular Systematics and Evolution of Regina and the Thamnophiine Snakes. Molecular Phylogenetics and Evolution. 21:408–423.

Andrews K.R., Good J.M., Miller M.R., Luikart G., Hohenlohe P.A. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics. 17:81–92.

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

Avise J.C. 2000. Phylogeography: The History and Formation of Species. Harvard University Press.

Avise J.C., Arnold J., Ball R.M., Bermingham E., Lamb T., Neigel J.E., Reeb C.A., Saunders N.C. 1987. Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annu. Rev. Ecol. Syst. 18:489–522.

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

Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.-H., Xie D., Suchard M.A., Rambaut A., Drummond A.J. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLOS Computational Biology. 10:e1003537.

Boulenger G.A. 1896. Catalogue of the Snakes in the British Museum (Natural History)... order of the Trustees.

Bronikowski A.M. 2000. Experimental Evidence for the Adaptive Evolution of Growth Rate in the Garter Snake Thamnophis Elegans. Evolution. 54:1760–1767.

Brunsfeld S.J., Sullivan J., Soltis D.E., Soltis P.S. 2001. Comparative phylogeography of northwestern North America: a synthesis. Special Publication-British Ecological Society. 14:319–340.

Bryant D., Bouckaert R., Felsenstein J., Rosenberg N.A., RoyChoudhury A. 2012. Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. Molecular Biology and Evolution. 29:1917–1932.

Buckley L.B., Jetz W. 2007. Environmental and historical constraints on global patterns of amphibian richness. Proceedings of the Royal Society B: Biological Sciences. 274:1167–1173.

Burbrink F.T. 2002. Phylogeographic analysis of the cornsnake (Elaphe guttata) complex as inferred from maximum likelihood and Bayesian analyses. Molecular Phylogenetics and Evolution. 25:465–476.

Burbrink F.T., Chan Y.L., Myers E.A., Ruane S., Smith B.T., Hickerson M.J. 2016. Asynchronous demographic responses to Pleistocene climate change in Eastern Nearctic vertebrates. Ecology letters. 19:1457–1467.

Burbrink F.T., Lawson R. 2007. How and when did Old World ratsnakes disperse into the New World? Molecular Phylogenetics and Evolution. 43:173–189.

Burbrink F.T., Lawson R., Slowinski J.B. 2000. Mitochondrial DNA Phylogeography of the Polytypic North American Rat Snake (elaphe Obsoleta): A Critique of the Subspecies Concept. Evolution. 54:2107–2118.

Burbrink F.T., Yao H., Ingrasci M., Bryson R.W., Guiher T.J., Ruane S. 2011. Speciation at the Mogollon Rim in the Arizona Mountain Kingsnake (Lampropeltis pyromelana). Molecular Phylogenetics and Evolution. 60:445–454.

Burnham K.P., Anderson D.R. 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research. 33:261–304.

Calsbeek R., Thompson J.N., Richardson J.E. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. Molecular Ecology. 12:1021–1029.

Castoe T.A., Bronikowski A.M., Brodie E.D., Edwards S.V., Pfrender M.E., Shapiro M.D., Pollock D.D., Warren W.C. 2011. A proposal to sequence the genome of a garter snake (Thamnophis sirtalis). Stand in Genomic Sci. 4:257–270.

Castoe T.A., Smith E.N., Brown R.M., Parkinson C.L. 2007. Higher-level phylogeny of Asian and American coralsnakes, their placement within the Elapidae (Squamata), and the systematic affinities of the enigmatic Asian coralsnake Hemibungarus calligaster (Wiegmann, 1834): RELATIONSHIPS OF CORALSNAKES AND HEMIBUNGARUS. Zoological Journal of the Linnean Society. 151:809–831.

Caye K., Deist T.M., Martins H., Michel O., François O. 2016. TESS3: fast inference of spatial population structure and genome scans for selection. Molecular Ecology Resources. 16:540–548.

Ceríaco L.M.P., Marques M.P., Schmitz A., Bauer A.M. 2017. The “Cobra-preta” of São Tomé Island, Gulf of Guinea, is a new species of Naja Laurenti, 1768 (Squamata: Elapidae). Zootaxa. 4324:121.

Chifman J., Kubatko L. 2014. Quartet Inference from SNP Data Under the Coalescent Model. Bioinformatics. 30:3317–3324.

Darriba D., Taboada G.L., Doallo R., Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature methods. 9:772–772.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution. 24:332–340.

Deutsch C.A., Tewksbury J.J., Huey R.B., Sheldon K.S., Ghalambor C.K., Haak D.C., Martin P.R. 2008. Impacts of climate warming on terrestrial ectotherms across latitude. PNAS. 105:6668–6672.

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

Earl D.A., vonHoldt B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genet Resour. 4:359–361.

Easterbrook D.J. 1969. Pleistocene chronology of the Puget lowland and San Juan islands, Washington. Geological Society of America Bulletin. 80:2273–2286.

Eaton D.A., Overcast I. 2016. ipyrad: Interactive assembly and analysis of RADseq datasets. Bioinformatics.

Edgehouse M., Latta L.C., Brodie E.D., Brodie E.D. 2014. Interspecific Aggression and Habitat Partitioning in Garter Snakes. PLoS One. 9.

Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution. 63:1–19.

Edwards S.V., Potter S., Schmitt C.J., Bragg J.G., Moritz C. 2016. Reticulation, divergence, and the phylogeography–phylogenetics continuum. Proc Natl Acad Sci USA. 113:8025–8032.

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology. 14:2611–2620.

Faircloth B.C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics. 32:786–788.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved Elements Anchor Thousands of Genetic Markers Spanning Multiple Evolutionary Timescales. Syst Biol. 61:717–726.

Felsenstein J. 1978. Cases in which Parsimony or Compatibility Methods will be Positively Misleading. Syst Biol. 27:401–410.

Fick S.E., Hijmans R.J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology. 37:4302–4315.

Fitch H.S. 1941. Geographic Variation in Garter Snakes of the Species Thamnophis sirtalis in the Pacific Coast Region of North America. American Midland Naturalist. 26:570.

Fitch H.S. 1981. Thamnophis sirtalis. Catalogue of American Amphibians and Reptiles (CAAR).

Flouri T., Jiao X., Rannala B., Yang Z. 2018. Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. Mol Biol Evol. 35:2585–2593.

Flouri T., Jiao X., Rannala B., Yang Z. 2019. A Bayesian Implementation of the Multispecies Coalescent Model with Introgression for Phylogenomic Analysis. Mol Biol Evol.

Gang W., Peng G., Ning X., JingZhuo Z., Zhuang S. 2011. Naja kaouthia: a new snake record to Guizhou. Guizhou Agricultural Sciences.:173–176.

Gaston K.J. 2003. The structure and dynamics of geographic ranges. Oxford: Oxford University Press.

GBIFAvailable from https://www.gbif.org/.

Gould S.J., Johnston R.F. 1972. Geographic Variation. Annual review of ecology and systematics. 3:457–498.

Gray L.N., Barley A.J., Poe S., Thomson R.C., Nieto-Montes de Oca A., Wang I.J. 2019. Phylogeography of a widespread lizard complex reflects patterns of both geographic and ecological isolation. Mol Ecol. 28:644–657.

Green R.E., Braun E.L., Armstrong J., Earl D., Nguyen N., Hickey G., Vandewege M.W., St. John J.A., Capella-Gutierrez S., Castoe T.A., Kern C., Fujita M.K., Opazo J.C., Jurka J., Kojima K.K., Caballero J., Hubley R.M., Smit A.F., Platt R.N., Lavoie C.A., Ramakodi M.P., Finger J.W., Suh A., Isberg S.R., Miles L., Chong A.Y., Jaratlerdsiri W., Gongora J., Moran C., Iriarte A., McCormack J., Burgess S.C., Edwards S.V., Lyons E., Williams C., Breen M., Howard J.T., Gresham C.R., Peterson D.G., Schmitz J., Pollock D.D., Haussler D., Triplett E.W., Zhang G., Irie N., Jarvis E.D., Brochu C.A., Schmidt C.J., McCarthy F.M., Faircloth B.C., Hoffmann F.G., Glenn T.C., Gabaldon T., Paten B., Ray D.A. 2014. Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. Science. 346:1254449–1254449.

Gregory P.T. 1978. Feeding habits and diet overlap of three species of garter snakes (Thamnophis) on Vancouver Island. Can. J. Zool. 56:1967–1974.

Gregory P.T. 1984. Habitat, diet, and composition of assemblages of garter snakes (Thamnophis) at eight sites on Vancouver Island. Can. J. Zool. 62:2013–2022.

Guo P., Liu Q., Xu Y., Jiang K., Hou M., Ding L., Alexander Pyron R., Burbrink F.T. 2012. Out of Asia: Natricine snakes support the Cenozoic Beringian Dispersal Hypothesis. Molecular Phylogenetics and Evolution. 63:825–833.

Gutenkunst R.N., Hernandez R.D., Williamson S.H., Bustamante C.D. 2009. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. PLoS Genet. 5.

Harris R.S. 2007. Improved pairwise Alignmnet of genomic DNA. .

Harvey M.G., Smith B.T., Glenn T.C., Faircloth B.C., Brumfield R.T. 2016. Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. Syst Biol. 65:910–924.

Head J.J., Holroyd P.A., Hutchison J.H., Ciochon R.L. 2005. First report of snakes (Serpentes) from the Late Middle Eocene Pondaung Formation, Myanmar. Journal of Vertebrate Paleontology. 25:246–250.

Heath T.A., Zwickl D.J., Kim J., Hillis D.M. 2008. Taxon Sampling Affects Inferences of Macroevolutionary Processes from Phylogenetic Trees. Syst Biol. 57:160–166.

Hewitt G. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society. 58:247–276.

Hickerson M.J., Carstens B.C., Cavender-Bares J., Crandall K.A., Graham C.H., Johnson J.B., Rissler L., Victoriano P.F., Yoder A.D. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Molecular Phylogenetics and Evolution. 54:291–301.

Holman J.A., others. 2000. Fossil snakes of North America: origin, evolution, distribution, paleoecology. Indiana University Press.

Huson D.H., Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol. 23:254–267.

Ineich I. 1995. Etat actuel de nos connaissances sur la classification des serpents venimeux. Bulletin de la Société herpétologique de France.:7–24.

Irwin D.E. 2002. Phylogeographic breaks without geographic barriers to gene flow. Evolution. 56:2383–2394.

Jackson N.D., Carstens B.C., Morales A.E., O'Meara B.C. 2017. Species Delimitation with Gene Flow. Syst Biol. 66:799–812.

Jakobsson M., Rosenberg N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 23:1801–1806.

Janzen F.J., Krenz J.G., Haselkorn T.S., Brodie E.D. 2002. Molecular phylogeography of common garter snakes (Thamnophis sirtalis) in western North America: implications for regional historical forces. Mol Ecol. 11:1739–1751.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 24:1403–1405.

Jouganous J., Long W., Ragsdale A.P., Gravel S. 2017. Inferring the Joint Demographic History of Multiple Populations: Beyond the Diffusion Approximation. Genetics. 206:1549–1567.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution. 30:772–780.

King R.B., Lawson R. 2001. Patterns of Population Subdivision and Gene Flow in Three Sympatric Natricine Snakes. Copeia. 2001:602–614.

Kishida T., Go Y., Tatsumoto S., Tatsumi K., Kuraku S., Toda M. 2019. Loss of olfaction in sea snakes provides new perspectives on the aquatic adaptation of amniotes. Proceedings of the Royal Society B. 286:20191828.

Kubatko L.S., Degnan J.H. 2007. Inconsistency of Phylogenetic Estimates from Concatenated Data under Coalescence. Syst Biol. 56:17–24.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2016. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol Biol Evol.:msw260.

Larsen K.W., Gregory P.T., Antoniak R. 1993. Reproductive Ecology of the Common Garter Snake Thamnophis sirtalis at the Northern Limit of Its Range. The American Midland Naturalist. 129:336–345.

Leaché A.D., Chavez A.S., Jones L.N., Grummer J.A., Gottscho A.D., Linkem C.W. 2015. Phylogenomics of Phrynosomatid Lizards: Conflicting Signals from Sequence Capture versus Restriction Site Associated DNA Sequencing. Genome Biology and Evolution. 7:706–719.

Leaché A.D., Zhu T., Rannala B., Yang Z. 2019. The Spectre of Too Many Species. Syst Biol. 68:168–181.

Lee M.S.Y., Sanders K.L., King B., Palci A. 2016. Diversification rates and phenotypic evolution in venomous snakes (Elapidae). Royal Society Open Science. 3:150277.

Lehner B., Verdin K., Jarvis A. 2008. New Global Hydrography Derived From Spaceborne Elevation Data. Eos, Transactions American Geophysical Union. 89:93–94.

Leopold E.B., Dunwiddie P.W., Whitlock C., Nickmann R., Watts W.A. 2016. Postglacial vegetation history of Orcas Island, northwestern Washington. Quaternary Research. 85:380–390.

Lobeck A.K. 1948. Physiographic Provinces of North America; Physiographic Diagram of North America. Geographical Press, Division of CS Hammond & Company.

Longbottom J., Shearer F.M., Devine M., Alcoba G., Chappuis F., Weiss D.J., Ray S.E., Ray N., Warrell D.A., Castañeda R.R. de, Williams D.J., Hay S.I., Pigott D.M. 2018. Vulnerability to snakebite envenoming: a global mapping of hotspots. The Lancet. 392:673–684.

Maddison W.P. 1997. Gene trees in species trees. Systematic biology. 46:523–536.

Manier M.K., Arnold S.J. 2005. Population genetic analysis identifies source–sink dynamics for two sympatric garter snake species (Thamnophis elegans and Thamnophis sirtalis). Molecular Ecology. 14:3965–3976.

Manthey J.D., Campillo L.C., Burns K.J., Moyle R.G. 2016. Comparison of Target-Capture and Restriction-Site Associated DNA Sequencing for Phylogenomics: A Test in Cardinalid Tanagers (Aves, Genus: Piranga). Syst Biol. 65:640–650.

Marth G.T., Czabarka E., Murvai J., Sherry S.T. 2004. The Allele Frequency Spectrum in Genome-Wide Human Variation Data Reveals Signals of Differential Demographic History in Three Large World Populations. Genetics. 166:351–372.

Mayr E. 1963. Animal species and evolution. Animal species and evolution.

McVay J.D., Flores-Villela O., Carstens B. 2015. Diversification of North American natricine snakes. Biol. J. Linn. Soc. 116:1–12.

Minton S.A. 1986. Origins of poisonous snakes: evidence from plasma and venom proteins. Natural toxins—animal, plant and microbial. Clarendon Press, Oxford.:3–21.

Mooi R.D., Wiens J.P., Casper G.S. 2011. Extreme Color Variation within Populations of the Common Gartersnake, Thamnophis sirtalis, in Central North America, with Implications for Subspecies Status. Copeia. 2011:187–200.

Moritz C., Langham G., Kearney M., Krockenberger A., VanDerWal J., Williams S. 2012. Integrating phylogeography and physiology reveals divergence of thermal traits between central and

peripheral lineages of tropical rainforest lizards. Philosophical Transactions of the Royal Society B: Biological Sciences. 367:1680–1687.

Mosher D.C., Hewitt A.T. 2004. Late Quaternary deglaciation and sea-level history of eastern Juan de Fuca Strait, Cascadia. Quaternary International. 121:23–39.

Mussmann S.M., Douglas M.R., Chafin T.K., Douglas M.E. 2019. BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data. Methods in Ecology and Evolution. 10:1808–1813.

Myers E.A., Weaver R.E., Alamillo H. 2013. Population Stability of the Northern Desert Nightsnake (Hypsiglena chlorophaea deserticola) during the Pleistocene. Journal of Herpetology. 47:432–439.

Nabhan A.R., Sarkar I.N. 2012. The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. Brief Bioinform. 13:122–134.

Nichols R. 2001. Gene trees and species trees are not the same. Trends in Ecology & Evolution. 16:358–364.

Paris J.R., Stevens J.R., Catchen J.M. 2017. Lost in parameter space: a road map for stacks. Methods in Ecology and Evolution. 8:1360–1373.

Pasquesi G.I.M., Adams R.H., Card D.C., Schield D.R., Corbin A.B., Perry B.W., Reyes-Velasco J., Ruggiero R.P., Vandewege M.W., Shortt J.A., Castoe T.A. 2018. Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. Nature Communications. 9:2774.

Perry B.W., Card D.C., McGlothlin J.W., Pasquesi G.I.M., Adams R.H., Schield D.R., Hales N.R., Corbin A.B., Demuth J.P., Hoffmann F.G., Vandewege M.W., Schott R.K., Bhattacharyya N., Chang B.S.W., Casewell N.R., Whiteley G., Reyes-Velasco J., Mackessy S.P., Gamble T., Storey K.B., Biggar K.K., Passow C.N., Kuo C.-H., McGaugh S.E., Bronikowski A.M., de Koning A.P.J., Edwards S.V., Pfrender M.E., Minx P., Brodie E.D., Brodie E.D., Warren W.C., Castoe T.A. 2018. Molecular Adaptations for Sensing and Securing Prey and Insight into Amniote Genome Diversity from the Garter Snake Genome. Genome Biology and Evolution. 10:2110–2129.

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S., Hoekstra H.E. 2012a. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLOS ONE. 7:e37135.

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S., Hoekstra H.E. 2012b. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS one. 7.

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

Placyk Jr. J.S., Burghardt G.M., Small R.L., King R.B., Casper G.S., Robinson J.W. 2007. Post-glacial recolonization of the Great Lakes region by the common gartersnake (Thamnophis sirtalis) inferred from mtDNA sequences. Molecular Phylogenetics and Evolution. 43:452–467.

Portik D.M., Leaché A.D., Rivera D., Barej M.F., Burger M., Hirschfeld M., Rödel M.-O., Blackburn D.C., Fujita M.K. 2017. Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. Molecular Ecology. 26:5245–5263.

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics. 155:945–959.

Pyron R.A., Burbrink F.T., Colli G.R., de Oca A.N.M., Vitt L.J., Kuczynski C.A., Wiens J.J. 2011. The phylogeny of advanced snakes (Colubroidea), with discovery of a new subfamily and comparison of support methods for likelihood trees. Molecular Phylogenetics and Evolution. 58:329–342.

Pyron R.A., Burbrink F.T., Wiens J.J. 2013a. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. BMC Evolutionary Biology. 13:93.

Pyron R.A., Kandambi H.K.D., Hendry C.R., Pushpamal V., Burbrink F.T., Somaweera R. 2013b. Genus-level phylogeny of snakes reveals the origins of species richness in Sri Lanka. Molecular Phylogenetics and Evolution. 66:969–978.

de Queiroz A., Lawson R., Lemos-Espinal J.A. 2002. Phylogenetic Relationships of North American Garter Snakes (Thamnophis) Based on Four Mitochondrial Genes: How Much DNA Sequence Is Enough? Molecular Phylogenetics and Evolution. 22:315–329.

Queiroz A., Lawson R. 1994. Phylogenetic relationships of the garter snakes based on DNA sequence and allozyme variation. Biological Journal of the Linnean Society. 53:209–229.

Queiroz A.D., Lawson R. 2008. A peninsula as an island: multiple forms of evidence for overwater colonization of Baja California by the gartersnake Thamnophis validus. BJLS. 95:409–424.

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic biology. 67:901.

Ricketts T.H., Dinerstein E., Olson D.M., Eichbaum W., Loucks C.J., Kavanaugh K., Hedao P., Hurley P., DellaSalla D., Abell R., Carney K., Walters S. 1999. Terrestrial Ecoregions of North America: A Conservation Assessment. Island Press.

Riddle B.R. 2016. Comparative phylogeography clarifies the complexity and problems of continental distribution that drove A. R. Wallace to favor islands. Proc Natl Acad Sci USA. 113:7970–7977.

Rochette N.C., Rivera-Colón A.G., Catchen J.M. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Molecular ecology. 28:4737–4754.

Rolland J., Silvestro D., Schluter D., Guisan A., Broennimann O., Salamin N. 2018. The impact of endothermy on the climatic niche evolution and the distribution of vertebrate diversity. Nature Ecology & Evolution. 2:459–464.

Rosenberg N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular ecology notes. 4:137–138.

Rossman D.A. 1996. The Garter Snakes: Evolution and Ecology. University of Oklahoma Press.

Ruane S., Austin C.C. 2017. Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. Molecular Ecology Resources. 17:1003–1008.

Ruthven A.G. 1908. Variations and genetic relationships of the garter-snakes. Govt. print. off.

Sanders K.L., Lee M.S.Y., Leys R., Foster R., Keogh J.S. 2008. Molecular phylogeny and divergence dates for Australasian elapids and sea snakes (hydrophiinae): evidence from seven genes for rapid evolutionary radiations. Journal of Evolutionary Biology. 21:682–695.

Santra V., Wüster W. 2017. Natural History: Naja kaouthia (Monocoled Cobra). Behavior / spitting. Herpetological Review. 48:455–456.

Schoener T.W. 1968. The Anolis Lizards of Bimini: Resource Partitioning in a Complex Fauna. Ecology. 49:704–726.

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

Shepard D.B., Burbrink F.T. 2009. Phylogeographic and demographic effects of Pleistocene climatic fluctuations in a montane salamander, *Plethodon fourchensis*. Mol Ecol. 18:2243–2262.

Slowinski J.B., Keogh J.S. 2000. Phylogenetic Relationships of Elapid Snakes Based on Cytochrome b mtDNA Sequences. Molecular Phylogenetics and Evolution. 15:157–164.

Smith B.T., Harvey M.G., Faircloth B.C., Glenn T.C., Brumfield R.T. 2014. Target Capture and Massively Parallel Sequencing of Ultraconserved Elements for Comparative Studies at Shallow Evolutionary Time Scales. Syst Biol. 63:83–95.

Soltis D.E., Gitzendanner M.A., Strenge D.D., Soltis P.S. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Pl Syst Evol. 206:353–373.

Soltis D.E., Morris A.B., McLACHLAN J.S., Manos P.S., Soltis P.S. 2006a. Comparative phylogeography of unglaciated eastern North America. Mol Ecol. 15:4261–4293.

Soltis D.E., Morris A.B., McLACHLAN J.S., Manos P.S., Soltis P.S. 2006b. Comparative phylogeography of unglaciated eastern North America. Molecular Ecology. 15:4261–4293.

Song S., Liu L., Edwards S.V., Wu S. 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proceedings of the National Academy of Sciences. 109:14942–14947.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.

Stange M., Sánchez-Villagra M.R., Salzburger W., Matschiner M. 2018. Bayesian Divergence-Time Estimation with Genome-Wide Single-Nucleotide Polymorphism Data of Sea Catfishes (Ariidae) Supports Miocene Closure of the Panamanian Isthmus. Syst Biol. 67:681–699.

Streicher J.W., Schulte J.A., Wiens J.J. 2016. How Should Genes and Taxa be Sampled for Phylogenomic Analyses with Missing Data? An Empirical Study in Iguanian Lizards. Syst Biol. 65:128–145.

Suchard M.A., Lemey P., Baele G., Ayres D.L., Drummond A.J., Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus evolution. 4:vey016.

Sukumaran J., Knowles L.L. 2017. Multispecies coalescent delimits structure, not species. PNAS. 114:1607–1612.

Suryamohan K., Krishnankutty S.P., Guillory J., Jevit M., Schröder M.S., Wu M., Kuriakose B., Mathew O.K., Perumal R.C., Koludarov I., others. 2020. The Indian cobra reference genome and transcriptome enables comprehensive identification of venom toxins. Nature Genetics.:1–12.

Swofford D.L. 2003. Phylogenetic analysis using parsimony (* and other methods). .

Syfert M.M., Smith M.J., Coomes D.A. 2013. The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. PLOS ONE. 8:e55158.

Trape J.-F., Chirio L., Broadley D.G., Wüster W. 2009. Phylogeography and systematic revision of the Egyptian cobra (Serpentes: Elapidae: Naja haje) species complex, with the description of a new species from West Africa. Zootaxa. 2236:1–25.

Uetz P., Cherikh S., Shea G., Ineich I., Campbell P.D., DORONIN I.V., Rosado J., Wynn A., Tighe K.A., McDiarmid R. 2019. A global catalog of primary reptile type specimens. Zootaxa. 4695:438–450.

Ullate-Agote A., Milinkovitch M.C., Tzika A.C. 2015. The genome sequence of the corn snake (Pantherophis guttatus), a valuable resource for EvoDevo studies in squamates. International Journal of Developmental Biology. 58:881–888.

Uyeda J.C. Connecting Microevolutionary Processes to Macroevolutionary Patterns Across Space and Time. :306.

Vonk F.J., Casewell N.R., Henkel C.V., Heimberg A.M., Jansen H.J., McCleary R.J.R., Kerkkamp H.M.E., Vos R.A., Guerreiro I., Calvete J.J., Wüster W., Woods A.E., Logan J.M., Harrison R.A., Castoe T.A., Koning A.P.J. de, Pollock D.D., Yandell M., Calderon D., Renjifo C., Currier R.B., Salgado D., Pla D., Sanz L., Hyder A.S., Ribeiro J.M.C., Arntzen J.W., Thillart G.E.E.J.M. van den, Boetzer M., Pirovano W., Dirks R.P., Spaink H.P., Duboule D., McGlinn E., Kini R.M., Richardson M.K. 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. PNAS. 110:20651–20656.

Voorhies M.R. 1990. Vertebrate biostratigraphy of the Ogallala Group in Nebraska. Geologic Framework and Regional Hydrology: Upper Cenozoic Blackwater Draw and Ogallala Forma- tion, Great Plains.:115–151.

Wallach V., Wüster W., Broadley D.G. 2009. In praise of subgenera: taxonomic status of cobras of the genus Naja Laurenti (Serpentes: Elapidae). Zootaxa. 2236:26–36.

Warren D.L., Glor R.E., Turelli M. 2008. Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. Evolution. 62:2868–2883.

Warren D.L., Glor R.E., Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography. 33:607–611.

Warren D.L., Seifert S.N. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecological Applications. 21:335–342.

Weese D.A., Fujita Y., Santos S.R. 2013. Multiple Colonizations Lead to Cryptic Biodiversity in an Island Ecosystem: Comparative Phylogeography of Anchialine Shrimp Species in the Ryukyu Archipelago, Japan. The Biological Bulletin. 225:24–41.

Wiens J.J. 2004. Speciation and Ecology Revisited: Phylogenetic Niche Conservatism and the Origin of Species. Evolution. 58:193–197.

Wiens J.J., Hutter C.R., Mulcahy D.G., Noonan B.P., Townsend T.M., Sites J.W., Reeder T.W. 2012. Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. Biology Letters. 8:1043–1046.

Wilson G.A., Rannala B. 2003. Bayesian Inference of Recent Migration Rates Using Multilocus Genotypes. Genetics. 163:1177–1191.

Wisz M.S., Pottier J., Kissling W.D., Pellissier L., Lenoir J., Damgaard C.F., Dormann C.F., Forchhammer M.C., Grytnes J.-A., Guisan A., Heikkinen R.K., Høye T.T., Kühn I., Luoto M., Maiorano L., Nilsson M.-C., Normand S., Öckinger E., Schmidt N.M., Termansen M., Timmermann A., Wardle D.A., Aastrup P., Svenning J.-C. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biological Reviews. 88:15–30.

Wood D.A., Vandergast A.G., Lemos Espinal J.A., Fisher R.N., Holycross A.T. 2011. Refugial isolation and divergence in the Narrowheaded Gartersnake species complex (Thamnophis rufipunctatus) as revealed by multilocus DNA sequence data: REFUGIAL ISOLATION IN NARROWHEADED GARTERSNAKES. Mol Ecol. 20:3856–3878.

Wüster W., Broadley D.G. 2007. Get an eyeful of this: a new species of giant spitting cobra from eastern and north-eastern Africa (Squamata: Serpentes: Elapidae: Naja). Zootaxa. 1532:51–68.

Wüster W., Chirio L., Trape J.-F., Ineich I., Jackson K., Greenbaum E., Barron C., Kusamba C., Nagy Z.T., Storey R., Hall C., Wüster C.E., Barlow A., Broadley D.G. 2018. Integration of nuclear and mitochondrial gene sequences and morphology reveals unexpected diversity in the forest cobra (Naja melanoleuca) species complex in Central and West Africa (Serpentes: Elapidae). Zootaxa. 4455:68–98.

Wüster W., Crookes S., Ineich I., Mané Y., Pook C.E., Trape J.-F., Broadley D.G. 2007. The phylogeny of cobras inferred from mitochondrial DNA sequences: Evolution of venom spitting and the phylogeography of the African spitting cobras (Serpentes: Elapidae: Naja nigricollis complex). Molecular Phylogenetics and Evolution. 45:437–453.

Wüster W., Thorpe R.S. 1992. Dentitional Phenomena in Cobras Revisited: Spitting and Fang Structure in the Asiatic Species of Naja (Serpentes: Elapidae). Herpetologica. 48:424–434.

Yap M.K.K., Tan N.H., Fung S.Y. 2011. Biochemical and toxinological characterization of Naja sumatrana (Equatorial spitting cobra) venom. Journal of Venomous Animals and Toxins including Tropical Diseases. 17:451–459.

Young B.A. 2004. The buccal buckle: the functional morphology of venom spitting in cobras. Journal of Experimental Biology. 207:3483–3494.

Zaher H., Murphy R.W., Arredondo J.C., Graboski R., Machado-Filho P.R., Mahlow K., Montingelli G.G., Quadros A.B., Orlov N.L., Wilkinson M., Zhang Y.-P., Grazziotin F.G. 2019. Large-scale molecular phylogeny, morphology, divergence-time estimation, and the fossil record of advanced caenophidian snakes (Squamata: Serpentes). PLOS ONE. 14:e0216148.

Zink R.M. 2002. Methods in Comparative Phylogeography, and Their Application to Studying Evolution in the North American Aridlands. Integr Comp Biol. 42:953–959

Fig 1. Sampling scheme for this study. Varied colors represent distinct Level II North American ecoregions with the distribution for *T. sirtalis* featured in the inset.

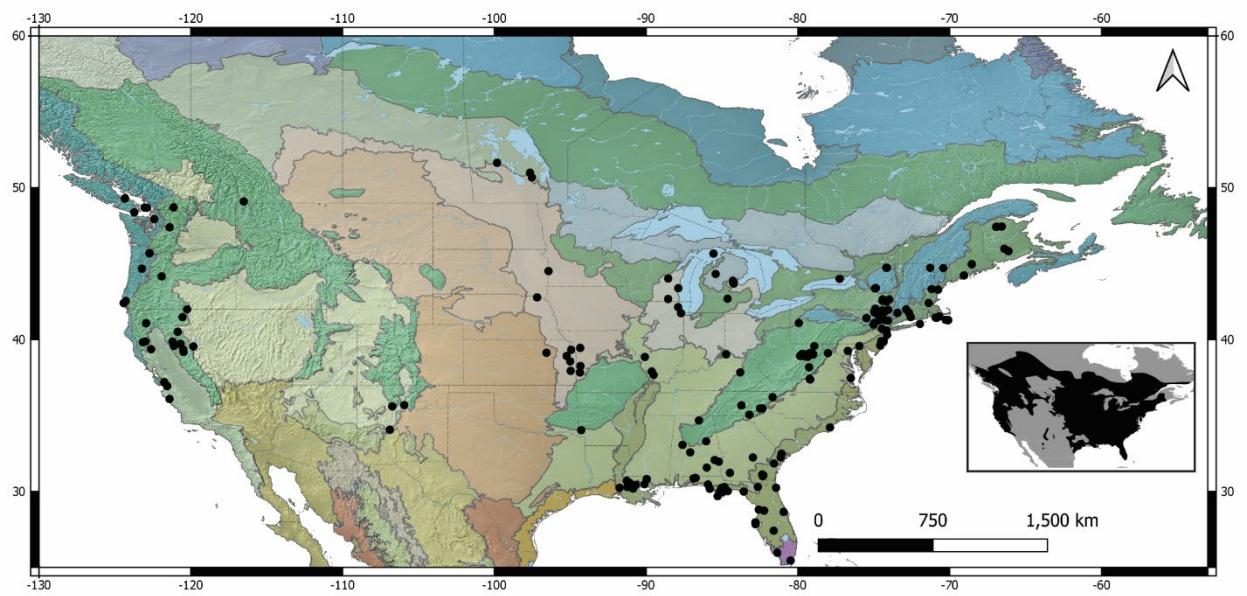


Fig 2. Population genetic structure results within *Thamnophis sirtalis* inferred from STRUCTURE (A-B) and TESS3 (C-D) under K=3 and K=4 models, the DAPC model with the highest BIC (E), and PCA (F).

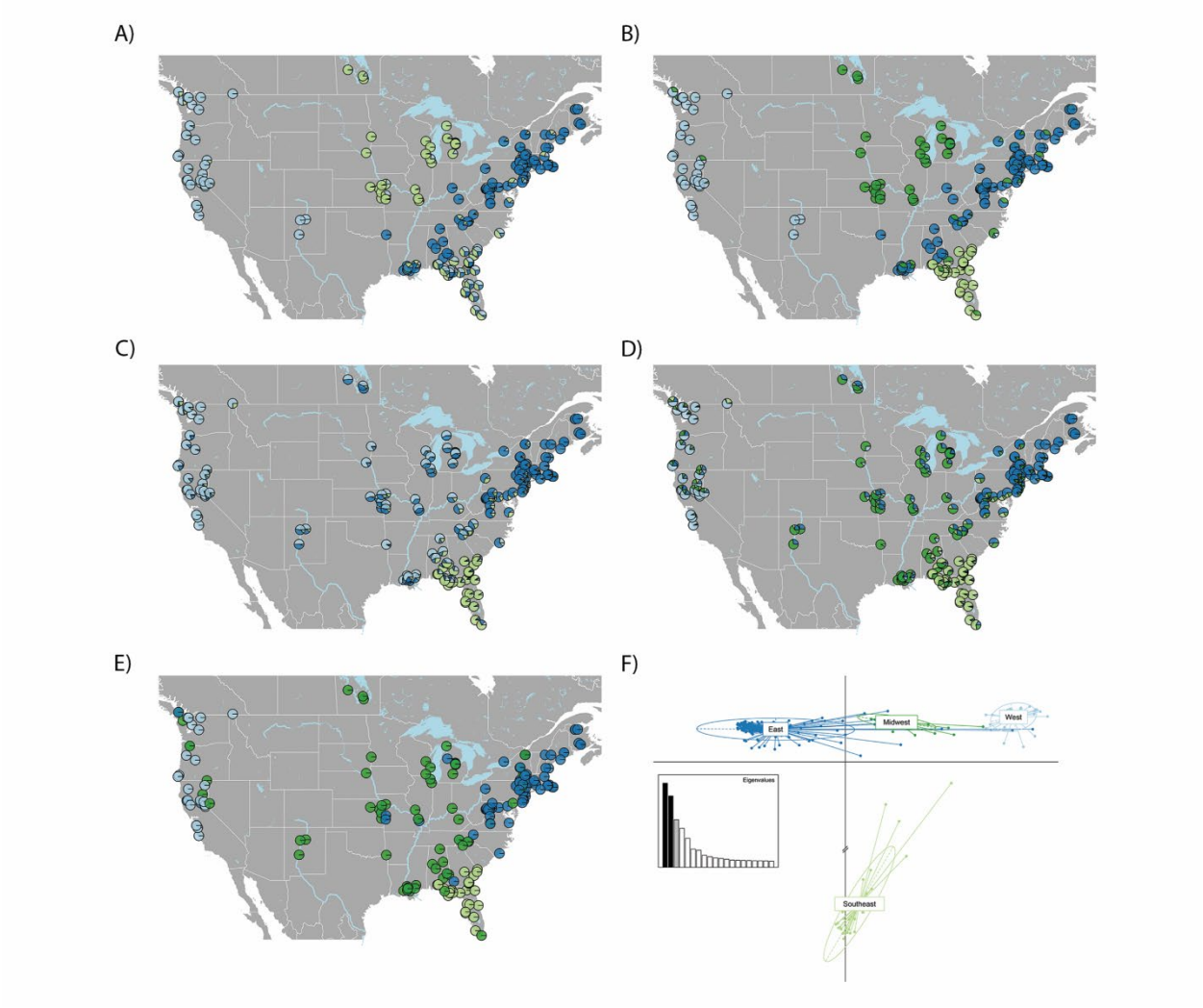


Fig 3. Phylogenetic relationships within *T. sirtalis*. **A)** *cytB* maximum clade credibility (MCC) tree estimated in BEAST v1.10. **B)** Consensus tree estimated by SNAPP v1.4.2 **C)** Visualization of BPP species tree (A01) and parameter (A00) estimation analyses.

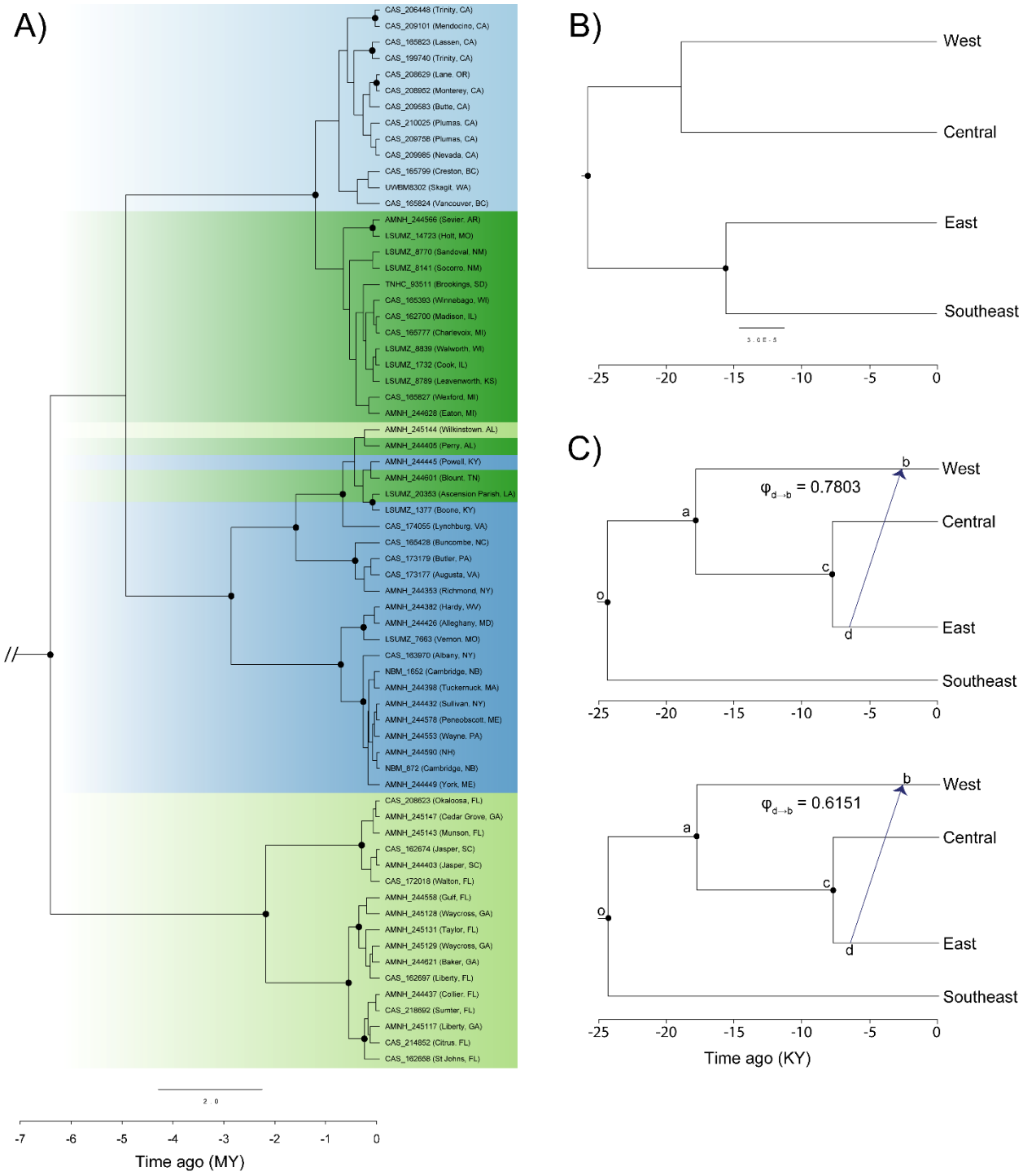


Fig 4. Highest ranked demographic models with residuals for phylogenetically defined *T. sirtalis* lineages using Moments. The reference population sizes (Nref) are scaled to median intraspecific divergence times inferred from BPP. Arrow width in the models reflect migration rates between populations, while red and blue residuals reflect the model's over- or under-prediction of allelic diversity.

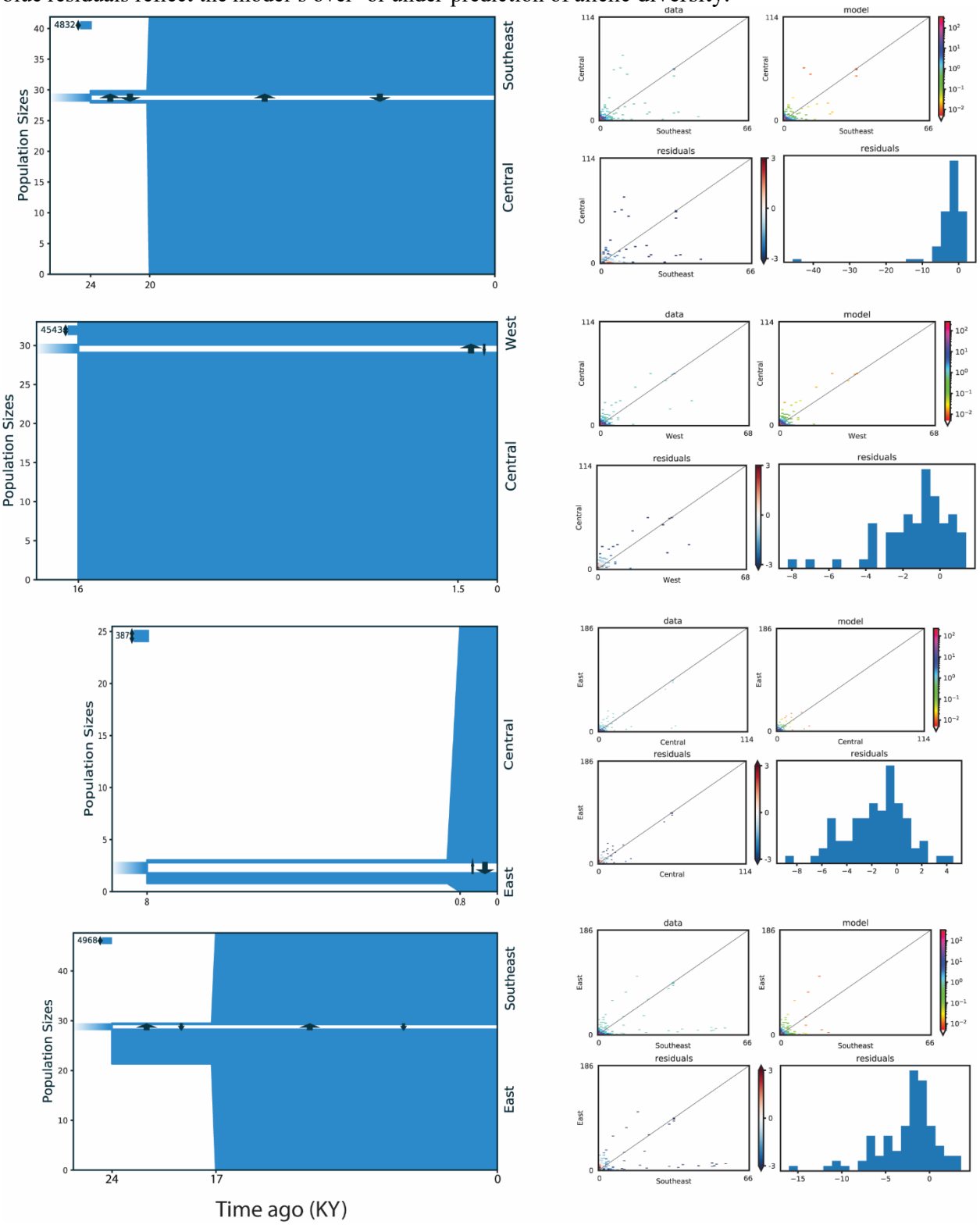


Fig 5. Species distribution models for *T. sirtalis* generated by MAXENT in ENMTools. Models were developed using spatially rarified research-grade georeferenced GBIF occurrences within perimeters corresponding to A) central, B) southeast, C) east, and D) west populations.

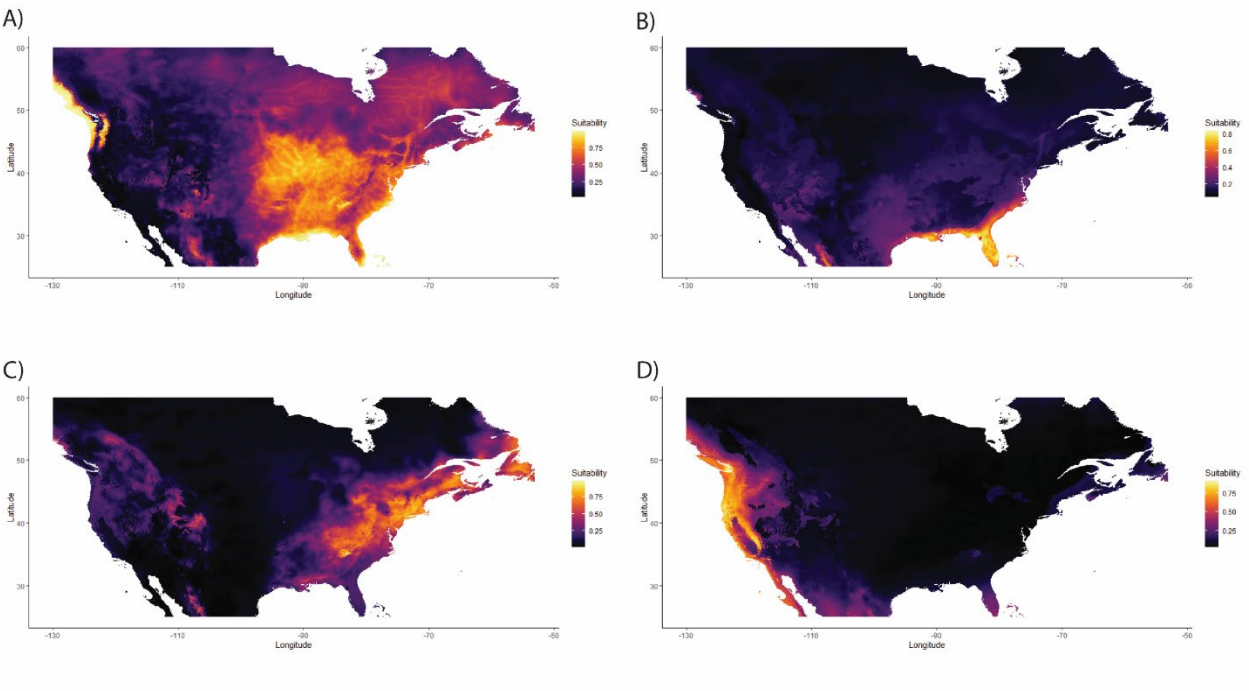


Table 1. Genetic differentiation (mean pairwise FST) between *Thamnophis sirtalis* populations from SNP data using the 'populations' module in STACKS.

Population pair	mean FST (missing data)	mean FST (no missing data)
East – Central	0.011387	0.0053516
East – Southeast	0.020739	0.0081842
East – West	0.022068	0.0095099
Central – Southeast	0.023234	0.0078016
Central – West	0.018548	0.0085273
Southeast – West	0.042005	0.013064

Table 2. Demographic model selection results representing ≥95% of the cumulative Akaike weights (AICω) for all two-way *T. sirtalis* population comparisons. nu1 and nu2 are population sizes before (a) and after (b) an initial split at scaled time T1 (in units of 2N generations) before secondary contact. T2 represents the scaled time between secondary contact and a third epoch. (m) refers to the individual migration rates between populations 1 and 2, with single values reflecting symmetrical migration rates.

Lineage comparison	Model	Log-likelihood	AIC	ΔAIC	AICω	nu1a	nu2a	nu1b	nu2b	m12	m21	T1	T2
Southeast x East	asymmetrical gene flow, pop size change	-437.3	890.6	0	0.9998	6.8903	0.193	28.2727	18.3438	0.1333	0.4828	0.3257	0.8813
Southeast x Central	population size change with symmetrical migration	-305.61	625.22	0	0.7956	0.2953	0.6236	28.2375	12.657	0.3767	-	0.1814	1.0596
	asymmetrical gene flow, population size change	-305.97	627.94	2.72	0.2042	22.9269	0.0522	22.9242	6.2052	0.2472	1.7803	0.2825	0.3398
East x Central	secondary contact, population size change with asymmetrical gene flow	-323.43	662.86	0	1	0.985	0.2433	1.7801	22.7127	19.1886	0.0535	4.6581	0.5535
Central x West	secondary contact with asymmetrical gene flow	-235.13	482.26	0	0.8068	29.1	2.9238	-	-	0.5208	7.1592	0.7952	0.0829
	asymmetrical gene flow, population size change	-234.84	485.68	3.42	0.1459	0.063	14.7271	4.0905	0.4716	0.2422	29.3725	1.5931	0.1314

Table S1. Specimen table for *Thamnophis sirtalis* samples used for ddRADseq. Population information reflects assignments from STRUCTURE analyses. Bolded Institution IDs reflect samples used in species tree analyses. Red coordinates indicate georeferenced centroid of city/county of collection in place of unavailable coordinates or specific locality data.

Institution ID	Population	City/County	State/Province	Country	Latitude	Longitude	Total Reads	% Reads Retained
AMNH_244382	East	Hardy	WV	USA	39.0187	-79.0012	7123470	95.65
AMNH_244398	East	Tuckernuck	MA	USA	41.2833	-70.0999	3762482	88.80
AMNH_244403	Southeast	Jasper	SC	USA	**32.4914**	**-81.0755**	2791886	92.22
AMNH_244405	East	Perry	AL	USA	32.5744	-87.0915	3766052	54.95
AMNH_244426	East	Alleghany	MD	USA	39.5685	-78.9167	16799576	89.24
AMNH_244432	East	Sullivan	NY	USA	**41.6897**	**-74.7805**	8201166	88.65
AMNH_244445	East	Powell	KY	USA	37.8415	-83.7938	1875597	86.46
AMNH_244449	East	York	ME	USA	43.2933	-70.7550	2760738	83.65
AMNH_244553	East	Wayne	PA	USA	41.4147	-75.4823	2968540	88.31
AMNH_244566	Central	Sevier	AR	USA	**34.0346**	**-94.2744**	4165761	86.49
AMNH_244578	East	Peneobscott	ME	USA	44.9662	-68.5419	11483859	69.87
AMNH_244601	Central	Blount	TN	USA	35.6701	-83.7187	23248571	90.60
AMNH_244602	Central	Blount	TN	USA	35.6701	-83.7187	1263059	59.2
AMNH_245144	East	Wilkinstown	AL	USA	31.5724	-86.0025	12538566	97.69
CAS_163970	East	Albany	NY	USA	42.6472	-73.9756	5250281	79.35
CAS_165428	East	Buncombe	NC	USA	35.4667	-82.4894	1033782	55.90
CAS_173177	East	Augusta	VA	USA	38.1740	-79.2559	1971949	84.53
CAS_173179	East	Butler	PA	USA	41.0945	-79.9349	3290774	42.34
CAS_174055	East	Lynchburg	VA	USA	37.4237	-79.2211	4173621	89.21
LSUMZ_1377	East	Boone	KY	USA	39.0285	-84.7227	4165960	92.03
LSUMZ_20353	East	Ascension Parish	LA	USA	**30.1844**	**-90.8593**	4918314	96.58
NBM_1652	East	New Brunswick	CAN	CAN	45.8244	-66.1208	2848844	91.73
NBM_872	East	New Brunswick	LA	CAN	45.8244	-66.1208	9635001	86.57
AMNH_244354	Central	Tuscaloosa	AL	USA	33.0650	-87.5974	6983209	90.67
AMNH_244355	East	Richmond	NY	USA	40.3200	-74.1200	13402765	71.93
AMNH_244356	East	Sullivan	NY	USA	41.5360	-74.5113	13736630	91.66

AMNH_244358	East	Nantucket	MA	USA	**41.2835**	**-70.0995**	4876716	87.38
AMNH_244360	East	Nantucket	MA	USA	**41.2835**	**-70.0995**	13938552	90.06
AMNH_244366	East	Clarke	VA	USA	39.1074	-78.0136	4698214	92.89
AMNH_244369	East	Ulster	NY	USA	41.9605	-74.0575	9855720	86.22
AMNH_244370	East	Delaware	NY	USA	42.0343	-74.8913	4614509	91.19
AMNH_244372	East	Franklin	NY	USA	44.7289	-74.1267	8245870	88.39
AMNH_244375	East	Tuckernuck	MA	USA	**41.2998**	**-70.2557**	2148876	74.53
AMNH_244381	East	Hardy	WV	USA	39.0430	-78.9804	6998433	94.59
AMNH_244383	East	Grant	WV	USA	38.9881	-79.2236	9352462	84.29
AMNH_244386	East	Elkins	WV	USA	38.9246	-79.8576	6127301	86.40
AMNH_244387	East	Tucker	WV	USA	39.0322	-79.5908	8706301	90.07
AMNH_244389	East	Randolph	WV	USA	38.9251	-79.7416	8357309	96.70
AMNH_244390	East	Seneca Rocks	WV	USA	38.8778	-79.3428	19090035	90.69
AMNH_244391	East	Seneca Rocks	WV	USA	38.8627	-79.3658	4052888	91.12
AMNH_244392	East	Parsons	WV	USA	39.0496	-79.7205	8835205	90.42
AMNH_244395	East	Harman	WV	USA	38.9149	-79.5318	10204531	93.70
AMNH_244400	East	Morris	NJ	USA	**40.7083**	**-74.4666**	2652710	84.32
AMNH_244405	Central	Perry	AL	USA	32.5744	-87.0915	2084114	58
AMNH_244407	East	Ulster	NY	USA	41.8301	-74.2912	10140431	88.46
AMNH_244409	East	Burlington	NJ	USA	39.9050	-74.4996	8612032	85.36
AMNH_244410	East	Albany	NY	USA	42.5186	-74.1388	7847877	95.88
AMNH_244413	East	Albany	NY	USA	42.6522	-74.4078	5285326	66.30
AMNH_244416	East	Gosnold	MA	USA	**41.4521**	**-70.8358**	8071413	64.24
AMNH_244418	East	Gosnold	MA	USA	**41.4521**	**-70.8358**	13704857	90.87
AMNH_244419	East	Gosnold	MA	USA	**41.4508**	**70.9227**	9520963	94.96
AMNH_244425	East	Albany	NY	USA	42.5237	-74.1546	5887352	96.63
AMNH_244428	East	Ulster	NY	USA	**41.8586**	**-74.3118**	6972500	86.86
AMNH_244429	East	Ulster	NY	USA	**41.8586**	**-74.3118**	3198281	89.77
AMNH_244431	East	Sullivan	NY	USA	**41.6897**	**-74.7805**	2436434	93.11
AMNH_244435	East	Ulster	NY	USA	41.9205	-74.4191	8561464	92.12
AMNH_244442	East	Sussex	NJ	USA	41.1026	-74.9615	8381827	89.76
AMNH_244554	East	Wayne	PA	USA	41.4147	-75.4823	1678681	90.18
AMNH_244562	East	Albany	NY	USA	42.5528	-74.4297	7777450	87.13

AMNH_244569	East	Coos	NH	USA	**44.7289**	**-71.2787**	4548544	89.31
AMNH_244580	East	Penobscot	ME	USA	44.9662	-68.5419	9862430	93.07
AMNH_244582	East	Lincoln	ME	USA	**44.2292**	**-69.0543**	2120695	93.35
AMNH_244584	East	Litchfield	CT	USA	41.7585	-73.4358	14533085	84.85
AMNH_244586	East	Middlesex	CT	USA	41.4844	-72.5609	1751797	75.76
AMNH_244588	East	Hartford	CT	USA	41.9930	-72.8634	4197040	93.68
AMNH_244591	East	Strafford	NH	USA	**43.3270**	**-71.1842**	1551374	78.13
AMNH_244603	East	Orange	NY	USA	41.2972	-74.2826	9373112	95.32
AMNH_244607	East	Sullivan	NY	USA	41.8248	-74.9640	3938415	97.23
AMNH_244608	East	Sullivan	NY	USA	41.7921	-74.9565	5971080	92.86
AMNH_244610	East	Suffolk	NY	USA	**41.0359**	**-71.9545**	2119526	91.23
AMNH_244612	East	Sullivan	NY	USA	41.5711	-74.4947	3596486	94.54
AMNH_244633	East	Warren	NJ	USA	40.9867	-75.0097	6526993	88.25
AMNH_245145	East	Eufaula	AL	USA	31.9612	-85.2087	5114940	93.32
AMNH_245146	East	Midway	AL	USA	32.0683	-85.4801	10693301	85.19
CAS_162657	East	Watauga	NC	USA	36.2111	-81.6675	5755565	80.20
CAS_165825	East	Albany	NY	USA	42.5131	-74.1919	2685923	52.13
CAS_174056	East	Lynchburg	VA	USA	37.3733	-79.1923	5113479	90.36
CAS_207275	East	Madison	AL	USA	34.6754	-86.5173	2379570	82.33
CAS_208628	East	Talladega	AL	USA	33.3017	-86.0398	4151730	89.13
LSUMZ_1823	Central	East Feliciana Parish	LA	USA	**30.7022**	**-91.2779**	16548441	86.95
LSUMZ_19004	Central	East Baton Rouge Parish	LA	USA	**30.5694**	**-91.0969**	7024920	86.61
LSUMZ_20419	Central	St. Martin Parish	LA	USA	**30.2396**	**-91.7539**	4473781	85.75
LSUMZ_2476	Central	Iberville Parish	LA	USA	30.2899	-91.2333	3112436	93.27
LSUMZ_2779	Central	St. Tammany Parish	LA	USA	30.4776	-90.1042	7769648	95.91
NBM_1651	East	Saint-Quentin	New Brunswick	CAN	47.4292	-66.9100	5861095	96.72
NBM_1687	East	Popple Depot	New Brunswick	CAN	47.4476	-66.5570	7820586	95.74
NBM_865	East	Fernmount	New Brunswick	CAN	45.9854	-66.3950	6794332	96.03
CAS_162699	Central	Jackson	IL	USA	37.6748	-89.5066	2371858	69.20

CAS_162701	Central	Jackson	IL	USA	37.9008	-89.6295	8105774	77.20
CAS_165393	Central	Winnebago	WI	USA	44.0247	-88.5426	3351959	68.87
CAS_165777	Central	Charlevoix	MI	USA	45.6646	-85.5568	6134334	73.36
LSUMZ_7663	Central	Vernon	MO	USA	37.8392	-94.3549	3682392	92.70
LSUMZ_8789	Central	Leavenworth	KS	USA	39.3304	-94.9364	1559345	95.83
TNHC_93511	Central	Brookings	SD	USA	44.5107	-96.4305	7545177	91.47
AMNH_244624	Central	Gladwin	MI	USA	43.8387	-84.2672	4728963	97.16
AMNH_244625	Central	Gladwin	MI	USA	43.8711	-84.2777	5873040	87.95
AMNH_244626	Central	Midland	MI	USA	43.6839	-84.2270	4701690	88.09
AMNH_244630	Central	Midland	MI	USA	43.7936	-84.2643	8871244	94.67
AMNH_244631	Central	Midland	MI	USA	43.7418	-84.2547	1941855	97.17
KU_332315	Central	Lawrence	KS	USA	38.9208	-95.2329	7908512	91.97
KU_332323	Central	Johnson	KS	USA	38.5655	-95.0021	4283366	84.15
KU_337110	Central	Manhattan	KS	USA	39.1305	-96.5639	7446126	90.27
LSUMZ_2702	Central	Ozaukee	WI	USA	43.3884	-87.8773	36171171	97.30
LSUMZ_2704	Central	Washington	WI	USA	30.4553	-91.1581	6426240	91.90
TNHC_86201	Central	Cedar	NE	USA	42.7788	-97.1785	9484928	95.17
AMNH_244437	Southeast	Collier	FL	USA	25.9593	-81.3551	7212509	96.17
AMNH_244558	Southeast	Gulf	FL	USA	29.6919	-85.2878	2337182	81.22
AMNH_245128	Southeast	Waycross	GA	USA	31.0683	-82.2715	5462099	93.17
AMNH_245129	Southeast	Waycross	GA	USA	31.1167	-82.3487	9202169	87.44
AMNH_245131	Southeast	Taylor	FL	USA	29.9953	-83.5642	9687372	73.60
AMNH_245143	Southeast	Munson	FL	USA	30.8575	-86.8710	8279457	95.31
AMNH_245147	Southeast	Cedar Grove	GA	USA	32.2368	-82.9582	6334534	88.18
CAS_162658	Southeast	St Johns	FL	USA	30.2409	-81.4363	5747312	75.00
CAS_172018	Southeast	Walton	FL	USA	30.4586	-85.9206	1817233	86.97
CAS_218692	Southeast	Sumter	FL	USA	28.7370	-82.2246	14126901	95.15
AMNH_245119	Southeast	Waycross	GA	USA	30.9947	-82.31269	2496569	93.19
AMNH_245122	Southeast	Mims	FL	USA	**28.6524**	**-80.9397**	3048584	83.51
AMNH_245125	Southeast	Sebring	FL	USA	27.4129	-81.5903	4757235	95.70
AMNH_245137	Southeast	Sumatra	FL	USA	29.9502	-84.9765	2111573	97.35
AMNH_245138	Southeast	Wilma	FL	USA	30.1054	-84.9862	4239840	81.85
AMNH_245141	Southeast	Liberty	FL	USA	30.2944	-84.8336	6899750	75.53

AMNH_245142	Southeast	Sopchoppy	FL	USA	30.0273	-84.6206	3570372	91.92
CAS_172023	Southeast	Pinellas	FL	USA	27.8385	-82.7849	2112433	66.42
CAS_207272	Southeast	Gulf	FL	USA	30.0002	-85.1263	3493346	87.06
LSUMZ_8064	Southeast	Pinellas	FL	USA	**27.9659**	**-82.8001**	6427979	80.61
LSUMZ_8226	Southeast	Bay	FL	USA	**30.1766**	**-85.8055**	450817	91.32
CAS_165799	West	Creston	BC	CAN	49.0943	-116.5060	3392967	65.64
CAS_165823	West	Lassen	CA	USA	40.5142	-120.8564	1657501	71.93
CAS_208629	West	Lane	OR	USA	44.1667	-121.9167	6466952	84.50
CAS_208952	West	Monterey	CA	USA	36.0927	-121.4010	2347215	80.30
CAS_209758	West	Plumas	CA	USA	39.7223	-120.6806	4840586	93.57
CAS_209985	West	Nevada	CA	USA	39.3538	-120.5036	3372028	91.27
CAS_160720	West	Vancouver Island	BC	CAN	49.3045	-124.3333	3428254	70.71
CAS_165810	West	Curry	OR	USA	**42.5441**	**-124.2782**	3722984	83.92
CAS_204786	West	Clark	WA	USA	45.6859	-122.7191	936827	76.93
CAS_206416	West	Placer	CA	USA	39.1864	-120.4751	3631117	80.30
CAS_206455	West	Modoc	CA	USA	41.4791	-120.5426	1458628	64.02
CAS_209170	West	Vann	CA	USA	39.2569	-122.9439	4286354	88.90
CAS_214287	West	Santa Clara	CA	USA	36.9428	-121.5546	2374158	92.86
CAS_214289	West	Santa Clara	CA	USA	36.9519	-121.5579	9413723	97.65
CAS_214827	West	Santa Clara	CA	USA	37.1984	-121.7419	1472697	95.22
CAS_214828	West	Santa Clara	CA	USA	37.1878	-121.7498	5773180	93.15
LSUMZ_8139	West	Santa Fe	NM	USA	35.6872	-105.9381	1161313	91.66
LSUMZ_8770	West	Sandoval	NM	USA	**35.6021**	**-106.7235**	6787533	84.48

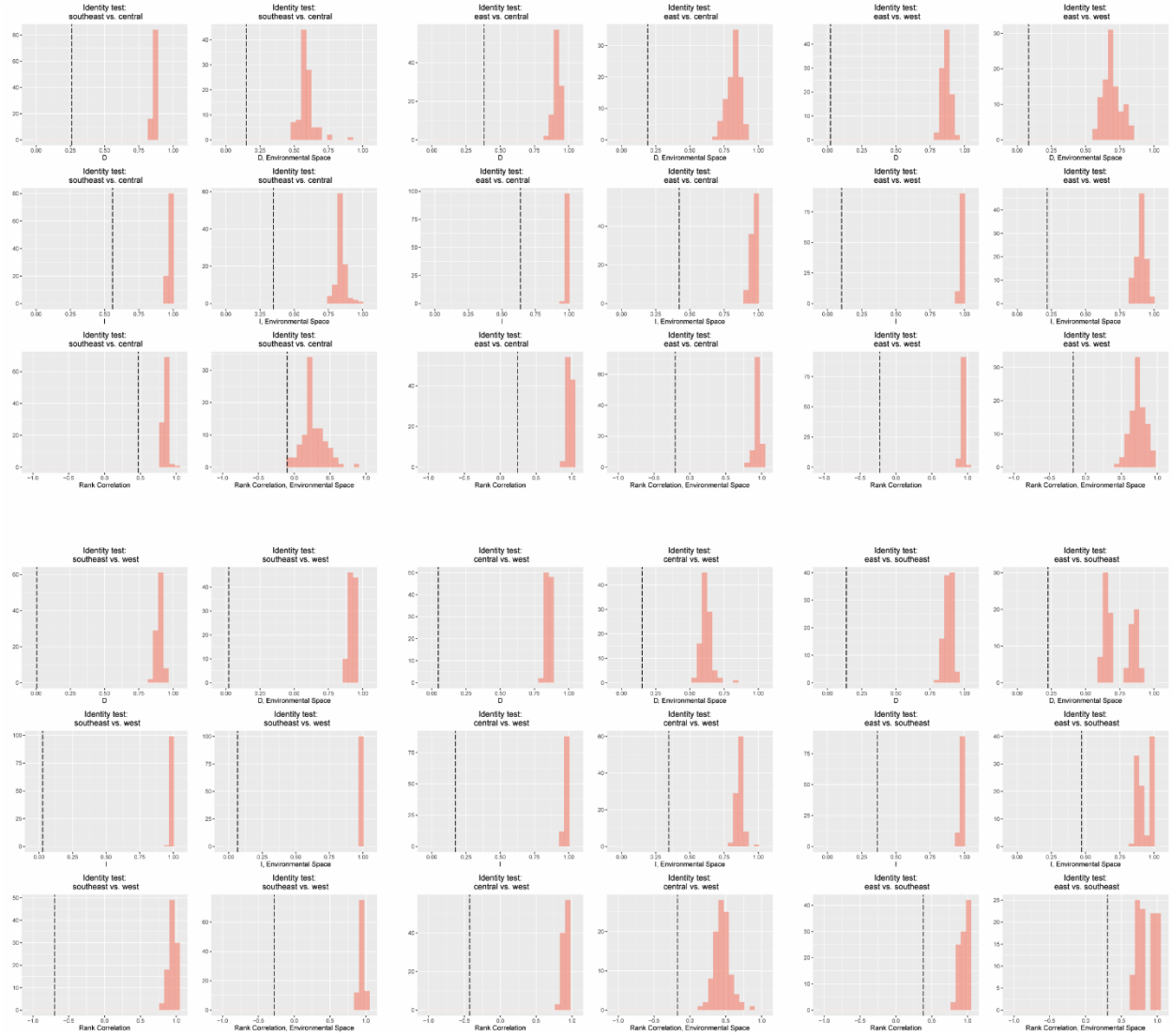
Table S2. Thamnophiine samples used for mtDNA gene tree.

Catalog # / Genbank ID	Species	Source Study
AF420069	*Adelophis foxi*	de Queiroz et al., 2002
AF402917	*Regina septemvittata*	Alfonso & Arnold, 2001
AF420085	*Thamnophis atratus*	de Queiroz et al., 2002
AF420089	*Thamnophis brachystoma*	de Queiroz et al., 2002
AF402923	*Thamnophis butleri*	Alfonso & Arnold, 2001
AF420108	*Thamnophis chrysocephalus*	de Queiroz et al., 2002
AF420103	*Thamnophis couchii*	de Queiroz et al., 2002
AF402924	*Thamnophis cyrtopsis*	Alfonso & Arnold, 2001
AF402925	*Thamnophis elegans*	Alfonso & Arnold, 2001
AF420117	*Thamnophis eques*	de Queiroz et al., 2002
AF420121	*Thamnophis errans*	de Queiroz et al., 2002
AF420125	*Thamnophis exsul*	de Queiroz et al., 2002
AF420129	*Thamnophis fulvus*	de Queiroz et al., 2002
AF420133	*Thamnophis gigas*	de Queiroz et al., 2002
AF420135	*Thamnophis godmani*	de Queiroz et al., 2002
AF420139	*Thamnophis hammondii*	de Queiroz et al., 2002
AF402926	*Thamnophis marcianus*	Alfonso & Arnold, 2001
AF420147	*Thamnophis melanogaster*	de Queiroz et al., 2002
AF420151	*Thamnophis mendax*	de Queiroz et al., 2002
AF420153	*Thamnophis nigronuchalis*	de Queiroz et al., 2002
AF402927	*Thamnophis ordinoides*	Alfonso & Arnold, 2001
AF420161	*Thamnophis proximus*	de Queiroz et al., 2002
AF420165	*Thamnophis pulchrilatus*	de Queiroz et al., 2002
AF420169	*Thamnophis radix*	de Queiroz et al., 2002
AF420173	*Thamnophis rufipunctatus*	de Queiroz et al., 2002
AF420177	*Thamnophis sauritus*	de Queiroz et al., 2002
AF420181	*Thamnophis scalaris*	de Queiroz et al., 2002
AF420189	*Thamnophis scaliger*	de Queiroz et al., 2002
AMNH_244353	*Thamnophis sirtalis*	This study
AMNH_244382	*Thamnophis sirtalis*	This study
AMNH_244398	*Thamnophis sirtalis*	This study
AMNH_244403	*Thamnophis sirtalis*	This study
AMNH_244405	*Thamnophis sirtalis*	This study
AMNH_244426	*Thamnophis sirtalis*	This study
AMNH_244432	*Thamnophis sirtalis*	This study
AMNH_244437	*Thamnophis sirtalis*	This study
AMNH_244445	*Thamnophis sirtalis*	This study
AMNH_244449	*Thamnophis sirtalis*	This study
AMNH_244553	*Thamnophis sirtalis*	This study
AMNH_244558	*Thamnophis sirtalis*	This study
AMNH_244566	*Thamnophis sirtalis*	This study
AMNH_244578	*Thamnophis sirtalis*	This study

AMNH_244590	*Thamnophis sirtalis*	This study
AMNH_244601	*Thamnophis sirtalis*	This study
AMNH_244621	*Thamnophis sirtalis*	This study
AMNH_244628	*Thamnophis sirtalis*	This study
AMNH_245117	*Thamnophis sirtalis*	This study
AMNH_245128	*Thamnophis sirtalis*	This study
AMNH_245129	*Thamnophis sirtalis*	This study
AMNH_245131	*Thamnophis sirtalis*	This study
AMNH_245143	*Thamnophis sirtalis*	This study
AMNH_245144	*Thamnophis sirtalis*	This study
AMNH_245147	*Thamnophis sirtalis*	This study
CAS_162658	*Thamnophis sirtalis*	This study
CAS_162674	*Thamnophis sirtalis*	This study
CAS_162697	*Thamnophis sirtalis*	This study
CAS_162700	*Thamnophis sirtalis*	This study
CAS_163970	*Thamnophis sirtalis*	This study
CAS_165393	*Thamnophis sirtalis*	This study
CAS_165428	*Thamnophis sirtalis*	This study
CAS_165777	*Thamnophis sirtalis*	This study
CAS_165799	*Thamnophis sirtalis*	This study
CAS_165823	*Thamnophis sirtalis*	This study
CAS_165824	*Thamnophis sirtalis*	This study
CAS_165827	*Thamnophis sirtalis*	This study
CAS_172016	*Thamnophis sirtalis*	This study
CAS_172018	*Thamnophis sirtalis*	This study
CAS_173177	*Thamnophis sirtalis*	This study
CAS_173179	*Thamnophis sirtalis*	This study
CAS_174055	*Thamnophis sirtalis*	This study
CAS_199740	*Thamnophis sirtalis*	This study
CAS_206448	*Thamnophis sirtalis*	This study
CAS_208623	*Thamnophis sirtalis*	This study
CAS_208629	*Thamnophis sirtalis*	This study
CAS_208952	*Thamnophis sirtalis*	This study
CAS_209101	*Thamnophis sirtalis*	This study
CAS_209583	*Thamnophis sirtalis*	This study
CAS_209758	*Thamnophis sirtalis*	This study
CAS_209985	*Thamnophis sirtalis*	This study
CAS_210025	*Thamnophis sirtalis*	This study
CAS_214852	*Thamnophis sirtalis*	This study
CAS_218692	*Thamnophis sirtalis*	This study
LSUMZ_1377	*Thamnophis sirtalis*	This study
LSUMZ_14723	*Thamnophis sirtalis*	This study
LSUMZ_1732	*Thamnophis sirtalis*	This study
LSUMZ_20353	*Thamnophis sirtalis*	This study
LSUMZ_7663	*Thamnophis sirtalis*	This study

LSUMZ_8141	*Thamnophis sirtalis*	This study
LSUMZ_8770	*Thamnophis sirtalis*	This study
LSUMZ_8789	*Thamnophis sirtalis*	This study
LSUMZ_8839	*Thamnophis sirtalis*	This study
NBM_1652	*Thamnophis sirtalis*	This study
NBM_872	*Thamnophis sirtalis*	This study
TNHC_93511	*Thamnophis sirtalis*	This study
UWBM_8302	*Thamnophis sirtalis*	This study
AF420197	*Thamnophis sumichrasti*	de Queiroz et al., 2002
EF417400	*Thamnophis validus*	de Queiroz & Lawson, 2008
EF417404	*Thamnophis validus*	de Queiroz & Lawson, 2008
EF417407	*Thamnophis validus*	de Queiroz & Lawson, 2008

Fig S1. Niche identity tests for each pairwise comparison for the 4 inferred *T. sirtalis* lineages using ENMtools.



Chapter 3: Phylogenomics of the cobras: Mixed models and markers resolve a recent radiation

Leonard N. Jones II[1], Adam D. Leaché[1], Wolfgang Wüster[2]

[1]Department of Biology & Burke Museum of Natural History and Culture, University of Washington, Seattle, WA, USA 98195

[2]Department of Biology, Bangor University, LL57 2DG, UK

INTRODUCTION

Phylogenetic inference seeks to disentangle the evolutionary relationships between living organisms to give context to their trait development and spatial distribution patterns. In recent radiations, a lack of accumulated genetic variation, ongoing gene flow, and gene tree discordance due to incomplete lineage sorting (ILS) can complicate phylogenetic inference (Maddison 1997; Nichols 2001; Edwards 2009). Furthermore, varying evolutionary rates among species complexes and unsampled lineages can further confound phylogenetic estimates (Huson and Bryant 2006; Heath et al. 2008; Nabhan and Sarkar 2012). Most recently, evolutionary biologists have benefitted from the development of 1) reduced representation sequencing methods to capture more of the genome, and 2) breakthroughs in species tree estimation under coalescent models more amenable to biological processes like ILS and introgression. However, these advances have also further complicated reaching consensus on many species relationships, since different genetic markers might carry different phylogenetic signals and lead to different conclusions regarding the origin of species and their characteristics.

Over the last decade, targeted sequence capture of ultraconserved elements (UCEs) (Faircloth et al. 2012) and double digest restriction associated DNA (ddRADseq) (Peterson et al. 2012b) arose as the most popular methods for generating reduced-representation genomic libraries for phylogenetic and population genetic inference. A body of work and debate comparing their utility at different temporal depths has been generated since their development. Multiple studies have found sufficient phylogenetic signal in the longer loci resulting from UCE capture to resolve relationships at deeper evolutionary timescales, while the generally larger number of short more variable loci collected by ddRADseq has enabled estimates of shallow phylogenies and population structure (Smith et al. 2014; Leaché et al. 2015; Manthey et al. 2016). Some work suggests an under-appreciated utility of UCEs for shallow systematics as well (Harvey et al. 2016). More recently, UCE targeting has proven beneficial for extracting data from formalin-fixed specimens (Ruane and Austin 2017).

The substantially greater number of loci acquired by sequence capture and ddRADseq methods has also led to advances in integrating the coalescent model into species tree estimation methods. Work comparing the utility of coalescent methods to concatenation, revealed phylogenetic

inconsistency (Kubatko and Degnan 2007) and sensitivity to parameterization (Leaché et al. 2015). Bayesian exploration of parameter space has been used to alleviate these downstream effects but the computational cost of integrating over all gene trees has restricted its implementations to small systems like populations and species complexes (Flouri et al. 2018). This has fueled considerable argument over whether the multispecies coalescent models population structure or delimits species (Jackson et al. 2017; Sukumaran and Knowles 2017; Leaché et al. 2019). Earlier models also operated under an assumption of no gene flow, but newer methods allow for introgression and more complex substitution models (Flouri et al. 2018).

Due to cost and computational constraints, joint applications of these reduced representation datatypes and coalescent methods to larger taxonomic groups consisting of multiple species complexes remain underrepresented in the literature. The thousands of short loci often recovered by ddRADseq make it amenable to resolving relationships within young radiations but equally prone to data loss due to polymorphisms at restriction site across divergent taxa resulting in loss of potentially informative haplotypes and subsequent underestimation of genetic diversity (Arnold et al. 2013; Andrews et al. 2016)(Andrews et al. 2016)(Andrews et al. 2016). Conversely, the longer loci potentially interrogated by sequence capture confer more complete datasets, but analysis of multiple informative sites within these longer loci can violate the assumptions of tight linkage/no recombination inherent to many commonly used coalescent methods (Bryant et al. 2012; Chifman and Kubatko 2014). Concatenation has been incorporated to get around this barrier but risks introducing inconsistent divergence time estimates that can drastically impact accurate divergence time estimations and tree topology, consequently affecting our understanding of diversification timing and mode in recent radiations and older divergences alike (Song et al. 2012; Stange et al. 2018). Larger taxonomic groups with divergent species complexes offer a unique opportunity to jointly compare divergence time estimates from reduced representation datatypes as well as between coalescent and concatenation methods.

Study system

The cobras are among the most well-known snakes in the world due to their elaborate arsenal of antipredator defenses—they are most easily identified by their iconic elaborate hooding displays,

but it is their bites that make the group one of considerable medical importance throughout their distribution. The core cobras and their close relatives significantly contribute to the near 2.5 million envenomations and > 100,000 fatalities throughout Africa and Asia, leading to the World Health Organization's recent re-designation of snakebite as a critically neglected tropical disease (Longbottom et al. 2018).

Cobras, in the broad sense, comprise multiple genera distributed throughout Africa and Asia. In a study seeking to resolve the global elapid relationships, Slowinski and Keogh first identified the "core cobra" clade (Slowinski and Keogh 2000), featuring 33 currently recognized species of the "true" cobra genus *Naja* found throughout Africa and Asia and three species-poor (1-3 species) genera found in East and Southern Africa and the Middle East. These genera, the shield-nosed cobras *Aspidelaps spp*, rinkahls *Hemachatus haemachatus*, and black desert cobra *Walterinnesia aegypti* exhibit hoods that are not as morphologically developed as those observed in *Naja.* The South/Southeast Asian king cobra *Ophiophagus hannah*, the most well-known of the cobras, was found to be an unambiguous distant relative to the core group and nested within a species poor Afro-Asian clade containing the African mambas *Dendroaspis*, the African garter snakes *Elapsoidea*, and the Southeast Asian barred coral snakes *Hemibungarus* (Slowinski and Keogh 2000; Castoe et al. 2007). Recent work has further validated this relationship with moderate to high support (Pyron et al. 2011, 2013a, 2013b; Lee et al. 2016). Considerable attention has also been given to resolving the relationships within regional species complexes of the core group where phylogeographic breaks have conferred recently timed divergences and led to substantial taxonomic revision within *Naja* (Wüster et al. 2007, Wüster and Broadley 2007, Wallach et al. 2009, Trape et al. 2009, Ceríaco et al. 2017, Wüster et al. 2018).

The most comprehensive phylogenetic investigation of the cobras to date sought to contextualize the development of the defensive spitting apparatus observed in multiple Asian and African species in a geographic context (Wüster et al. 2007). Understanding the evolutionary origins of this trait are of interest due to both medical interest and the morphological, behavioral, and toxicological complexity of its form: the efficacy of this trait is contingent on co-evolution of 1) anatomical changes to the delivery system (e.g. fang modifications that facilitate upwards venom expulsion towards a would-be threat), 2) conservation of the spitting behavior itself, 3)

functional changes in venom proteolytic complexes that cause loss of vision (Wüster and Thorpe 1992, Young 2004, Yap et al. 2011). The spitting cobras' associations with grassland and arid environments served as a model trait to test hypotheses regarding the evolution of the spitting defense in response to large mammals that accompanied the rise of grassland formations throughout Africa. Divergence time estimation of the African and Asiatic *Naja* radiations using mtDNA inferred asynchronously timed colonization events of the *Naja* and the large land mammals purported to trigger the spitting response, leading Wüster et al to conclude that defensive spitting did not evolve as an adaptation to their appearance. However, an understanding of the relationships between and within these radiations have evaded consensus resolution due to the recency of the diversification events and the lack of genomic resources for many species. Notable variation in defensive spitting morphology and behavior exists within *Naja* and between *Naja* and *Hemachatus*. The spitting Asiatic *Naja* are generally distributed throughout mainland and insular Southeast Asia. Spanning both South and Southeast Asia is *Naja kaouthia*, with a distribution spanning eastern India through mainland Southeast Asia and the Malay Pensinsula. Spitting behavior has recently been documented in eastern Indian *N. kaouthia* populations which otherwise do not exhibit the specialized dentitional modifications to facilitate venom ejection (e.g. reduced venom exit orifice) found in the obligate spitting species of Southeast Asia (Santra and Wüster 2017). In the easternmost part of its distribution, it exists in sympatry with *Naja atra*, which also exhibits population-specific non-specialized spitting adaptations (Gang et al. 2011). The taxonomic placement of both species within the Asiatic *Naja* has varied considerably among the studies in which it has been included due in part to differences in sampling and genomic data (Slowinski and Keogh 2000; Wüster et al. 2007; Lee et al. 2016; Zaher et al. 2019). This ambiguity raises the question of whether the underdeveloped condition of spitting apparatus in *Naja kaouthia* represents an intermediate phenotype, a partial loss, or gain of function independent of that of the specialized Southeast Asian spitters. Resolving the relationships and inferring accurate divergence times within this radiation might characterize the tempo and mode of the expansion of *Naja* into Asia from Africa as well as the evolution of the spitting apparatus within the group.

Here we reevaluate the evolutionary relationships of the cobras and closely related elapids (family Elapidae) from targeted sequence capture of UCEs and SNPs called derived from

ddRADseq data. We assess the stability of divergence times from prior mtDNA and concatenation-based studies by estimating clade ages under the multispecies coalescent model. Specifically, we evaluate the age and topology of the Asiatic *Naja* radiation. We use phylogenetic inference of both reduced representation methods to 1) attempt to resolve the phylogenetic relationships of the cobras *sensu lato* (*Aspidelaps*, *Hemachatus*, *Walterinnesia*, *Naja*) 2) determine wether the African origin for *Naja* precedes multiple Asian colonization events (Minton 1986, Ineich 1995) 2) evaluate the ancestral state and origin of the defensive venom spitting apparatus.

MATERIALS & METHODS

Sampling and Genomic Data Generation

We assembled field and museum-collected tissues representing the core cobra group, closely related elapids, and distantly related colubrids (*Thamnophis sirtalis* and *T. ordinoides*, family Colubridae) to ensure accurate divergence time estimation (Table 1) and isolated gDNA using NaCl extraction. We assessed extracted DNA quality via gel electrophoresis and quantity with a Qubit dsDNA BR assay (Life Technologies Inc.) prior to library preparation.

UCE capture and bioinformatics

We sonicated ~500ng gDNA to a target peak of 400bp using a Bioruptor Pico (Diagenode Inc. # B01060001) and prepared libraries using an Illumina Truseq Nano HT library preparation kit (Illumina Inc. # FC-121-9010DOC). We used probes targeting a total of 585 loci sourced from both a library of ~5,000 vertebrate UCEs and 44 additional genes from the squamate Tree of Life project (Wiens et al. 2012). We incubated the RNA probes and samples for 24 hours in a blocking mixture of forward and reverse compliments of the Illumina Truseq Nano Adapters with indices replaced by inosines, and chicken-derived competitor DNA (Chicken Hybloc, Applied Genetics Lab Inc.) to reduce non-specific and repetitive binding. We enriched hybridized libraries with Truseq adapter primers and Phusion Taq polymerase (New England Biolabs Inc. # M0530) for 20 cycles using standard protocols and cleaned the resulting PCR

product with SeraPure Sera-Mag Speedbeads. We quantified the final libraries using quantitative polymerase chain reaction (qPCR) (Applied Biosystems Inc.) with primers targeting five loci mapping to different target regions in the *Anolis* genome from which the probe subset was generated (see DISCUSSION). We qualified the resulting libraries on an Agilent TapeStation 2200 (Agilent Tech.). We pooled samples in equimolar ratios to sequence 150bp paired end reads on an Illumina HiSeq4000 lane at the QB3 facility at UC Berkeley.

We demultiplexed raw sequences with Casava (Illumina), trimmed off low-quality sequence ends and residual adapter sequences with Trimmomatic (Bolger et al. 2014). The cleaned paired reads were organized by individual and then assembled with the de novo assembler ABYSS (Simpson et al., 2009) with k-mer values of 60. We used phyluce (Faircloth 2016) to assemble loci across species and aligned species-specific assemblies to the probe sequences using the program LASTZ (Harris 2007). After creating an SQL relational database of assembly-to-probe matches for each species, we queried the database for loci that were shared for a minimum of three species across all samples, and for those that were present across all species. We performed multiple sequence alignments for each locus using MAFFT (Katoh and Standley 2013) and trimmed ragged ends to reduce missing or incomplete data.

To increase our sampling scheme and reliability of downstream phylogenetic estimates and divergence time estimation, we bioinformatically extracted targeted loci from multiple samples from additional studies (Vonk et al. 2013, Kishida et al. 2019, Suryamohan et al. 2020, Ullate-Agote et al. 2015) made publicly available on the Sequence Read Archive representing key ingroup and outgroup species to anchor the phylogeny.

ddRADseq library preparation and bioinformatics

From a reduced set of samples for which there remained sufficient gDNA, we generated ddRADseq libraries. We first digested 500-1000 ng DNA with SbfI and MspI (New England Biolabs Inc # R3642, R0106) and ligated sample-specific barcodes to individual samples. Samples ligated to unique barcodes were then pooled together in groups of up to 8, and pools were size selected for fragments in the range of 415-515 bp with a Pippin Prep system (Sage

Science Inc). We amplified the size-selected products with a Phusion High Fidelity Taq polymerase kit (New England Biolabs Inc # M0530) with Illumina primers that introduce unique multiplexing indices to each pool for 30 cycles at 51º C annealing temperature. Final libraries were also with an Agilent 2200 Tapestation before sequencing across 3 Illumina HiSeq 4000 lanes (50-bp single end reads) at UC Berkeley QB3.

We processed raw reads using ipyrad v0.9 (Eaton and Overcast 2016). We first discarded reads with a quality score limit below 33 or with a single base pair difference in adapter and barcode sequences. We also applied a 5 percent maximum threshold of uncalled bases per locus, and discarded loci exhibiting low coverage (read depths < 6) or excess diploid haplotype diversity (>2). After filtering, we generated a reference-based assembly guided by alignment to the *Ophiophagus hannah* reference genome, one of the most complete elapid genomes to date (Accession: GCA_000516915.1, Vonk et al. 2013). We generated datasets consisting of loci present in 50% of the samples to account for the potentially high data attrition that can result from overly stringent missing data thresholds in datasets with highly divergent taxa. Prior to phylogenetic analyses, we generated reduced UCE and ddRADseq-derived SNP datasets featuring the same samples to enable comparison between phylogenies inferred by the two methods.

Phylogenetic inference and divergence time estimation using concatenation

We estimated divergence times within the cobras using concatenation of UCE data due to the size and taxon coverage of the dataset. The pitfalls stemming from concatenation-based methods are well-documented (Kubatko & Degnan 2007, Degnan & Rosenberg 2009), but the species and loci numbers here preclude inference with modern multispecies coalescent-based methods, which remain computationally intractable for taxon sampling schemes of this size. To account for substitution rate variation among the UCEs, we selected partitioning schemes using ParitionFinder2 (Lanfear et al. 2016), from the data using candidate partitions representing each UCE. We first evaluated the information content within 3 datasets representing 50%, 75%, and 95% of taxon coverage for each locus to address the tradeoff between information content and data completeness and its impact on downstream analyses in divergent groups (Streicher et al.

2016). We determined the data from the 75% and 95% thresholds to not contain enough information to differentiate between highly divergent taxa and consequently selected the 50% threshold for analysis. We first estimated a Maximum Likelihood starting tree from the data using RaxML (Stamatakis 2006). We compared GTR, GTR+G, and GTR+G+I substitution models, and searched for partitioning schemes for the loci under the fast relaxed clustering algorithm ('rcluserf'). We ranked the fit of the substitution using BIC. The highest ranked partitioning schemes were implemented in BEAST v1.10 (Drummond and Rambaut 2007). We used a clade of colubrid species *Thamnophis sirtalis, Thamnophis ordinoides*, and *Pantherophis guttatus* as a distant outgroup clade and estimated the tmrca of *Naja* and Elapidae constraining the monophyly of the latter. We applied a gamma distribution prior (shape=40; scale=1, offset=0) to the tmrca(Elapidae), a slight relaxation of constraints from previous studies (Lee et al. 2016) based on fossil records of caenophidians found in ~34-37.2 MA deposits (Head et al. 2005, Sanders et al. 2008). Similarly, we constrained the divergence of *Thamnophis sirtalis* and *Thamnophis ordinoides* using a wide normal prior (mean=14; SD=1), based on the earliest known *Thamnophis* fossil from the Medial Barstovian (13-14.5 MA) (Voorhies 1990). We conducted duplicate Bayesian MCMC analyses with randomized starting seeds for 50,000,000 cycles, sampling every 1000. We assessed convergence of parameters and congruence between the different runs in Traver v1.7 (Rambaut et al. 2018).

Trait evolution

To characterize the evolution of spitting ability, we coded all species in the UCE dataset (the most comprehensive dataset for the core cobra group) for spitting ability as present or absent based on the literature, and applied an additional intermediate character designation to *Naja kaouthia*, *N. atra,* and *Hemachatus haemachatus* due to their non-specialized dentition properties and heterogeneity in spitting behavior (Wüster and Thorpe 1992; Young 2004; Santra and Wüster 2017). We fit the equal rates ("ER"), all rates different ("ARD"), and symmetric rates ("SYM") models to the tree and determined the best model by AIC ranking using the R package 'ape' (Paradis et al., 2004). We also performed stochastic character mapping (Huelsenbeck et al., 2003) by summarizing a distribution of 1000 character maps applied to the BEAST phylogeny under the selected model using the R package 'phytools' (Revell 2012).

Phylogenetic inference: UCE vs ddRADseq comparison

For the most comprehensive UCE and SNP datasets, we inferred phylogenetic relationships using SVDQuartets (Chifman and Kubatko 2014) implemented in PAUP* 4.0 (Swofford 2003). This method estimates an unrooted tree by evaluating all possible quartets of taxon relationships from the data under the coalescent model. An advantage of this method here is its applicability to large genome-wide datasets, for which traditional Bayesian MCMC methods do not scale well regarding computational power and time requirements. We analyzed the individual complete UCE and SNP datasets as well as two datasets reduced to samples shared between the two. We evaluated all possible quartets and assessed uncertainty of species relationships with 100 bootstrap replicates. We grouped species represented by multiple individuals and putative species complexes (i.e. *Naja melanoleuca*) to single taxa to reduce computation time.

Resolution of Naja subclades using the MSC

For the *Naja melanoleuca* species complex and the Asiatic *Naja*, we inferred relationships under the MSC model using BPP v4.2 (Flouri et al. 2018). To start, we generated two UCE and SNP datasets for each clade, one with *Ophiophagus hannah* as an outgroup and one without. For *N. melanoleuca*, we partitioned individuals by country/locality (Ghana, Senegal, São Tomé and Príncipe, Cameroon, and Congo). We included *Naja subfulva* and *Naja annulata* due to their established proximal relationships to *N. melanoleuca* (with the former rendering *N. melanoleuca sensu lato* paraphyletic). Asiatic *Naja* were partitioned by species. We analyzed all datasets with and without *Ophiophagus hannah* as distant outgroups in analyses to verify stability of inferred relationships under varied coalescent depths. We ran species tree estimation (A01) and then parameter estimation (A00) analyses for each dataset, using the best tree inferred by the former as a fixed topology required by the latter. A01 analyses require a starting tree as a prior, so we used a summary tree from 100 bootstrap replicates of an SVDQuartets analysis for each subclade dataset. We assumed a GTR substitution model and applied inverse gamma (α, β) priors $\theta \sim (3,0.0075)$ and $\tau \sim (3, 0.025)$. We ran the analysis for 50,000 generations, sampling every 2, with 5000 discarded as burn in. We conducted four replicate runs under different starting seeds for

each analysis to assess convergence, and the MCMC of each run were combined and re-summarized to produce the fixed species tree topology required as input for the A00 analysis. We ran A00 analyses with these relationships fixed under the same priors and sampling rate and burnin percentage over a longer MCMC chain of 100,000 iterations to account for the joint estimation of more individual parameters. Replicate runs were combined and summarized to maximize ESS values.

RESULTS

Genomic data

Detailed summary statistics for UCE and ddRADseq datasets are shown in **Table 1**. Under the 50% missing data threshold (where 50% of all individuals must be represented in a locus), >95% of the 585 targeted loci were retained (per sample average UCEs retained = 313) across all UCE datasets, with 10,119 informative sites in the complete dataset. The UCE datasets are consistent in their capturing most of the targeted loci (95-98%) and mean loci/sample but vary widely in information content as a function of the inclusion/exclusion of *Ophiophagus hannah*. The ddRAD datasets expectedly contained substantially more loci and informative sites---contained an average of 8090 retained loci (per sample average = 2720), with the complete dataset containing 3,973 unlinked SNPs.

Phylogenetic relationships and divergence times: concatenation

The time-calibrated BEAST tree estimated from concatenated UCE data (**Figure 2**) supports a late Eocene split between Elapidae and Colubridae (the latter represented by *Thamnophis*) timed at ~44 MY (95% HPD: 29.24-57.66) in agreement with the most recent comprehensive evolutionary investigation of advanced snakes (Zaher et al. 2019). We recover a moderately supported (PP=0.88) ~37 MY split between the ancestor of the core cobra group and that of the Australasian elapids (*Bungarus*, *Demansia*, *Acanthophis*, *Notechis, Pseudechis, Oxyuranus Micropechis, Laticauda*). This estimate recapitulates findings of prior studies with broader elapid representation (Lee et al., 2016), but is substantially younger than prior estimates from mtDNA

with targeted core cobra group sampling (Wuster et al., 2007). We find moderate support (PP = .84) for the sister relationship of the African and Asiatic *Naja* clades, and a less robustly supported (PP = .80) sister relationship of the rarely sampled *Pseudohaje* to *Naja.* Generally, concatenated UCEs fail to resolve the deeper relationships within the core cobras, with low support for the phylogenetic placements of *Hemachatus*, *Walterinnesia*, and *Aspidelaps*.

Trait evolution

Results of ancestral state reconstruction of defensive spitting are shown in **Figure 3**. Maximum likelihood and stochastic character mapping analyses were highly correlated for spitting origins, (Pearson's $r = 0.999$, $P = 2.2 \times 10^{-16}$). The "ER" model was the highest ranked model by AIC in the maximum likelihood analysis. Under this model, stochastic character mapping between the three spitting states unambiguously support multiple origins of venom spitting, largely in agreement with the prevailing hypotheses of independent development of the apparatus in the Asiatic, African, and non-*Naja* cobra lineages.

Phylogenetic relationships under the coalescent model: UCEs vs SNPs comparison

Phylogenetic relationships inferred from UCE and SNP datasets using SVDQuartets are shown in **Figures 4** and **5**, respectively, and a comparison between analyses of the two datasets reduced to their shared samples is shown in **Figure 6**. Analyses of both datatypes support an Asian origin for the elapids and the monophyly of the core cobra group, reflecting the consensus of prior genus and family level investigations (Wüster et al. 2007, Lee et al. 2016, Slowinski & Keogh 2000, Pyron et al. 2011) and the tree inferred from the concatenated UCE data, but differ widely in species relationships within Asiatic *Naja*. The UCE topology reflects relationships historically inferred from few loci or exclusively mtDNA, while the SNP tree topology implausibly places *Hemachatus* and *Dendroaspis* within *Naja* with low support.

Bayesian resolution of Naja subclade relationships under the coalescent model

Results from the BPP A01 and A00 analyses of the *N. melanoleuca* species complex and the Asiatic *Naja* are shown in **Figures 6** and **7**, respectively. The UCE and SNP trees for *N. melanoleuca* differ between each other substantially in topology and support with the addition of *Ophiophagus hannah* as an outgroup, with the tree estimated from the SNP dataset excluding *O. hannah* completely resolving the relationships within the species complex (**Figure 7A**). Relationships inferred from this tree are like those of the only population-level investigation of the group to date (Wuster et al., 2018). Among the Asiatic *Naja*, relationships within UCE trees with and without *Ophiophagus hannah* are incongruent in topology, with both featuring weakly supported and implausible relationships in trees at nearly all nodes. By contrast, the trees estimated from SNP data recover the same fully resolved tree topology (PP = 1.0 for all relationships) regardless of of the presence of *O. hannah*. The SNP topology for the Asiatic *Naja* supports three clades: a South and Central Asian clade containing the westernmost occurring species (*Naja naja* and *N. oxiana*,, respectively), a Southeast Asian clade featuring all specialized spitting species (*N. mandalayensis, N. philippinensis, N. samarensis, N. sputatrix, N. sumatrana, N. siamensis*), and a third clade ranging from East India to Southern China consisting of the intermediate spitting species *N. kaouthia* and *N. atra* and non-spitting *N. sagittifera*.

DISCUSSION

Here we present the most comprehensive phylogeny of the cobras and their close relatives to date in coverage of species and genomic data. We present multiple lines of evidence from which a consensus relationship can be drawn, and rather than weigh the merits of sequence capture of UCEs and ddRADseq against each other, we exploit their strengths to address longstanding hypotheses across multiple evolutionary timescales. Expectedly, in our analysis of concatenated UCE loci we report similar but poorly supported relationships between the non-*Naja* members of the core cobra group found in prior investigations, possibly representing a case where more data introduces more uncertainty to the tree at this depth (Felsenstein 1978). In our coalescent analyses UCEs and SNPs produced substantially different phylogenies, with the former failing to resolve relationships within the Asiatic *Naja* and the latter strongly supporting multiple implausible relationships within the core cobras. The Bayesian BPP analyses of subset Asiatic *Naja* and the *N. melanoleuca* complex data point to the strengths of interrogating shallow scale

relationships from SNPs and the occasional lack of sufficient phylogenetic signal in UCEs to infer shallow divergences between species. We consider sampling, specifically outgroup choice, in our shallow coalescent analyses and reveal the stability of tree topology using SNPs compared to UCEs across outgroup inclusion/exclusion analyses.

The multispecies coalescent and stability of phylogenetic relationships

Following the growing body of work evaluating phylogenetic estimation from SNP-based and UCE capture methods, we encounter similar phylogenetic disagreement as observed in prior squamate investigations (Leaché et al. 2015). In search of the "true" relationships between the cobras we find a consistent mix of substantial disagreement within *Naja* but persistent stability of the subgenera *Boulengerina*, *Afronaja*, Asiatic *Naja*, and *Ureaus*, echoing the perspective of support for subgenera designation in taxonomic research in medically important systems (Wallach et al. 2009). It is reasonable to assume that a continued push towards greater genomic representation and the development of computational tools to feasibly analyze them might resolve some of the problematic relationships observed. We acknowledge that a caveat of this study's interrogated loci is the use of a subset of UCEs originally optimized for phylogenetic analysis of Iguanian lizards by mapping to *Anolis* and *Sceloporus* genomes (Leaché et al. 2015) to infer relationships within elapid snakes, which show extremely high net diversification rates among reptiles (Sanders et al. 2008, Lee et al. 2016). Despite that bias, the unrooted relationships inferred from UCE and SNP data under the coalescent model reach agreement in multiple small-scale divergences within both African and Asian radiations. The phylogeny of the expanded and concatenated UCE data inferred by BEAST supports the monophyly of all cobra subgenera, but slightly deviate from the estimates under the coalescent. Subgenera *Afronaja* plus *Boulengerina* and *Naja* and *Uraeus* form two weakly supported clades. In agreement with the coalescent tree, *Dendroaspis* and the core cobra group form a monophyletic clade to which *Ophiophagus* is sister with moderate support. These analyses failed to reach convergence after a combined 70,000,000 cycles with low ESS values for parameter estimates related to the core cobra clade divergence times in particular. However, mean estimates for tmrca(Elapidae) and tmrca(*Naja*) fall within range of the results of all recent from which these calibration priors came, and further substantiate a middle and late Eocene origin for the elapids and *Naja*, respectively. Taken

together with the time calibrated concatenation analysis of the UCE data, we find strong support for relationships within historically inferred cobra subclades previously estimated from orders of magnitude fewer loci, and we conclude that both methods add valuable insights into a historically taxonomic moving target.

Phylogenetic status of spitting

Generally, our analyses of the core cobras and their close relative support the hypothesis of three separate origins of spitting within the elapids (Wüster and Thorpe 1992). In resolving the relationships within the Asiatic *Naja*, we provide geographic context to this hypothesis and find support for a model of eastward expansion from Central Asia, with spitting developing in the peninsula-bound SE Asian lineage. A sister clade to that lineage comprises two "intermediate" spitters *N. atra* and *N. kaouthia* and the non-spitting *N. sagittifera*. Under this topology one could parsimoniously interpret this clade as a descendant of an ancestor containing a precursor phenotype to that of the more obligate spitters. Within the broader core cobras, the placement of *Hemachatus haematchatus* relative to *Naja* evades consensus.

Following our conclusion of continued support for the sister group relationship of the African and Asiatic *Naja* clades, we examine the evolutionary origins of defensive spitting in the core cobra group and morphological characteristics within the Asiatic *Naja* radiation. Within this group, SNP data strongly support (PP = 1.0) a sister relationship containing a clade of obligate spitting cobra species and a clade featuring three species that exhibit a spectrum of behavioral and morphological adaptations for defensive spitting (from none to population-specific intermediate spitting). Our ancestral state reconstruction results support an ancestral state of specialization, rendering the distribution of non- to intermediate spitting species *N. sagittifera*, *N. kaouthia*, and *N. atra* descendants of a reversal event. Alternatively, one could conclude from the distribution of traits across the fully supported Asiatic *Naja* SNP tree as one of stepwise progression from a lack of spitting (*Naja oxiana* and *Naja naja*) to obligate defensive spitting exhibited in the Southeast Asiatic *Naja*, with the intermediate clade representing the phenotypes of transitional forms. Assuming the outgroup relationship of the monotypic *Hemachatus haemachatus* to *Naja*, we echo the prevailing hypothesis in interpreting the development of

defensive spitting in *Hemachatus* as independent due to the lack of specialization as observed in African *Naja*. The uncertainty around the development of the spitting apparatus points to the need for a more detailed and integrative characterization of the spitting apparatus. Here we reduce the trait to three broad categories by consideration of dentition and a binary measure of ubiquity of the behavior in the species, but have not included similarity in venom composition among species as a metric for spitting specialization and do not have population level sampling for species with population level variation of spitting behavior. Furthermore population-level resolution of the spitting apparatus and more population genomic-focused studies of the species that contain them might uncover a role of hybridization in the trait's distribution. Little is known about the barriers to gene flow or lack thereof between the more regionally confined Asiatic *Naja* due in part to the paucity of modern genomic resources phylogenetic status of the for the group. The relationships within Asiatic *Naja* are only resolved with one dataset in this study, but in the context of the geographic distribution and the spectrum of spitting apparatus development, we cautiously conclude the relationships inferred by the Asiatic subset of SNP data to be the most plausible. We stress the need for additional research that considers population-level behavioral and morphological variation in spitting clades.

REFERENCES

Alexander Pyron R., Burbrink F.T. 2009. Lineage diversification in a widespread species: roles for niche divergence and conservatism in the common kingsnake, *Lampropeltis getula*. Mol Ecol. 18:3443–3457.

Alfaro M.E., Arnold S.J. 2001. Molecular Systematics and Evolution of Regina and the Thamnophiine Snakes. Molecular Phylogenetics and Evolution. 21:408–423.

Andrews K.R., Good J.M., Miller M.R., Luikart G., Hohenlohe P.A. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. Nature Reviews Genetics. 17:81–92.

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

Avise J.C. 2000. Phylogeography: The History and Formation of Species. Harvard University Press.

Avise J.C., Arnold J., Ball R.M., Bermingham E., Lamb T., Neigel J.E., Reeb C.A., Saunders N.C. 1987. Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annu. Rev. Ecol. Syst. 18:489–522.

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

Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.-H., Xie D., Suchard M.A., Rambaut A., Drummond A.J. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLOS Computational Biology. 10:e1003537.

Boulenger G.A. 1896. Catalogue of the Snakes in the British Museum (Natural History)... order of the Trustees.

Bronikowski A.M. 2000. Experimental Evidence for the Adaptive Evolution of Growth Rate in the Garter Snake Thamnophis Elegans. Evolution. 54:1760–1767.

Brunsfeld S.J., Sullivan J., Soltis D.E., Soltis P.S. 2001. Comparative phylogeography of northwestern North America: a synthesis. Special Publication-British Ecological Society. 14:319–340.

Bryant D., Bouckaert R., Felsenstein J., Rosenberg N.A., RoyChoudhury A. 2012. Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. Molecular Biology and Evolution. 29:1917–1932.

Buckley L.B., Jetz W. 2007. Environmental and historical constraints on global patterns of amphibian richness. Proceedings of the Royal Society B: Biological Sciences. 274:1167–1173.

Burbrink F.T. 2002. Phylogeographic analysis of the cornsnake (Elaphe guttata) complex as inferred from maximum likelihood and Bayesian analyses. Molecular Phylogenetics and Evolution. 25:465–476.

Burbrink F.T., Chan Y.L., Myers E.A., Ruane S., Smith B.T., Hickerson M.J. 2016. Asynchronous demographic responses to Pleistocene climate change in Eastern Nearctic vertebrates. Ecology letters. 19:1457–1467.

Burbrink F.T., Lawson R. 2007. How and when did Old World ratsnakes disperse into the New World? Molecular Phylogenetics and Evolution. 43:173–189.

Burbrink F.T., Lawson R., Slowinski J.B. 2000. Mitochondrial DNA Phylogeography of the Polytypic North American Rat Snake (elaphe Obsoleta): A Critique of the Subspecies Concept. Evolution. 54:2107–2118.

Burbrink F.T., Yao H., Ingrasci M., Bryson R.W., Guiher T.J., Ruane S. 2011. Speciation at the Mogollon Rim in the Arizona Mountain Kingsnake (Lampropeltis pyromelana). Molecular Phylogenetics and Evolution. 60:445–454.

Burnham K.P., Anderson D.R. 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research. 33:261–304.

Calsbeek R., Thompson J.N., Richardson J.E. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. Molecular Ecology. 12:1021–1029.

Castoe T.A., Bronikowski A.M., Brodie E.D., Edwards S.V., Pfrender M.E., Shapiro M.D., Pollock D.D., Warren W.C. 2011. A proposal to sequence the genome of a garter snake (Thamnophis sirtalis). Stand in Genomic Sci. 4:257–270.

Castoe T.A., Smith E.N., Brown R.M., Parkinson C.L. 2007. Higher-level phylogeny of Asian and American coralsnakes, their placement within the Elapidae (Squamata), and the systematic affinities of the enigmatic Asian coralsnake Hemibungarus calligaster (Wiegmann, 1834): RELATIONSHIPS OF CORALSNAKES AND HEMIBUNGARUS. Zoological Journal of the Linnean Society. 151:809–831.

Caye K., Deist T.M., Martins H., Michel O., François O. 2016. TESS3: fast inference of spatial population structure and genome scans for selection. Molecular Ecology Resources. 16:540–548.

Ceríaco L.M.P., Marques M.P., Schmitz A., Bauer A.M. 2017. The “Cobra-preta” of São Tomé Island, Gulf of Guinea, is a new species of Naja Laurenti, 1768 (Squamata: Elapidae). Zootaxa. 4324:121.

Chifman J., Kubatko L. 2014. Quartet Inference from SNP Data Under the Coalescent Model. Bioinformatics. 30:3317–3324.

Darriba D., Taboada G.L., Doallo R., Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature methods. 9:772–772.

Degnan J.H., Rosenberg N.A. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution. 24:332–340.

Deutsch C.A., Tewksbury J.J., Huey R.B., Sheldon K.S., Ghalambor C.K., Haak D.C., Martin P.R. 2008. Impacts of climate warming on terrestrial ectotherms across latitude. PNAS. 105:6668–6672.

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

Earl D.A., vonHoldt B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genet Resour. 4:359–361.

Easterbrook D.J. 1969. Pleistocene chronology of the Puget lowland and San Juan islands, Washington. Geological Society of America Bulletin. 80:2273–2286.

Eaton D.A., Overcast I. 2016. ipyrad: Interactive assembly and analysis of RADseq datasets. Bioinformatics.

Edgehouse M., Latta L.C., Brodie E.D., Brodie E.D. 2014. Interspecific Aggression and Habitat Partitioning in Garter Snakes. PLoS One. 9.

Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution. 63:1–19.

Edwards S.V., Potter S., Schmitt C.J., Bragg J.G., Moritz C. 2016. Reticulation, divergence, and the phylogeography–phylogenetics continuum. Proc Natl Acad Sci USA. 113:8025–8032.

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology. 14:2611–2620.

Faircloth B.C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics. 32:786–788.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved Elements Anchor Thousands of Genetic Markers Spanning Multiple Evolutionary Timescales. Syst Biol. 61:717–726.

Felsenstein J. 1978. Cases in which Parsimony or Compatibility Methods will be Positively Misleading. Syst Biol. 27:401–410.

Fick S.E., Hijmans R.J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology. 37:4302–4315.

Fitch H.S. 1941. Geographic Variation in Garter Snakes of the Species Thamnophis sirtalis in the Pacific Coast Region of North America. American Midland Naturalist. 26:570.

Fitch H.S. 1981. Thamnophis sirtalis. Catalogue of American Amphibians and Reptiles (CAAR).

Flouri T., Jiao X., Rannala B., Yang Z. 2018. Species Tree Inference with BPP Using Genomic Sequences and the Multispecies Coalescent. Mol Biol Evol. 35:2585–2593.

Flouri T., Jiao X., Rannala B., Yang Z. 2019. A Bayesian Implementation of the Multispecies Coalescent Model with Introgression for Phylogenomic Analysis. Mol Biol Evol.

Gang W., Peng G., Ning X., JingZhuo Z., Zhuang S. 2011. Naja kaouthia: a new snake record to Guizhou. Guizhou Agricultural Sciences.:173–176.

Gaston K.J. 2003. The structure and dynamics of geographic ranges. Oxford: Oxford University Press.

GBIFAvailable from https://www.gbif.org/.

Gould S.J., Johnston R.F. 1972. Geographic Variation. Annual review of ecology and systematics. 3:457–498.

Gray L.N., Barley A.J., Poe S., Thomson R.C., Nieto-Montes de Oca A., Wang I.J. 2019. Phylogeography of a widespread lizard complex reflects patterns of both geographic and ecological isolation. Mol Ecol. 28:644–657.

Green R.E., Braun E.L., Armstrong J., Earl D., Nguyen N., Hickey G., Vandewege M.W., St. John J.A., Capella-Gutierrez S., Castoe T.A., Kern C., Fujita M.K., Opazo J.C., Jurka J., Kojima K.K., Caballero J., Hubley R.M., Smit A.F., Platt R.N., Lavoie C.A., Ramakodi M.P., Finger J.W., Suh A., Isberg S.R., Miles L., Chong A.Y., Jaratlerdsiri W., Gongora J., Moran C., Iriarte A., McCormack J., Burgess S.C., Edwards S.V., Lyons E., Williams C., Breen M., Howard J.T., Gresham C.R., Peterson D.G., Schmitz J., Pollock D.D., Haussler D., Triplett E.W., Zhang G., Irie N., Jarvis E.D., Brochu C.A., Schmidt C.J., McCarthy F.M., Faircloth B.C., Hoffmann F.G., Glenn T.C., Gabaldon T., Paten B., Ray D.A. 2014. Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. Science. 346:1254449–1254449.

Gregory P.T. 1978. Feeding habits and diet overlap of three species of garter snakes (Thamnophis) on Vancouver Island. Can. J. Zool. 56:1967–1974.

Gregory P.T. 1984. Habitat, diet, and composition of assemblages of garter snakes (Thamnophis) at eight sites on Vancouver Island. Can. J. Zool. 62:2013–2022.

Guo P., Liu Q., Xu Y., Jiang K., Hou M., Ding L., Alexander Pyron R., Burbrink F.T. 2012. Out of Asia: Natricine snakes support the Cenozoic Beringian Dispersal Hypothesis. Molecular Phylogenetics and Evolution. 63:825–833.

Gutenkunst R.N., Hernandez R.D., Williamson S.H., Bustamante C.D. 2009. Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. PLoS Genet. 5.

Harris R.S. 2007. Improved pairwise Alignmnet of genomic DNA. .

Harvey M.G., Smith B.T., Glenn T.C., Faircloth B.C., Brumfield R.T. 2016. Sequence Capture versus Restriction Site Associated DNA Sequencing for Shallow Systematics. Syst Biol. 65:910–924.

Head J.J., Holroyd P.A., Hutchison J.H., Ciochon R.L. 2005. First report of snakes (Serpentes) from the Late Middle Eocene Pondaung Formation, Myanmar. Journal of Vertebrate Paleontology. 25:246–250.

Heath T.A., Zwickl D.J., Kim J., Hillis D.M. 2008. Taxon Sampling Affects Inferences of Macroevolutionary Processes from Phylogenetic Trees. Syst Biol. 57:160–166.

Hewitt G. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society. 58:247–276.

Hickerson M.J., Carstens B.C., Cavender-Bares J., Crandall K.A., Graham C.H., Johnson J.B., Rissler L., Victoriano P.F., Yoder A.D. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. Molecular Phylogenetics and Evolution. 54:291–301.

Holman J.A., others. 2000. Fossil snakes of North America: origin, evolution, distribution, paleoecology. Indiana University Press.

Huson D.H., Bryant D. 2006. Application of Phylogenetic Networks in Evolutionary Studies. Mol Biol Evol. 23:254–267.

Ineich I. 1995. Etat actuel de nos connaissances sur la classification des serpents venimeux. Bulletin de la Société herpétologique de France.:7–24.

Irwin D.E. 2002. Phylogeographic breaks without geographic barriers to gene flow. Evolution. 56:2383–2394.

Jackson N.D., Carstens B.C., Morales A.E., O'Meara B.C. 2017. Species Delimitation with Gene Flow. Syst Biol. 66:799–812.

Jakobsson M., Rosenberg N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 23:1801–1806.

Janzen F.J., Krenz J.G., Haselkorn T.S., Brodie E.D. 2002. Molecular phylogeography of common garter snakes (Thamnophis sirtalis) in western North America: implications for regional historical forces. Mol Ecol. 11:1739–1751.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 24:1403–1405.

Jouganous J., Long W., Ragsdale A.P., Gravel S. 2017. Inferring the Joint Demographic History of Multiple Populations: Beyond the Diffusion Approximation. Genetics. 206:1549–1567.

Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution. 30:772–780.

King R.B., Lawson R. 2001. Patterns of Population Subdivision and Gene Flow in Three Sympatric Natricine Snakes. Copeia. 2001:602–614.

Kishida T., Go Y., Tatsumoto S., Tatsumi K., Kuraku S., Toda M. 2019. Loss of olfaction in sea snakes provides new perspectives on the aquatic adaptation of amniotes. Proceedings of the Royal Society B. 286:20191828.

Kubatko L.S., Degnan J.H. 2007. Inconsistency of Phylogenetic Estimates from Concatenated Data under Coalescence. Syst Biol. 56:17–24.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2016. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol Biol Evol.:msw260.

Larsen K.W., Gregory P.T., Antoniak R. 1993. Reproductive Ecology of the Common Garter Snake Thamnophis sirtalis at the Northern Limit of Its Range. The American Midland Naturalist. 129:336–345.

Leaché A.D., Chavez A.S., Jones L.N., Grummer J.A., Gottscho A.D., Linkem C.W. 2015. Phylogenomics of Phrynosomatid Lizards: Conflicting Signals from Sequence Capture versus Restriction Site Associated DNA Sequencing. Genome Biology and Evolution. 7:706–719.

Leaché A.D., Zhu T., Rannala B., Yang Z. 2019. The Spectre of Too Many Species. Syst Biol. 68:168–181.

Lee M.S.Y., Sanders K.L., King B., Palci A. 2016. Diversification rates and phenotypic evolution in venomous snakes (Elapidae). Royal Society Open Science. 3:150277.

Lehner B., Verdin K., Jarvis A. 2008. New Global Hydrography Derived From Spaceborne Elevation Data. Eos, Transactions American Geophysical Union. 89:93–94.

Leopold E.B., Dunwiddie P.W., Whitlock C., Nickmann R., Watts W.A. 2016. Postglacial vegetation history of Orcas Island, northwestern Washington. Quaternary Research. 85:380–390.

Lobeck A.K. 1948. Physiographic Provinces of North America; Physiographic Diagram of North America. Geographical Press, Division of CS Hammond & Company.

Longbottom J., Shearer F.M., Devine M., Alcoba G., Chappuis F., Weiss D.J., Ray S.E., Ray N., Warrell D.A., Castañeda R.R. de, Williams D.J., Hay S.I., Pigott D.M. 2018. Vulnerability to snakebite envenoming: a global mapping of hotspots. The Lancet. 392:673–684.

Maddison W.P. 1997. Gene trees in species trees. Systematic biology. 46:523–536.

Manier M.K., Arnold S.J. 2005. Population genetic analysis identifies source–sink dynamics for two sympatric garter snake species (Thamnophis elegans and Thamnophis sirtalis). Molecular Ecology. 14:3965–3976.

Manthey J.D., Campillo L.C., Burns K.J., Moyle R.G. 2016. Comparison of Target-Capture and Restriction-Site Associated DNA Sequencing for Phylogenomics: A Test in Cardinalid Tanagers (Aves, Genus: Piranga). Syst Biol. 65:640–650.

Marth G.T., Czabarka E., Murvai J., Sherry S.T. 2004. The Allele Frequency Spectrum in Genome-Wide Human Variation Data Reveals Signals of Differential Demographic History in Three Large World Populations. Genetics. 166:351–372.

Mayr E. 1963. Animal species and evolution. Animal species and evolution.

McVay J.D., Flores-Villela O., Carstens B. 2015. Diversification of North American natricine snakes. Biol. J. Linn. Soc. 116:1–12.

Minton S.A. 1986. Origins of poisonous snakes: evidence from plasma and venom proteins. Natural toxins—animal, plant and microbial. Clarendon Press, Oxford.:3–21.

Mooi R.D., Wiens J.P., Casper G.S. 2011. Extreme Color Variation within Populations of the Common Gartersnake, Thamnophis sirtalis, in Central North America, with Implications for Subspecies Status. Copeia. 2011:187–200.

Moritz C., Langham G., Kearney M., Krockenberger A., VanDerWal J., Williams S. 2012. Integrating phylogeography and physiology reveals divergence of thermal traits between central and

peripheral lineages of tropical rainforest lizards. Philosophical Transactions of the Royal Society B: Biological Sciences. 367:1680–1687.

Mosher D.C., Hewitt A.T. 2004. Late Quaternary deglaciation and sea-level history of eastern Juan de Fuca Strait, Cascadia. Quaternary International. 121:23–39.

Mussmann S.M., Douglas M.R., Chafin T.K., Douglas M.E. 2019. BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data. Methods in Ecology and Evolution. 10:1808–1813.

Myers E.A., Weaver R.E., Alamillo H. 2013. Population Stability of the Northern Desert Nightsnake (Hypsiglena chlorophaea deserticola) during the Pleistocene. Journal of Herpetology. 47:432–439.

Nabhan A.R., Sarkar I.N. 2012. The impact of taxon sampling on phylogenetic inference: a review of two decades of controversy. Brief Bioinform. 13:122–134.

Nichols R. 2001. Gene trees and species trees are not the same. Trends in Ecology & Evolution. 16:358–364.

Paris J.R., Stevens J.R., Catchen J.M. 2017. Lost in parameter space: a road map for stacks. Methods in Ecology and Evolution. 8:1360–1373.

Pasquesi G.I.M., Adams R.H., Card D.C., Schield D.R., Corbin A.B., Perry B.W., Reyes-Velasco J., Ruggiero R.P., Vandewege M.W., Shortt J.A., Castoe T.A. 2018. Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. Nature Communications. 9:2774.

Perry B.W., Card D.C., McGlothlin J.W., Pasquesi G.I.M., Adams R.H., Schield D.R., Hales N.R., Corbin A.B., Demuth J.P., Hoffmann F.G., Vandewege M.W., Schott R.K., Bhattacharyya N., Chang B.S.W., Casewell N.R., Whiteley G., Reyes-Velasco J., Mackessy S.P., Gamble T., Storey K.B., Biggar K.K., Passow C.N., Kuo C.-H., McGaugh S.E., Bronikowski A.M., de Koning A.P.J., Edwards S.V., Pfrender M.E., Minx P., Brodie E.D., Brodie E.D., Warren W.C., Castoe T.A. 2018. Molecular Adaptations for Sensing and Securing Prey and Insight into Amniote Genome Diversity from the Garter Snake Genome. Genome Biology and Evolution. 10:2110–2129.

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S., Hoekstra H.E. 2012a. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. PLOS ONE. 7:e37135.

Peterson B.K., Weber J.N., Kay E.H., Fisher H.S., Hoekstra H.E. 2012b. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS one. 7.

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

Placyk Jr. J.S., Burghardt G.M., Small R.L., King R.B., Casper G.S., Robinson J.W. 2007. Post-glacial recolonization of the Great Lakes region by the common gartersnake (Thamnophis sirtalis) inferred from mtDNA sequences. Molecular Phylogenetics and Evolution. 43:452–467.

Portik D.M., Leaché A.D., Rivera D., Barej M.F., Burger M., Hirschfeld M., Rödel M.-O., Blackburn D.C., Fujita M.K. 2017. Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. Molecular Ecology. 26:5245–5263.

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics. 155:945–959.

Pyron R.A., Burbrink F.T., Colli G.R., de Oca A.N.M., Vitt L.J., Kuczynski C.A., Wiens J.J. 2011. The phylogeny of advanced snakes (Colubroidea), with discovery of a new subfamily and comparison of support methods for likelihood trees. Molecular Phylogenetics and Evolution. 58:329–342.

Pyron R.A., Burbrink F.T., Wiens J.J. 2013a. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. BMC Evolutionary Biology. 13:93.

Pyron R.A., Kandambi H.K.D., Hendry C.R., Pushpamal V., Burbrink F.T., Somaweera R. 2013b. Genus-level phylogeny of snakes reveals the origins of species richness in Sri Lanka. Molecular Phylogenetics and Evolution. 66:969–978.

de Queiroz A., Lawson R., Lemos-Espinal J.A. 2002. Phylogenetic Relationships of North American Garter Snakes (Thamnophis) Based on Four Mitochondrial Genes: How Much DNA Sequence Is Enough? Molecular Phylogenetics and Evolution. 22:315–329.

Queiroz A., Lawson R. 1994. Phylogenetic relationships of the garter snakes based on DNA sequence and allozyme variation. Biological Journal of the Linnean Society. 53:209–229.

Queiroz A.D., Lawson R. 2008. A peninsula as an island: multiple forms of evidence for overwater colonization of Baja California by the gartersnake Thamnophis validus. BJLS. 95:409–424.

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic biology. 67:901.

Ricketts T.H., Dinerstein E., Olson D.M., Eichbaum W., Loucks C.J., Kavanaugh K., Hedao P., Hurley P., DellaSalla D., Abell R., Carney K., Walters S. 1999. Terrestrial Ecoregions of North America: A Conservation Assessment. Island Press.

Riddle B.R. 2016. Comparative phylogeography clarifies the complexity and problems of continental distribution that drove A. R. Wallace to favor islands. Proc Natl Acad Sci USA. 113:7970–7977.

Rochette N.C., Rivera-Colón A.G., Catchen J.M. 2019. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. Molecular ecology. 28:4737–4754.

Rolland J., Silvestro D., Schluter D., Guisan A., Broennimann O., Salamin N. 2018. The impact of endothermy on the climatic niche evolution and the distribution of vertebrate diversity. Nature Ecology & Evolution. 2:459–464.

Rosenberg N.A. 2004. DISTRUCT: a program for the graphical display of population structure. Molecular ecology notes. 4:137–138.

Rossman D.A. 1996. The Garter Snakes: Evolution and Ecology. University of Oklahoma Press.

Ruane S., Austin C.C. 2017. Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. Molecular Ecology Resources. 17:1003–1008.

Ruthven A.G. 1908. Variations and genetic relationships of the garter-snakes. Govt. print. off.

Sanders K.L., Lee M.S.Y., Leys R., Foster R., Keogh J.S. 2008. Molecular phylogeny and divergence dates for Australasian elapids and sea snakes (hydrophiinae): evidence from seven genes for rapid evolutionary radiations. Journal of Evolutionary Biology. 21:682–695.

Santra V., Wüster W. 2017. Natural History: Naja kaouthia (Monocoled Cobra). Behavior / spitting. Herpetological Review. 48:455–456.

Schoener T.W. 1968. The Anolis Lizards of Bimini: Resource Partitioning in a Complex Fauna. Ecology. 49:704–726.

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

Shepard D.B., Burbrink F.T. 2009. Phylogeographic and demographic effects of Pleistocene climatic fluctuations in a montane salamander, *Plethodon fourchensis*. Mol Ecol. 18:2243–2262.

Slowinski J.B., Keogh J.S. 2000. Phylogenetic Relationships of Elapid Snakes Based on Cytochrome b mtDNA Sequences. Molecular Phylogenetics and Evolution. 15:157–164.

Smith B.T., Harvey M.G., Faircloth B.C., Glenn T.C., Brumfield R.T. 2014. Target Capture and Massively Parallel Sequencing of Ultraconserved Elements for Comparative Studies at Shallow Evolutionary Time Scales. Syst Biol. 63:83–95.

Soltis D.E., Gitzendanner M.A., Strenge D.D., Soltis P.S. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Pl Syst Evol. 206:353–373.

Soltis D.E., Morris A.B., McLACHLAN J.S., Manos P.S., Soltis P.S. 2006a. Comparative phylogeography of unglaciated eastern North America. Mol Ecol. 15:4261–4293.

Soltis D.E., Morris A.B., McLACHLAN J.S., Manos P.S., Soltis P.S. 2006b. Comparative phylogeography of unglaciated eastern North America. Molecular Ecology. 15:4261–4293.

Song S., Liu L., Edwards S.V., Wu S. 2012. Resolving conflict in eutherian mammal phylogeny using phylogenomics and the multispecies coalescent model. Proceedings of the National Academy of Sciences. 109:14942–14947.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.

Stange M., Sánchez-Villagra M.R., Salzburger W., Matschiner M. 2018. Bayesian Divergence-Time Estimation with Genome-Wide Single-Nucleotide Polymorphism Data of Sea Catfishes (Ariidae) Supports Miocene Closure of the Panamanian Isthmus. Syst Biol. 67:681–699.

Streicher J.W., Schulte J.A., Wiens J.J. 2016. How Should Genes and Taxa be Sampled for Phylogenomic Analyses with Missing Data? An Empirical Study in Iguanian Lizards. Syst Biol. 65:128–145.

Suchard M.A., Lemey P., Baele G., Ayres D.L., Drummond A.J., Rambaut A. 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus evolution. 4:vey016.

Sukumaran J., Knowles L.L. 2017. Multispecies coalescent delimits structure, not species. PNAS. 114:1607–1612.

Suryamohan K., Krishnankutty S.P., Guillory J., Jevit M., Schröder M.S., Wu M., Kuriakose B., Mathew O.K., Perumal R.C., Koludarov I., others. 2020. The Indian cobra reference genome and transcriptome enables comprehensive identification of venom toxins. Nature Genetics.:1–12.

Swofford D.L. 2003. Phylogenetic analysis using parsimony (* and other methods). .

Syfert M.M., Smith M.J., Coomes D.A. 2013. The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. PLOS ONE. 8:e55158.

Trape J.-F., Chirio L., Broadley D.G., Wüster W. 2009. Phylogeography and systematic revision of the Egyptian cobra (Serpentes: Elapidae: Naja haje) species complex, with the description of a new species from West Africa. Zootaxa. 2236:1–25.

Uetz P., Cherikh S., Shea G., Ineich I., Campbell P.D., DORONIN I.V., Rosado J., Wynn A., Tighe K.A., McDiarmid R. 2019. A global catalog of primary reptile type specimens. Zootaxa. 4695:438–450.

Ullate-Agote A., Milinkovitch M.C., Tzika A.C. 2015. The genome sequence of the corn snake (Pantherophis guttatus), a valuable resource for EvoDevo studies in squamates. International Journal of Developmental Biology. 58:881–888.

Uyeda J.C. Connecting Microevolutionary Processes to Macroevolutionary Patterns Across Space and Time. :306.

Vonk F.J., Casewell N.R., Henkel C.V., Heimberg A.M., Jansen H.J., McCleary R.J.R., Kerkkamp H.M.E., Vos R.A., Guerreiro I., Calvete J.J., Wüster W., Woods A.E., Logan J.M., Harrison R.A., Castoe T.A., Koning A.P.J. de, Pollock D.D., Yandell M., Calderon D., Renjifo C., Currier R.B., Salgado D., Pla D., Sanz L., Hyder A.S., Ribeiro J.M.C., Arntzen J.W., Thillart G.E.E.J.M. van den, Boetzer M., Pirovano W., Dirks R.P., Spaink H.P., Duboule D., McGlinn E., Kini R.M., Richardson M.K. 2013. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. PNAS. 110:20651–20656.

Voorhies M.R. 1990. Vertebrate biostratigraphy of the Ogallala Group in Nebraska. Geologic Framework and Regional Hydrology: Upper Cenozoic Blackwater Draw and Ogallala Forma- tion, Great Plains.:115–151.

Wallach V., Wüster W., Broadley D.G. 2009. In praise of subgenera: taxonomic status of cobras of the genus Naja Laurenti (Serpentes: Elapidae). Zootaxa. 2236:26–36.

Warren D.L., Glor R.E., Turelli M. 2008. Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. Evolution. 62:2868–2883.

Warren D.L., Glor R.E., Turelli M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. Ecography. 33:607–611.

Warren D.L., Seifert S.N. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecological Applications. 21:335–342.

Weese D.A., Fujita Y., Santos S.R. 2013. Multiple Colonizations Lead to Cryptic Biodiversity in an Island Ecosystem: Comparative Phylogeography of Anchialine Shrimp Species in the Ryukyu Archipelago, Japan. The Biological Bulletin. 225:24–41.

Wiens J.J. 2004. Speciation and Ecology Revisited: Phylogenetic Niche Conservatism and the Origin of Species. Evolution. 58:193–197.

Wiens J.J., Hutter C.R., Mulcahy D.G., Noonan B.P., Townsend T.M., Sites J.W., Reeder T.W. 2012. Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. Biology Letters. 8:1043–1046.

Wilson G.A., Rannala B. 2003. Bayesian Inference of Recent Migration Rates Using Multilocus Genotypes. Genetics. 163:1177–1191.

Wisz M.S., Pottier J., Kissling W.D., Pellissier L., Lenoir J., Damgaard C.F., Dormann C.F., Forchhammer M.C., Grytnes J.-A., Guisan A., Heikkinen R.K., Høye T.T., Kühn I., Luoto M., Maiorano L., Nilsson M.-C., Normand S., Öckinger E., Schmidt N.M., Termansen M., Timmermann A., Wardle D.A., Aastrup P., Svenning J.-C. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biological Reviews. 88:15–30.

Wood D.A., Vandergast A.G., Lemos Espinal J.A., Fisher R.N., Holycross A.T. 2011. Refugial isolation and divergence in the Narrowheaded Gartersnake species complex (Thamnophis rufipunctatus) as revealed by multilocus DNA sequence data: REFUGIAL ISOLATION IN NARROWHEADED GARTERSNAKES. Mol Ecol. 20:3856–3878.

Wüster W., Broadley D.G. 2007. Get an eyeful of this: a new species of giant spitting cobra from eastern and north-eastern Africa (Squamata: Serpentes: Elapidae: Naja). Zootaxa. 1532:51–68.

Wüster W., Chirio L., Trape J.-F., Ineich I., Jackson K., Greenbaum E., Barron C., Kusamba C., Nagy Z.T., Storey R., Hall C., Wüster C.E., Barlow A., Broadley D.G. 2018. Integration of nuclear and mitochondrial gene sequences and morphology reveals unexpected diversity in the forest cobra (Naja melanoleuca) species complex in Central and West Africa (Serpentes: Elapidae). Zootaxa. 4455:68–98.

Wüster W., Crookes S., Ineich I., Mané Y., Pook C.E., Trape J.-F., Broadley D.G. 2007. The phylogeny of cobras inferred from mitochondrial DNA sequences: Evolution of venom spitting and the phylogeography of the African spitting cobras (Serpentes: Elapidae: Naja nigricollis complex). Molecular Phylogenetics and Evolution. 45:437–453.

Wüster W., Thorpe R.S. 1992. Dentitional Phenomena in Cobras Revisited: Spitting and Fang Structure in the Asiatic Species of Naja (Serpentes: Elapidae). Herpetologica. 48:424–434.

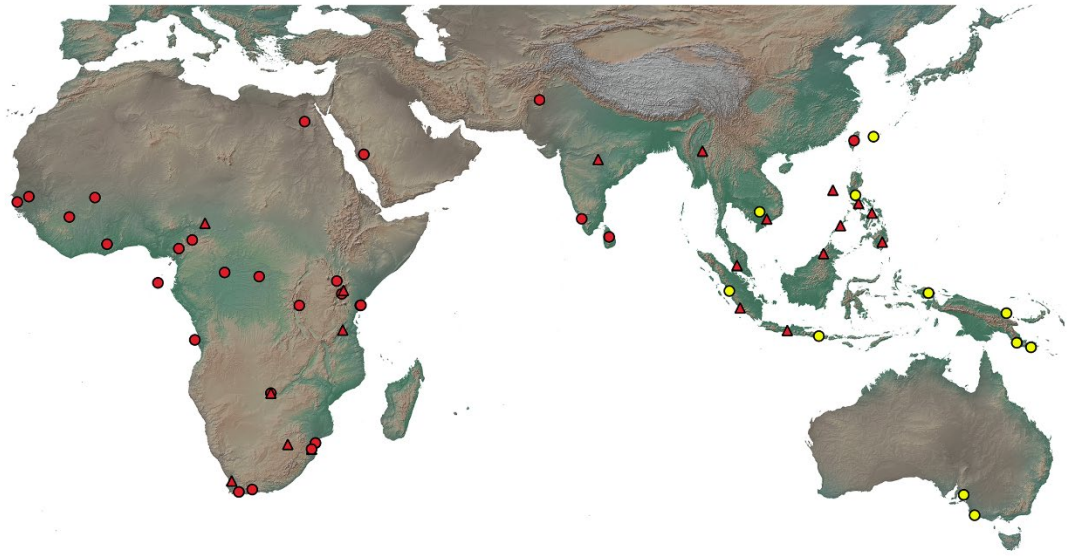
Yap M.K.K., Tan N.H., Fung S.Y. 2011. Biochemical and toxinological characterization of Naja sumatrana (Equatorial spitting cobra) venom. Journal of Venomous Animals and Toxins including Tropical Diseases. 17:451–459.

Young B.A. 2004. The buccal buckle: the functional morphology of venom spitting in cobras. Journal of Experimental Biology. 207:3483–3494.

Zaher H., Murphy R.W., Arredondo J.C., Graboski R., Machado-Filho P.R., Mahlow K., Montingelli G.G., Quadros A.B., Orlov N.L., Wilkinson M., Zhang Y.-P., Grazziotin F.G. 2019. Large-scale molecular phylogeny, morphology, divergence-time estimation, and the fossil record of advanced caenophidian snakes (Squamata: Serpentes). PLOS ONE. 14:e0216148.

Zink R.M. 2002. Methods in Comparative Phylogeography, and Their Application to Studying Evolution in the North American Aridlands. Integr Comp Biol. 42:953–959.

Fig 1. Geographic distribution of samples analyzed in this study. Core cobras are marked red with spitting cobras indicated by triangles. Outgroup elapids are represented in yellow.



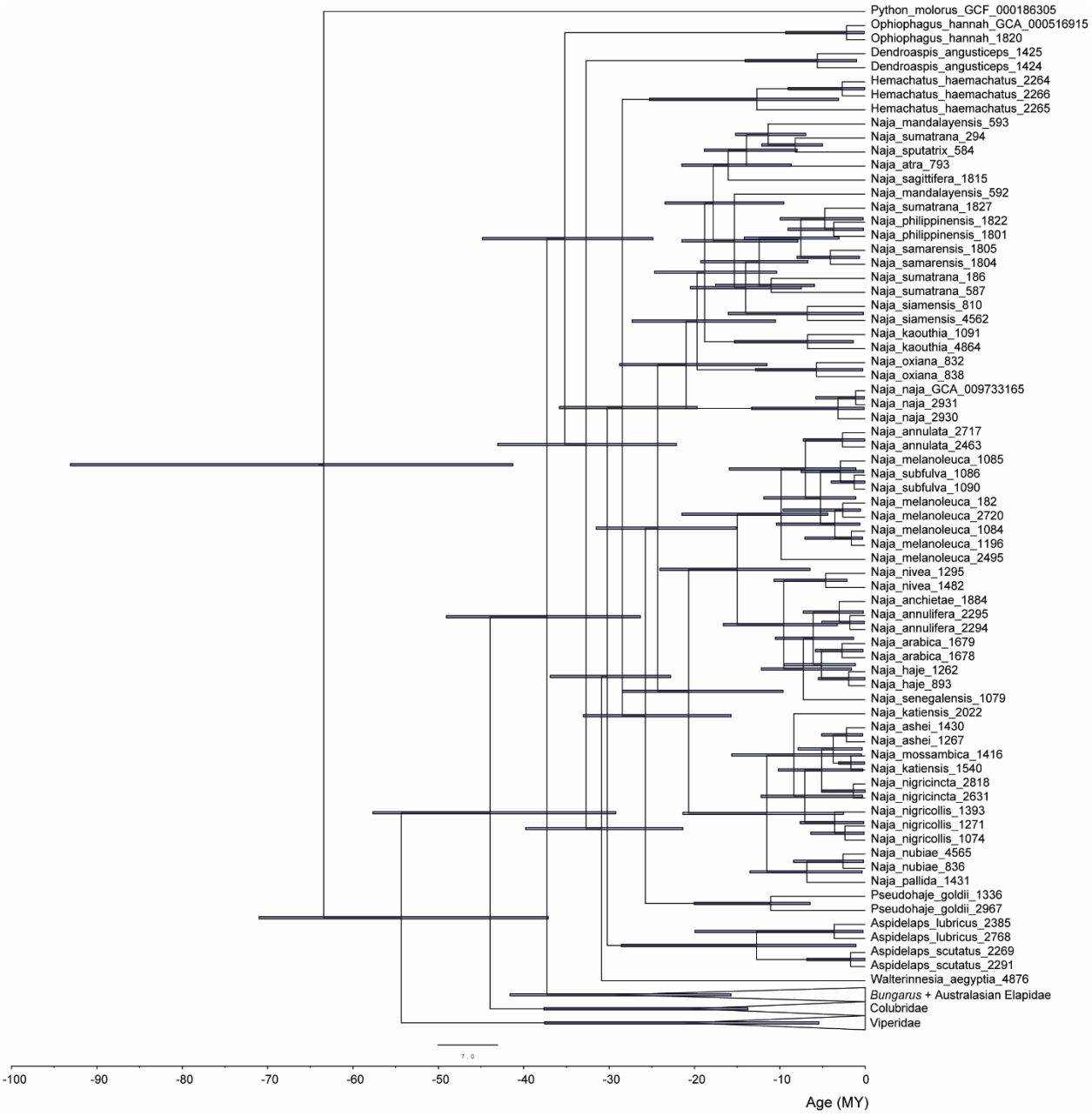


Fig 2. Time-calibrated phylogeny of the cobras inferred from concatenated UCE data using BEAST. Grey bars represent 95% confidence intervals for the associated node age. The fully annotated tree file featuring node posterior and age estimates is available in the supplement.

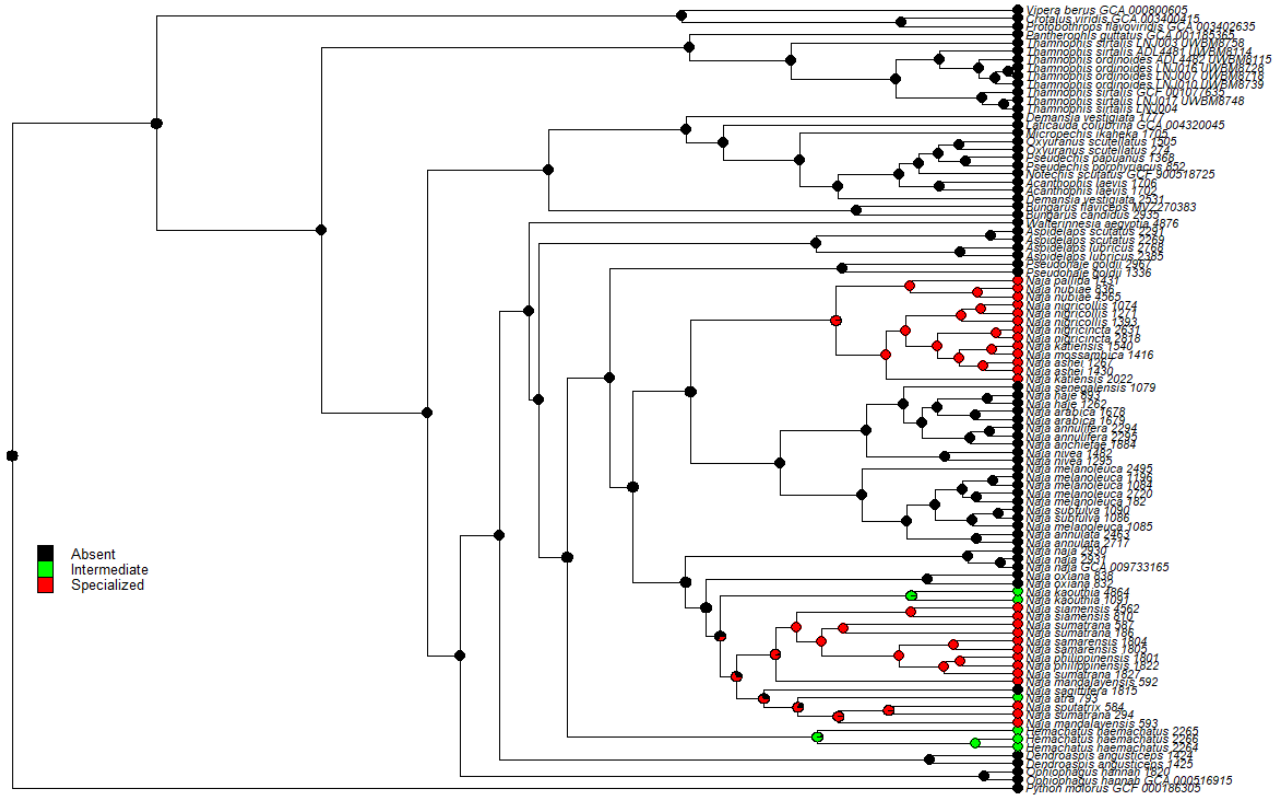


Fig 3. Results of the ancestral state reconstruction of the defensive spitting behavior in the cobras. Stochastic character mapping analyses under the "ER" model support two separate origins of specialized spitting from an absent ancestral state in *Naja*, as well as a separate evolution of intermediate spitting ability observed in *Hemachatus*.

Table 1. Summary of UCE capture and ddRADseq assemblies used in the study. Missing data thresholds of 50% were applied to all datasets to maximize information content and minimize the impact of low coverage samples on phylogenetic inference.

DDRAD data	**# species (individuals)**	**Retained loci (% of prefiltered)**	**mean loci/sample**	**Unlinked SNPs**
Complete	57 (39)	3978 (2.04)	2935.12	3973
UCE-DDRAD shared samples	27 (29)	8746 (5.97)	5372.79	8698
Asiatic *Naja*	13 (11)	14898 (21.51)	10406	12542
Asiatic *Naja* + *Ophiophagus hannah*	14 (12)	15211 (20.39)	10112.57	13144
N. melanoleuca complex + *N. fulva* + *N. annulata*	7 (10)	15295 (25.65)	10848.8	12770
N. melanoleuca complex + *N. fulva* + *N. annulata* + *Ophiophagus hannah*	8 (11)	12205 (18.8)	8867.09	10824
UCE data	**# species (individuals)**	**Retained UCEs (% of prefiltered)**	**mean loci/sample**	**Informative sites**
Complete	51 (96)	552 (95.83)	332.85	10119
UCE-DDRAD shared samples	27 (29)	555 (96.69)	352.6	2954
Asiatic *Naja*	11 (19)	546 (96.30)	207.66	879
Asiatic *Naja* + *Ophiophagus hannah*	12 (21)	552 (97.01)	328.67	1463
N. melanoleuca complex + *N. fulva* + *N. annulata*	7 (10)	557 (98.76)	304.21	319
N. melanoleuca complex + *N. fulva* + *N. annulata* + *Ophiophagus hannah*	8 (12)	557 (98.76)	352.77	1389

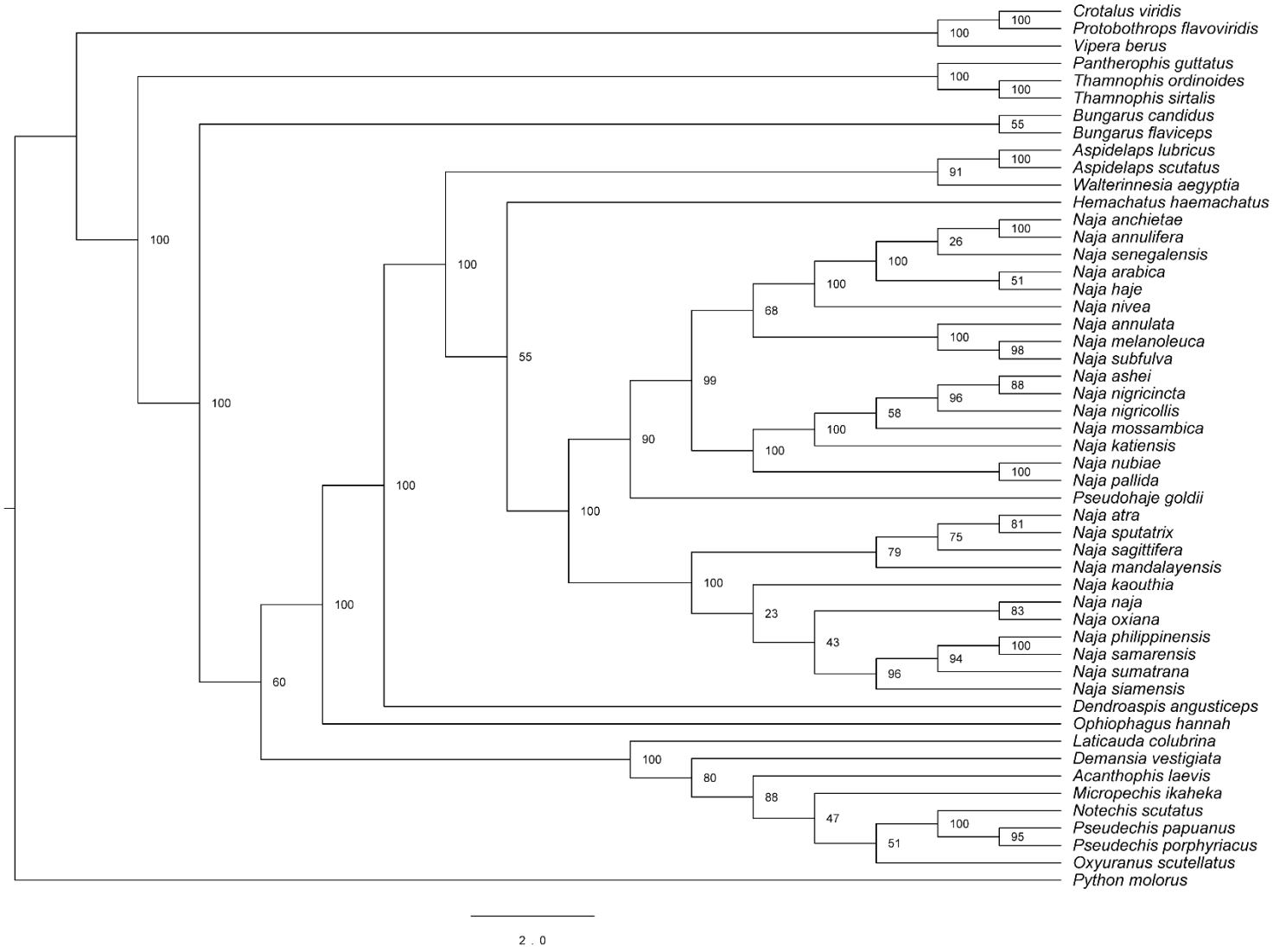


Fig 4. Unrooted phylogeny of the cobras inferred from UCE data under the coalescent model using SVDQuartets. Node values reflect bootstrap support values after 500 replicates.

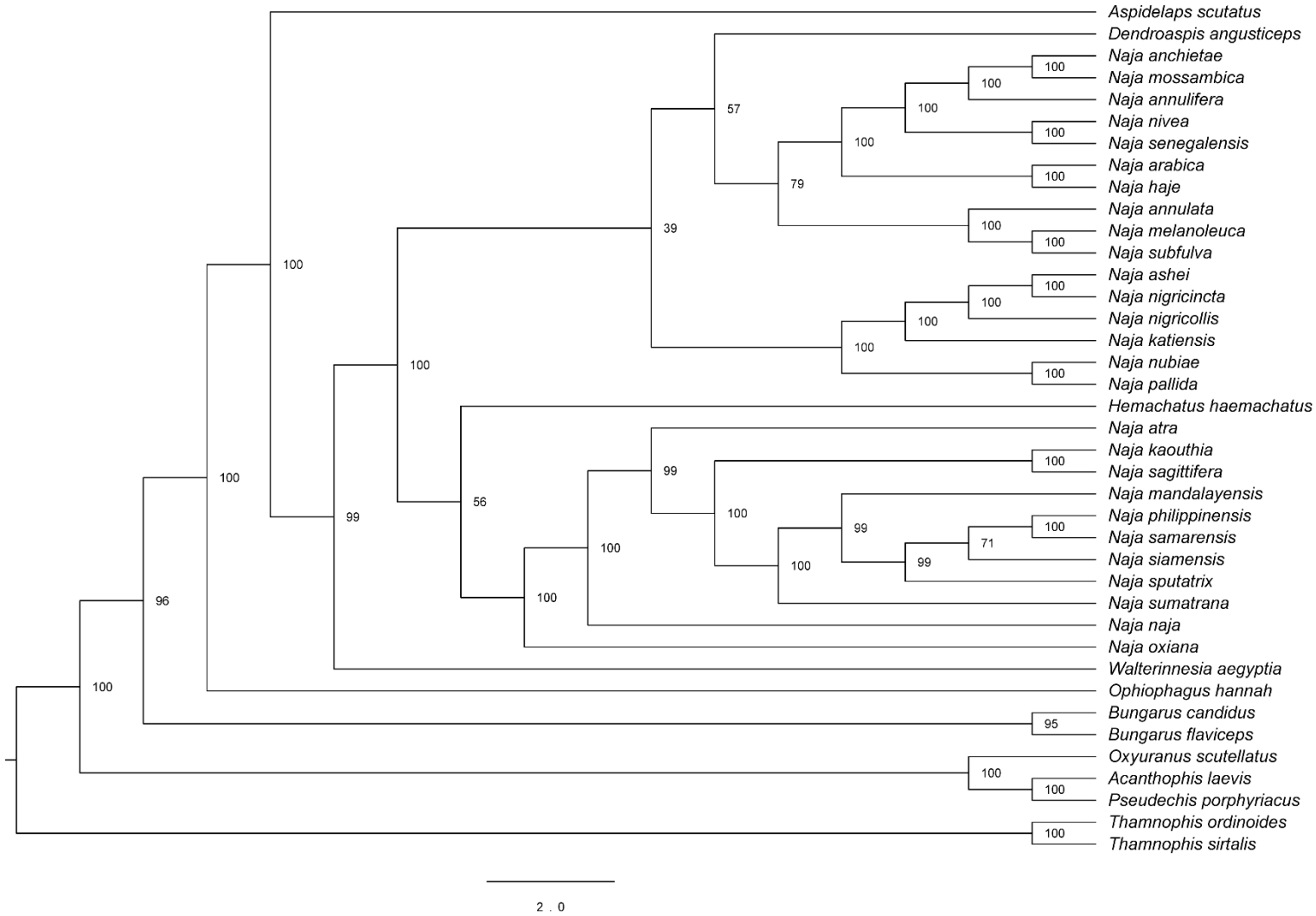


Fig 5. Unrooted phylogeny of the cobras inferred from SNP data using SVDQuartets. Node values reflect bootstrap support values after 500 replicates.

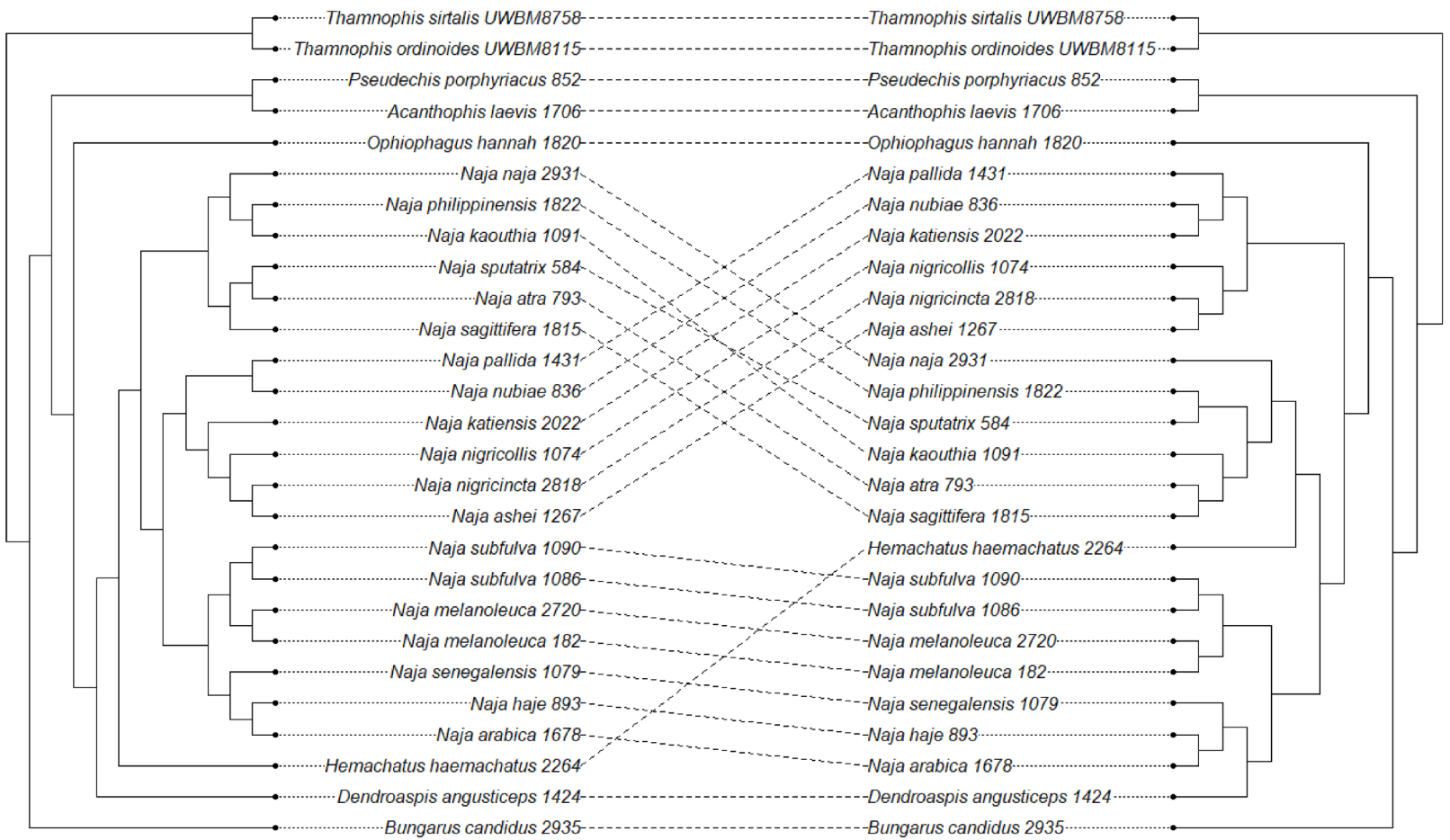


Fig 6. SVDQuartets phylogenies estimated from UCE (Left) and SNP (Right) data reduced to individuals shared across datasets. The UCE tree broadly recovers a tree topology like previous estimates, while the SNP tree topology is characterized by weakly supported and erroneous placements for *Hemachatus*, *Dendroaspis*. Contrary to the topology of the tree estimated from the complete SNP dataset, the topology of the sample-reduced SNP tree renders the African *Naja* clade non-monophyletic.

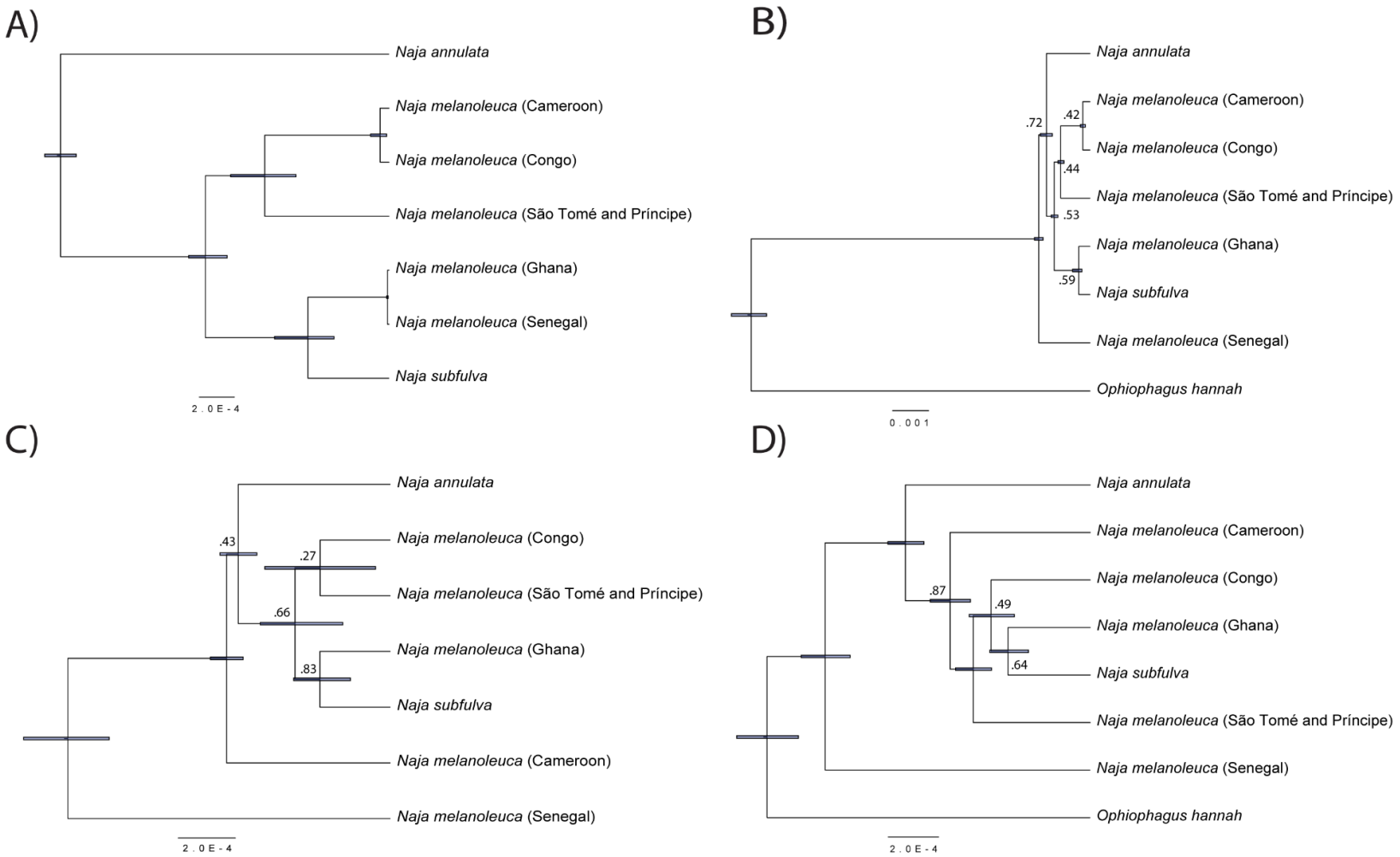


Fig 7. Phylogenetic relationships of the *Naja melanoleuca* complex inferred from SNP (A-B) and UCE (C-D) data using BPP. Node values represent posterior probabilities < .90 of the best tree produced from the A01 analysis, and node bars reflect 97.5% HPD calculated in the A00 analysis. Relationships inferred from UCE data are not highly supported. In both UCE and SNP data, phylogenetic relationships within *Naja* and their support values change with the addition of *Ophiophagus hannah* as an outgroup.

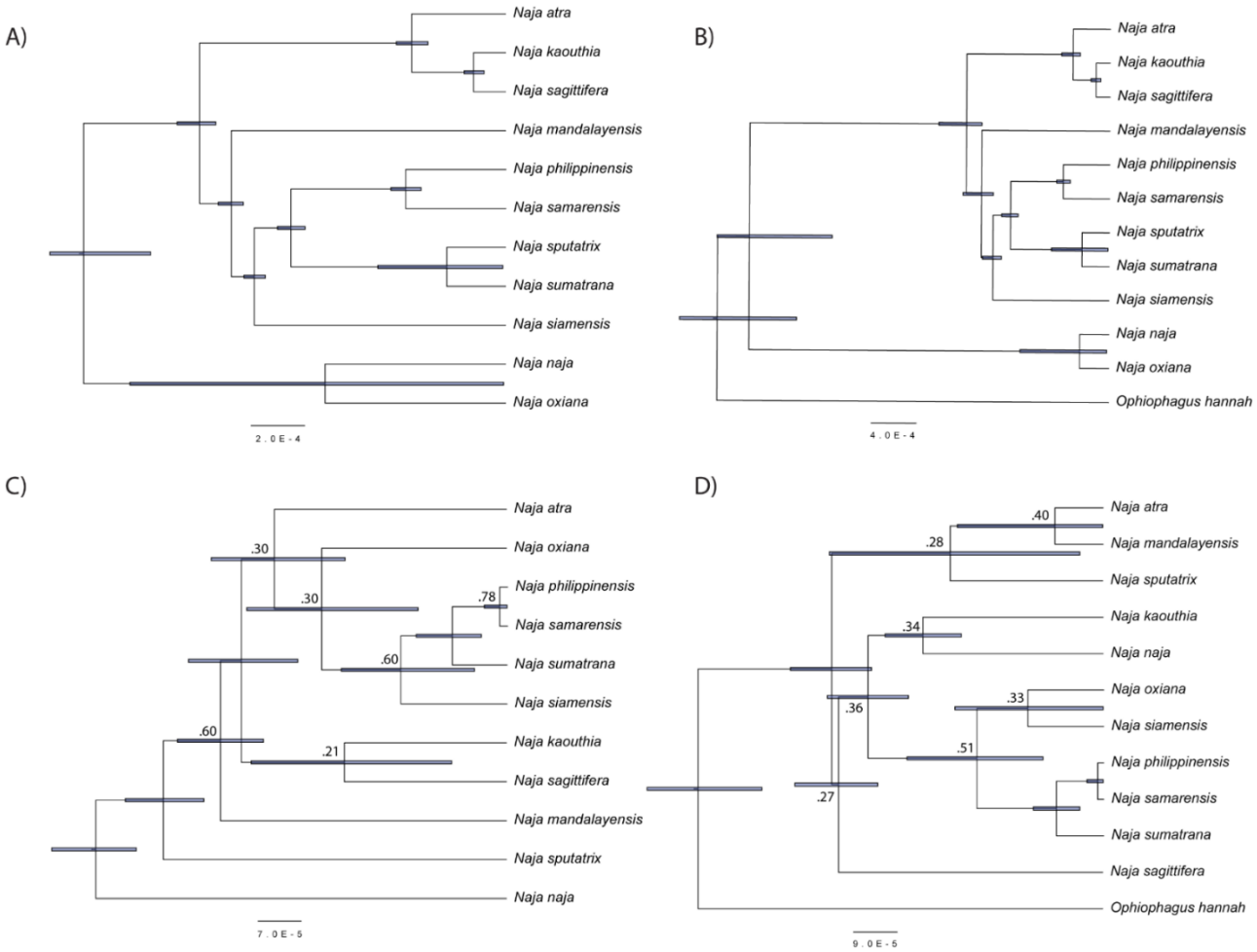


Fig 8. Phylogenetic relationships of the Asiatic *Naja* inferred from SNP (A-B) and UCE (C-D) data using BPP. Node values represent posterior probabilities < .90 of the best tree produced from the A01 analysis, and node bars reflect 97.5% HPD calculated in the A00 analysis. Relationships inferred from UCE data are not highly supported. In both UCE and SNP data, phylogenetic relationships within *Naja* and their support values change with the addition of *Ophiophagus hannah* as an outgroup.

Table S1. Sample table for the study.

Sample/Accession #	Species	Locality	Country	lat	long
1702	*Acanthophis laevis*	Karkar Island, Madang Province	Papua New Guinea	-4.6499974	145.9666628
1706	*Acanthophis laevis*	Karkar Island, Madang Province	Papua New Guinea	-4.6499974	145.9666628
2385	*Aspidelaps lubricus*	Central Oudtshoorn	South Africa	-33.60047	22.19955
2768	*Aspidelaps lubricus*	Western Cape	South Africa	-34	20
2269	*Aspidelaps scutatus fulafula*	Maputo	Mozambique	-25.96553	32.58322
2291	*Aspidelaps scutatus fulafula*	Maputo	Mozambique	-25.96553	32.58322
2935	*Bungarus candidus*	Chup Kampong Cam	Cambodia	11.99339	105.4635
MVZ270383	*Bungarus flaviceps*	West Sumatra	Indonesia	-1	100.5
1777	*Demansia vestigiata*	Padi Padi palm oil plantation, Milne Bay Province	PNG	-10.25	150
2531	*Demansia vestigiata*	Vanumai, Central Province	PNG	-9.5	147.666667
1424	*Dendroaspis angusticeps*	Watamu	Kenya	-3.342547	40.027416
1425	*Dendroaspis angusticeps*	Mida Creek, Watamu	Kenya	-3.3654365	39.9283292
2264	*Hemachatus haemachatus*	Nisela	Swaziland	-26.9834435	31.9373203
2266	*Hemachatus haemachatus*	Johannesburg	South Africa	-26.195246	28.034088
GCA4320045	*Laticauda colubrina*	Ishigaki Island	Japan	24.340556	124.155556
1705	*Micropechis ikaheka*	Karkar Island, Madang Province	Papua New Guinea	-4.6499974	145.9666628
1884	*Naja anchietae*	Kazungula	Botswana	-17.7833	25.2667
2463	*Naja annulata*				
2717	*Naja annulata*	Impongui, Likouala	Congo	2.043924	17.668887
2294	*Naja annulifera*	Nisela	Swaziland	-26.9834435	31.9373203
2295	*Naja annulifera*	Nisela	Swaziland	-26.9834435	31.9373203
1678	*Naja arabica*	Taif	Saudi Arabia	21.437273	40.512714
1679	*Naja arabica*	Taif	Saudi Arabia	21.437273	40.512714
1267	*Naja ashei*	Watamu	Kenya	-3.342547	40.027416

1430	*Naja ashei*	Watamu	Kenya	-3.342547	40.027416
793	*Naja atra*		Taiwan	23.6978092	120.9605179
893	*Naja haje*		Egypt	26.8205528	30.8024979
1262	*Naja haje*	Athi River	Kenya	-1.4717563	36.9035304
1091	*Naja kaouthia*		India	20.5936832	78.962883
4864	*Naja kaouthia*	Ampang, Selangor	Malaysia	3.1588	101.757446
1540	*Naja katiensis*	Doussoudiana	Mali	11.149332	-7.8123716
2022	*Naja katiensis*		Senegal	14.4974012	-14.4523621
592	*Naja mandalayensis*	Near Mandalay	Burma	21.97473	96.08359
593	*Naja mandalayensis*	Near Mandalay	Burma	21.97473	96.08359
182	*Naja melanoleuca*		Cameroon	7.3697219	12.354722
1084	*Naja melanoleuca*	N'kawkaw, Kumasi	Ghana	6.68848	-1.62443
1085	*Naja melanoleuca*	N'kawkaw, Kumasi	Ghana	6.68848	-1.62443
1196	*Naja melanoleuca*		São Tomé	0.33654	6.72732
2495	*Naja melanoleuca*	Keur sen gueye	Senegal	13.6064763	-16.3513438
2720	*Naja melanoleuca*	Impongui, Likouala	Congo	2.043924	17.668887
1416	*Naja mossambica*	Mikumi National Park	Tanzania	-7.413322	37.055714
1892	*Naja mossambica*	Kazungula	Botswana	-17.7833	25.2667
2930	*Naja naja*		Sri Lanka	7.873054	80.7717972
2931	*Naja naja*	Karachi	Pakistan	30.3753204	69.3451157
GCA9733165	*Naja naja*	Kerala	India	10.850516	76.27108
2631	*Naja nigricincta nigricincta*				
2818	*Naja nigricincta woodi*	Clanwilliam	South Africa	-32.1960966	18.8339153
1074	*Naja nigricollis ssp.*	Lara, Kaélé	Cameroon	10.10917	14.45083
1271	*Naja nigricollis*	Makuyu	Kenya	-0.9	37.183333
1393	*Naja nigricollis crawshayi*	Mfuwe	Zambia	19.469	-154.833
1295	*Naja nivea*	Southwest Cape	South Africa	-34	20
1482	*Naja nivea*	25 km NE Oudsthoorn, Western Cape	South Africa	-34	20

836	*Naja nubiae*				
4565	*Naja nubiae*	unknown	unknown		
832	*Naja oxiana*				
838	*Naja oxiana*				
1081	*Naja pallida*		Kenya		
1431	*Naja pallida*	Lake Baringo	Kenya	0.633333	36.083333
1801	*Naja philippinensis*	Bicol, Luzon	Philippines	15.5465287	117.4868428
1822	*Naja philippinensis*	Marinduque Island	Philippines	13.384635	121.7006577
1815	*Naja sagittifera*				
1804	*Naja samarensis*	Samal Island, Davao, Mindanao	Philippines	7.0452255	125.5902958
1805	*Naja samarensis*	Samar	Philippines	11.79859	123.8923119
1079	*Naja senegalensis*	Bandriaguara	Mali	14.35	-3.611111
810	*Naja siamensis*	Southern	Vietnam	10.82302	106.62965
4562	*Naja siamensis*	unknown	unknown		
584	*Naja sputatrix*	Java	Indonesia	-7.491667	110.004444
1086	*Naja subfulva*	Bamenda	Cameroon	5.9597	10.14597
1090	*Naja subfulva*		Burundi	-3.3730559	29.9188862
186	*Naja sumatrana*	Tenom, Sabah	Malaysia	5.133333	115.95
294	*Naja sumatrana*	Benkulu	Sumatra	-3.788892	102.266579
587	*Naja sumatrana*	Zoo Negara, Ampang, Selangor	Malaysia	3.2089018	101.7542854
1827	*Naja sumatrana*	Puerto Princesa, Palawan	Philippines	9.740696	118.730072
GCF900518725	*Notechis scutatus*				
1820	*Ophiophagus hannah*	Wawa Dam	Philippines	14.727778	121.191667
GCA516915	*Ophiophagus hannah*				
274	*Oxyuranus scutellatus canni*	Merauke, Irian Jaya	Indonesia	-1.336115	133.174716
1505	*Oxyuranus scutellatus canni*	Moreguina, Central Province	Papua New Guinea	-9.5	147.666667
GCA1185365	*Pantherophis guttatus*				
1368	*Pseudechis papuanus*	Irian Jaya	Indonesia	-1.336115	133.174716
852	*Pseudechis porphyriacus*	Barossa Valley	Australia	-34.4768	138.9606

1336	*Pseudohaje goldii*		Uganda		
2967	*Pseudohaje goldii*	Pango Aluquem, Bengo Province	Angola		
GCA900518735	*Pseudonaja textilis*	Barossa Valley	Australia	-34.4768	138.9606
GCF186305	*Python molorus bivittatus*	unknown (obtained from the pet trade)			
UWBM8115	*Thamnophis ordinoides*	San Juan County, WA	United States		
UWBM8718	*Thamnophis ordinoides*	San Juan County, WA	United States		
UWBM8739	*Thamnophis ordinoides*	San Juan County, WA	United States		
UWBM8728	*Thamnophis ordinoides*	San Juan County, WA	United States		
UWBM8758	*Thamnophis sirtalis*	San Juan County, WA	United States		
GCF1077635	*Thamnophis sirtalis*	Benton County, OR	United States		
GCA800605	*Vipera berus*		Sweden		
4876	*Walterinnesia aegyptia*				