THE VENOMOUS RATTLESNAKES AND NEW WORLD CORALSNAKES: MOLECULAR SYSTEMATICS, BIOGEOGRAPHY AND THE EVOLUTION OF VENOM SYSTEMS

by

JACOBO REYES-VELASCO

Presented to the Faculty of the Graduate School of Biology

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2015

Acknowledgements

During my time at UTA I have received tremendous help and support from a number of people. Dr. Jonathan A. Campbell and Dr. Todd A. Castoe have been incredible advisors and I couldn't have asked for better better mentors. My commetee members have always been very helpful: Dr. Matthew K. Fujita, Dr. Dr. Jeffrey E. Demuth and Dr. Eric N. Smith. Previous committee members include Dr. Paul Chippindale, Dr. Laura Gough and Dr. James Robinson. I am specially thankfull to Dr. Smith for his help and support even before I was a graduate student at UTA. The rest of the faculty and staff at the Life Science department at UTA have also been a bless – Dr. Matthew Walsh and Dr. Walter Schargel have been very helpful in answering several statistic related questions. Without the ladies in Life Science office I would probably been kicked out of UTA long time ago. Paulette Batten, Linda Taylor, Gloria Burlinham, Sherri Echols, Zaida Alsina and Derik Austin. I am in great debt to all of them. I am very greatefull to all other staff members in the Biology Department, specially Rachel Wostl and Jill Castoe.

Without the help of multiple researchers across many countries it would have been impossible for me to acomplish my research goals. I am especially grateful to many researchers and collection managers from several US and Mexican collections, including Carl Franklin (UTA), Edmundo Pérez Rámoz and Oscar Flores-Villela (MZFC), as well as the following institutions: IPN, IBUNAM, LSU, KU, UTEP and UT Tyler. Many other researchers kindly provided important tissues, most prominently Carlos Hernandez, Ivan Ahumada-Carrillo, Uri García, Luis Canseco. Peter T. Nguyen, Augusto Barragan, George Tsai and Reem Abunhandara provided great help in the lab and in the museum. My friends Alexander Hermosillo, Oscar Avila, Gabriela Zamora-Silva, Chistoph Grünwald, Jason Jones and Ginny Weatherman have been incredible over the years and play a fundamental piece in the development of my academic career. My parents

Aranzazu Velasco-Lafarga and Oscar Reyes-Ruvalcaba, my sister Urzula and my late grandmother Guadalupe Lafarga (Abu) have always been incredible in helping me acomplish all of my goals. This dissertation is dedicated to them, as well as to the rest of my family.

Last but not least are all the gradstudents and other friends from UT Arlington. I have learn so much from them during the last 5 years, I don't know were to beggin to thank them all. They include Alex Hall, Assiatu Bah, Audra Andrew, Christian Cox, Coleman Sheehy, Claudia Marquez, Contessa Ricci, Corey Lawns, Corey Roelke, Daren Card, Diego Martin-Perez, Drew Schield, Elija Wostl, Emmanuela Mujica, Eric Watson, Heath Blackmon, James Titus-McQuillan, Jeffrey Streicher, Jesse Meik, Kyle O'Connell, Kyle Shaney, Liomari Diaz, Manish Yelekar, Paul Pasichnyk, Rachel Wostl, Rich Adams, Shannon Beston, Utpal Smart and Walter Schargle. I especially need to thank David Sánchez and Ruben Tovar, *mis parceros*, for so much during these last 5 years.

May 8, 2015

Abstract

THE VENOMOUS RATTLESNAKES AND NEW WORLD CORALSNAKES:

MOLECULAR SYSTEMATICS, BIOGEOGRAPHY

AND THE EVOLUTION OF

VENOM SYSTEMS

Jacobo Reyes-Velasco, PhD

The University of Texas at Arlington, 2015

Supervising Professors: Todd A. Castoe and Jonathan A. Campbell

Middle America is one of the most biologically diverse regions in the world, however it also possessessome of the highest rates of biodiversity loss. Reptiles are especially diverse in this region, and are a very important component of the biota, from a biological as well as a cultural perspective. Venomous snakes are the most studied group of reptiles in Mexico and Central America, as they are of great ecological and medical importance. Nevertheless, our understanding of the evolutionary relationships of many groups of venomous snakes is superficial. Understanding these relationships is no trivial matter as it is fundamental to appropriately address broader evolutionary questions. Herein I examine the systematic relationships of the two most diverse lineages of venomous snakes in Middle America, the rattlesnakes (genus *Crotalus*) and the coralsnakes (genus *Micrurus*). In particular, I studied the phylogenetic relationships of an obscure lineage of rattlesnakes, the longtailed rattlesnakes, as well as the phylogeny and species limits of coralsnakes in the *diastema* species complex of the genus *Micrurus.* I used nuclear and mitochondrial genes to test if the longtailed rattlesnakes form a monophyletic group and to estimate their closest relatives within the genus. By doing so, I

provide the most robust molecular phylogeny for the rattlesnakes to date. Relationships among coralsnakes are known to be difficult to estimate with the use of DNA markers obtained from traditional sequencing techniques. Therefore, I use a combination of traditional sequencing of mitochondrial DNA and next generation sequencing (in the form of double digest restriction associated DNA sequencing, ddradseq), to elucidate evolutionary relationships and species limits within the *diastema* species complex of coralsnakes. Lastly, I made use of the Burmese python genome as a proxy to understand how non-venom genes became recruited into venom systems in the most recent ancestor of both coralsnakes and rattlesnakes.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Phylogenetic relationships of the enigmatic longtailed rattlesnakes (*Crotalus ericsmithi, C. lannomi* and *C. stejnegeri*) and insigths into the phylogeny of rattlesnakes

Introduction

Rattlesnakes are a unique and distinctive group of venomous snakes exclusive to the Western Hemisphere that have intrigued biologists and laymen alike for centuries. Their distinctive morphological features, potent venom, and wide geographic range have contributed to both their medical and cultural importance (Greene 2000). Rattlesnakes range from Canada to Argentina, and include 36 species placed within two genera, *Crotalus* and *Sistrurus* (The Reptile Database, accessed March 2013; (Uetz and Etzold 1996). Since Linnaeus described the first rattlesnake species in 1758 they have become among the most studied group of reptiles (e.g. >2,000 citations on PubMed). Multiple hypotheses concerning the systematic relationships of the rattlesnakes have been proposed based on internal and external morphology (Gloyd 1940; Klauber 1956; Brattstrom 1964; Klauber 1972), venom properties (Githens and George 1931; Minton 1956; Foote and MacMahon 1977), immunological and electrophoretic data (Cadle 1992; Minton 1992) and molecular data (Parkinson 1999; Murphy et al. 2002; Parkinson 2002; Castoe and Parkinson 2006). Despite substantial attention, a cohesive and well-supported phylogenetic hypothesis for the relationships among rattlesnake species remains absent, particularly at deeper nodes in the rattlesnake tree. Among published phylogenies there is much conflict between morphological and molecular-based analyses, and even among molecular-based estimates (Murphy et al. 2002; Castoe and Parkinson 2006).

It is notable that the majority of molecular phylogenies that include rattlesnakes (e.g. (Castoe and Parkinson 2006; Lawing and Polly 2011; Pyron et al. 2011) have recycled the same GenBank sequences of earlier studies, many published more than a

decade ago (Parkinson 1999; Murphy et al. 2002). Thus, despite many studies including rattlesnake DNA sequences, there have been little new data added to refine estimates of rattlesnake relationships. In addition to the issue of minimal additions to gene sequences being used to resolve rattlesnake phylogeny, there are several rare rattlesnake species that have never been included in any molecular study, and this systematic exclusion of lineages may lead to decreased accuracy of inferred phylogenies (Rannala et al. 1998; Zwickl and Hillis 2002).

One group of species collectively referred to as the "longtailed" rattlesnakes has never been included in a molecular phylogenetic analysis and contains the rarest rattlesnake species in museums worldwide. The group is composed of three species that inhabit the coastal foothills of western Mexico (Campbell and Lamar 2004a; Campbell and Flores-Villela 2008; Reyes-Velasco et al. 2010), from Sinaloa to Guerrero (see Fig. 1). Although at least one species of the longtailed rattlesnakes has been known to science for more than 115 years (Boulenger 1896), they have remained particularly rare in biological collections. The first species to be described was *C. stejnegeri* Dunn (1919); this species inhabits the lower foothills of the Sierra Madre Occidental, in the Mexican states of Durango and Sinaloa (and possibly Nayarit, Sonora and Chihuahua). It is known from fewer than 15 specimens, and had not been collected since 1976 (Armstrong and Murphy 1979). The second species, *C. lannomi,* was described from a single specimen collected in the state of Jalisco in 1966 (Tanner 1966). For many years no additional specimens were reported until the species was recently rediscovered in the mountains of Colima (Reyes-Velasco et al. 2010). The third species, *C. ericsmithi*, was recently described from a single specimen collected in the Sierra Madre del Sur of Guerrero (Campbell and Flores-Villela 2008). Newly acquired material for all three species has also shown that some characters used to distinguish these species from one another are not consistent; for example, head scalation and coloration characters show overlap among species (Reyes-Velasco et al. 2010).

The phylogenetic relationships among the longtailed rattlesnakes, and their position in the rattlesnake phylogeny, have been historically difficult to establish based solely on morphological analysis of the small number of available specimens (Gloyd 1940). Several authors have proposed close affinities between *C. stejnegeri* and the Mexican lance-headed rattlesnake, *C. polystictus,* as well as with members of the *C. triseriatus* group (Dunn 1919; Amaral 1929; Brattstrom 1964; Klauber 1972). In the description of *C. lannomi*, Tanner (1966) suggested a close relationship between *C. lannomi* and *C. stejnegeri*. Later, other authors suggested that these two species were among the most basally-diverged rattlesnake lineage, but were not each other's closest relatives (Klauber 1972; Stille 1987). Most recently, Campbell and Flores-Villela (2008) proposed that *C. stejnegeri, C. lannomi* and the newly described *C. ericsmithi* were closely related, although no explicit inferences were made regarding their relationships to other rattlesnake species. Recent fieldwork in Mexico has substantially increased the number of specimens of longtailed rattlesnakes, thereby facilitating the inclusion of these enigmatic species in molecular phylogenetic analyses and providing the first opportunity to examine previous hypotheses about relationships of longtailed rattlesnake species.

In this study we bring new mitochondrial and nuclear gene sequence data from all three longtailed rattlesnake species to bear on questions relating to the relationships among these species and their placement in the phylogeny of rattlesnakes. We also add new data to supplement existing GenBank sequences for several other rattlesnake species, to fill in sampling for major lineages and to replace GenBank data we identify as questionable. With this data set we evaluate the following questions: 1) Are the longtailed rattlesnake species valid and moderately divergent from one another? 2) Do the three longtailed rattlesnake species form a monophyletic group, and if so, how are they related to one another? 3) Where do longtailed rattlesnakes fall within the phylogeny of all rattlesnakes and what lineages are most closely related to them? 4) When did longtailed

rattlesnakes diverge from one another and from other rattlesnake lineages? 5) Can estimated divergence times be plausibly linked to spatio-temporal biogeographic events?

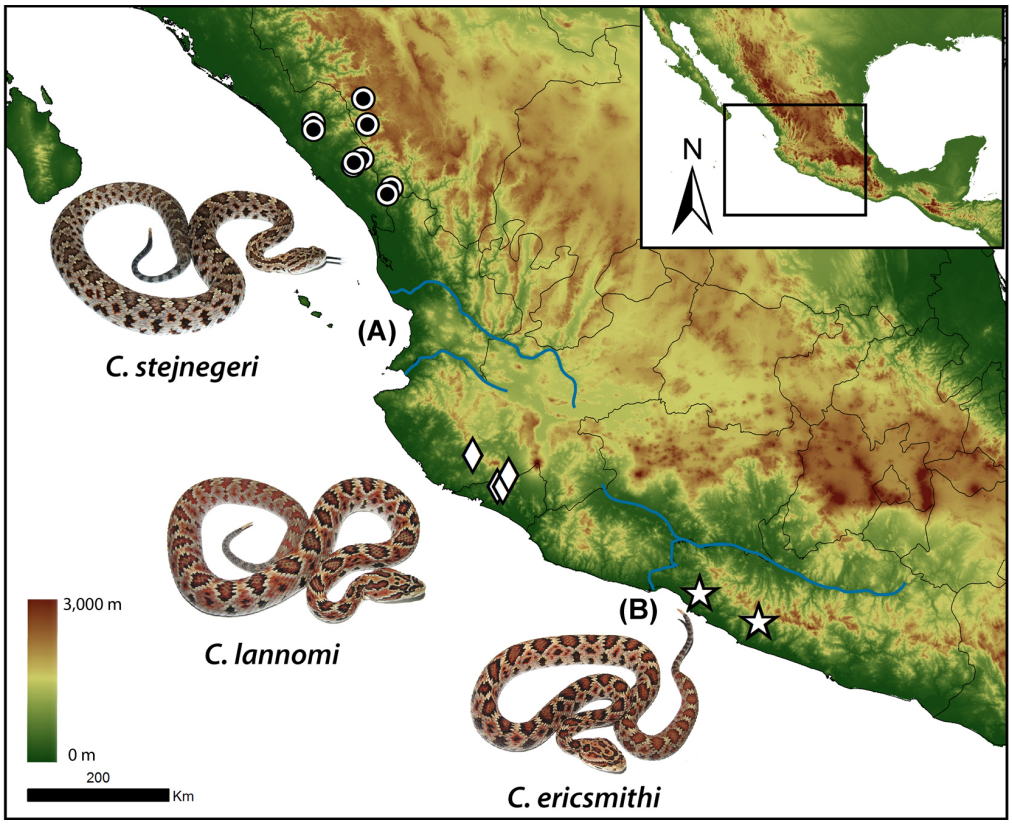


Figure 1 - Map of central Mexico showing topographic relief and indicating the known ranges of each of the three species of longtailed rattlesnakes, as well as possible biogeographic barriers. Rivers are indicated in blue, as either (A) Rio Grande de Santiago-Rio Ameca, or (B) Rio Balsas. Icons represent the only known localities of the longtailed rattlesnakes: circles – *Crotalus stejnegeri*; diamonds – *C. lannomi*; stars – *C. ericsmithi*.

Materials and methods

Taxon sampling

We collected all three species of longtailed rattlesnakes, including two specimens of *C. ericsmithi,* three specimens of *C. lannomi* and three specimens of *C. stejnegeri,* between 2007 and 2011. These specimens represent the only individuals of two of the species (*C. ericsmithi* and *C. lannomi*) and three of only four specimens of *C. stejnegeri* known to have been collected in over 30 years (Campbell and Flores-Villela 2008; Villa and Uriarte-Garzon 2011). Tissue samples (muscle or liver) were preserved in either 95% ethanol or tissue lysis buffer (Burbrink and Castoe 2009). Whole preserved specimens were fixed in formalin and deposited at the Museo de Zoologia, Faculta de Ciencias, Universidad Nacional Autónoma de México (MZFC-UNAM) and the University of Texas at Arlington Amphibian and Reptile Diversity Research Center (UTA-ARDRC). We obtained tissues of additional species of rattlesnakes from the frozen tissue collection at the UTA-ARDRC. In addition to new data generated, we retrieved DNA sequences from GenBank from other *Crotalus* species and outgroup taxa. Except in the case of the longtailed rattlesnakes, all sequences (from multiple voucher individuals in some cases) from a particular taxon were combined to represent that taxon in phylogenetic analyses. Data for all specimens and sequences used in this study are provided in the Supplementary Table 1.

Laboratory techniques

Genomic DNA was extracted using the Qiagen DNeasy kit (Valencia, CA, USA). We PCR amplified and sequenced three mitochondrial DNA fragments, including ATPase subunits 6 and 8 (ATP6_8), cytochrome B (cyt-b), and NADH dehydrogenase subunit 4 (ND4). We also amplified and sequenced three nuclear gene fragments: oocyte maturation factor mos (C-mos), neurotrophin-3 (NT3) and recombination activating gene-1 (RAG-1). Gene fragments were amplified using previously published primer sets and

PCR protocols (Supplementary Table 2). Bi-directional sequencing of DNA fragments was performed by the University of Texas Arlington Genomics Core Facility on an ABI 3130 capillary sequencer (Applied Biosystems). Raw sequence chromatographs were trimmed for quality, assembled, and consensus sequences for gene fragments were estimated using Sequencer 4.8 (Gen Codes Corp., Ann Arbor, MI, USA).

Screening problematic GenBank sequences

In preliminary analyses of sequences, we discovered multiple instances in which GenBank sequences appeared to have either been labeled incorrectly upon original deposition, or to represent anomalous or chimeric sequences. In the supplementary material we summarize the evidence for these assumptions (Supplementary Table 3). Many discrepancies were diagnosed by a first-pass phylogenetic screening of all *Crotalus* GenBank sequences using neighbor joining; problematic sequences were identified when the same species or lineage clustered with taxa known to be distantly related (rather than grouping with conspecific or congeneric species) or where species known to be distantly related had identical sequences for rapidly-evolving mitochondrial loci. Other problematic sequences were identified by blastn searches against the NCBI nr database in which only portions of their length aligned to other rattlesnakes, or where they aligned to non-rattlesnake species (details in Supplementary Table 3). Many discrepancies involved apparent mismatching of information between that listed in GenBank and that provided in the referenced publications (i.e., in many cases the original publication results seemed correct but the GenBank details were incorrect), but in other instances, mislabeling of sequences or presumed contamination appear have been responsible for the errors. Several of these problematic sequences could be the cause of erroneous phylogenies in previous works, for example the nesting of *Crotalus enyo* in the *C. durissus* group (Murphy et al. 2002; Castoe and Parkinson 2006), or the apparent paraphyly of *Crotalus* found by (Parkinson 1999).

Based on concerns with some existing data on GenBank for several rattlesnake species, we took multiple steps to increase our confidence in the quality of the data used in this study. First, we filled in new data from six species that seemed particularly phylogenetically unstable (based on preliminary analyses): *C. adamanteus, C. cerastes, C. enyo, C. horridus, C. polystictus* and *C. willardi.* Second, we generated new data for seven species that we identified as having questionable GenBank accessions or missing data: *C. aquilus, C. atrox, C. basiliscus, C. pricei, C. scutulatus, C. tigris* and *C. triseriatus.* Lastly, we excluded the following sequences from GenBank due to probable errors: AF259175.1 (cyt-b of *C. enyo*), HM631837.1 (ND4 of *C. horridus*) and HQ257775.1 (ND4 of *C. triseriatus armstrongi*).

Phylogenetic analysis

We aligned all sequences using ClustalW (Thompson et al. 1994). All protein-coding genes where translated to their predicted amino acid sequences to check for the presence of stop codons (none were detected). Only two individual sequences of nuclear genes had heterozygote sites, *Crotalus lannomi* (JRV-BM) and *C. scutulatus* (JAC-29076), both in the NT3 loci. We phased these sequences manually (based on re-analysis of the raw chromatogram files) and included each individual allele separately in downstream analyses. We used TOPALi version 2 (Milne et al. 2009) to test for recombination in nuclear loci using the difference of sums of squares (DSS) method with a sliding window of 100-bp and 10-bp step size. No significant recombination was detected in any of the nuclear loci. Best-fit models of nucleotide evolution for each gene (or partition) were estimated using Akaike Information Criterion (AIC) in the program JModelTest (Posada 2008). Individual gene fragments were concatenated using Sequence Matrix (Vaidya et al. 2011). When all genes were concatenated the total length of aligned positions was 3,496 bases. The final data matrix was ca. 71% complete at the level of gene loci per species, and 68% complete at the nucleotide level.

We estimated phylogenetic trees using Bayesian Metropolis-Hastings coupled Markov chain Monte Carlo inference (BI) and Maximum likelihood (ML) phylogenetic approaches using all concatenated genes. BI was used to estimate the posterior probabilities of phylogenetic trees based on a total of 10^8 generations Metropolis-coupled Markov chain Monte Carlo (MCMC) with MrBayes version 3.2.1 (Huelsenbeck and Ronquist 2001). BI analyses consisted of four simultaneous runs, each with four chains (three heated and one cold), sampled every 1,000 generations. We visualized the output from BI in the program TRACER v. 1.5 (Drummond and Rambaut 2007) to verify that independent runs had converged. Potential scale reduction factor (PSRF) estimates comparing chain likelihood values indicated convergence by 10^7 generations. We therefore conservatively discarded the first 25% of BI samples as burn-in. A majority-rule consensus phylogram was estimated from the combination of the post-burnin samples from the four BI runs. ML analysis was performed with raxmlGUI 1.3 (Silvestro and Michalak 2012). Nodal support for ML analyses was assessed using the rapid bootstrap algorithm with 10^4 replicates (Stamatakis et al. 2008).

We estimated BI phylogenetic trees in MrBayes for each individual locus separately, and also ran independent analyses for both the mitochondrial (ATP6_8, cyt-b, ND4) and nuclear (C-mos, NT3 and RAG-1) data sets. We conducted further BI analyses on the concatenated set of all loci combined. For the sake of discussion, nodes with ≥95% Bayesian posterior probabilities were considered to be strongly supported (Felsenstein 2004); in the ML analysis, nodes with ≥70% bootstrap support were considered strongly supported (Hillis and Bull 1993).

We used comparisons of tree likelihoods for different tree topologies to evaluate relative support for alternative trees. For this, we implemented the stepping-stone sampling method (Xie et al. 2011) in MrBayes v3.2 to estimate the marginal likelihood for each topological constraint. For each hypothesis, we evaluated the complete

concatenated dataset using the best-fit partitioned model based on $5x10^5$ generations of each of the 49 steps, sampling every 1,000 generations, for a total of $2x10^7$ generations.

Species tree analysis

In addition to concatenated phylogenetic inferences made using MrBayes, we also implemented a multispecies coalescent model to estimate the 'species tree' based on multi-locus data. Given the lack of substantial intraspecific sampling and the moderate amount of missing data, our dataset is not particularly well suited for species tree analysis. We therefore use species tree analyses as a means to further explore phylogenetic signal in the data, but with the above caveats. We used the program *BEAST (Heled and Drummond 2010), within the BEAST software package (Drummond and Rambaut 2007), to estimate a species tree from the three separate nuclear loci (C-mos, NT3 and RAG-1) plus a concatenated mitochondrial dataset (ATP6_8, cyt-b and ND4) that was treated as a fourth locus. We used a relaxed molecular clock model for all loci and an HKY + Γ model of nucleotide substitution for each data partition, with the exception of RAG-1, for which we used a JC model. We chose the models of nucleotide substitution based on the Akaike Information Criterion (AIC) estimated using JModelTest (Posada 2008). The tree prior was set to the Yule process, and other priors in *BEAST were set to default values. Analyses were run in duplicate, each for $1x10^9$ generations, sampling every 20,000 generations, for a total of $5x10^4$ sampled trees. We used TreeAnnotator v1.7.4 (Drummond and Rambaut 2007) to discard the first 10% of the samples as burn-in, and to map nodal support for the remaining samples on the tree.

Allele networks

Parsimony haplotype networks for the nuclear genes C-mos and NT3 data sets of the longtailed rattlesnakes were calculated using TCS 1.21 (Clement et al. 2000). All three species of longtailed rattlesnakes shared an identical haplotype of the Rag-1 gene,

so it was excluded from this analysis. Haplotype networks were inferred using a statistical parsimony framework (Templeton 1998), with gaps treated as missing data and a connection limit of 95%. Identical sequences were collapsed into a unique haplotype set.

Phyllogenetic hypothesis test

To evaluate relative evidence for different hypotheses regarding the phylogenetic placement of the longtailed rattlesnakes among *Crotalus* species, and the relationships among the three longtailed species, we used Bayes Factors in MrBayes to compare the likelihood of alternative trees based on the concatenated dataset. We used the criterion of 2ln [*bf*] ranging from two to six as positive evidence, six to ten as strong evidence and >10 as very strong evidence against the alternative hypotheses (Kass and Raftery 1995; Miller and Bergsten 2012).

Divergence dating

We performed a likelihood ratio test to test the null hypothesis that substitutions in the genes used follow a strict molecular clock of evolution. At a significance threshold of $p < 0.05$, the set of all mitochondrial genes rejected the strict molecular clock, while the set of all nuclear genes failed to reject it. The concatenated analysis of all genes also rejected the strict molecular clock. We estimated divergence times across the rattlesnake phylogeny using BEAST 2 (Drummond and Rambaut 2007) instead of incorporating divergence estimation in our *BEAST analysis of the species tree. We took this approach because our species tree analysis had considerable missing data (>30%), which presumably contributed to the failure of these *BEAST runs to reach convergence in >1 billion generations. Additionally, most nodes in our species tree analysis had extremely low support values. Because preliminary analyses in BEAST 2 implementing nucleotide models partitioned across genes and codon position showed poor convergence, our final analysis used an unpartitioned model to estimate divergences using the entire dataset.

The concatenated data set rejected the strict clock hypothesis, so we implemented an uncorrelated log-normal relaxed clock model with a Yule tree prior using the HKY + Γ model of sequence evolution applied to the combined data set. Two independent analyses were run for $1x10^8$ generations, sampling every 10,000 generations. Dates used to constrain nodes were obtained from estimates based on the fossil record or biogeographic divergence events published in previous studies (Holman, 2000; Castoe et al., 2007; (Parmley and Holman 2007), and many of the priors we use here follow a recent study that has dated a similar phylogenetic tree (Bryson et al. 2011b). We used the program Tracer v. 1.5 (Drummond and Rambaut 2007) to confirm stationarity of the Markov chain Monte Carlo (MCMC) analysis, adequate effective sample sizes of the posterior (>200 for each estimated parameter), and the appropriate percent to discard as burn-in (which we estimated conservatively to be 10%, or 1,000 trees). We used two fossil and one geologic calibration for our divergence estimates. First, we used the oldest *Sistrurus* fossil (Late Miocene, Claredonian; (Parmley and Holman 2007). We constrained the ancestral node of *Sistrurus* with a zero offset of 8 million years ago (mya), with a log-normal mean of 0.01, and a log-normal standard deviation of 0.76, resulting on a median age of 7 my and a 95% prior credible interval (PCI) that extended to the Late Clarendonian, ~11.5 mya (Holman, 2000). Second, we used the oldest fossil of *Agkistrodon contortrix* (Late Hemphillian; (Holman 2000). This node was constrained with a zero offset of 6 mya, a log-normal mean of 0.01, and a log-normal standard deviation of 0.42, resulting on a median age of 7 my and a 95% PCI that extended to the Late Hemphillian, ~8 mya (Holman, 2000). Third, we used the estimated time of divergence between *C. atrox* and *C. ruber* as approximately 3.2 mya (Castoe et al. 2007b). This node was given an offset of 3.2, a normal mean of 0 and a normal standard deviation of 1, resulting on a median age of 3.2 my and a 95% PCI that extended to ~4.8 mya. After discarding burn-in samples, the trees and parameter estimates from the independent runs were combined using LogCombiner v. 1.7.4 (Drummond and Rambaut

2007). We summarized parameter values of the samples from the posterior on the maximum clade credibility tree using the program TreeAnnotator v. 1.7.4 (Drummond and Rambaut 2007).

Revision of skeletal material

The absence of teeth in the palatine bone has been considered a synapomorphy uniting *Crotalus polystictus* and *C. stejnegeri* (Klauber, 1952; Brattstrom, 1964). To re-evaluate this supposition, we looked for the presence or absence of teeth in the palatine bone in the skulls of specimens of ten species of the genus *Crotalus,* as well as one species of each of the genera *Sistrurus* and *Agkistrodon*. All specimens are deposited at the UTA-ARDRC. A list of the specimens examined and their locality data is given in Supplementary Table 4.

Results

DNA sequence characteristics

The combined set of mitochondrial loci contained 1610 bp, 801 of which were variable. The total length of ATP6_8 was 444 bp, with 245 (45%) variable sites. For cyt-b, the total length was 564 bp, with 260 (46%) variable sites. The total length of ND4 was 602 bp, with 296 (49%) variable sites. The combined set of nuclear loci contained 1887 bp, 91 of which were variable. The C-mos fragment contained 553 bp, with 29 (5%) variable sites. NT3 had a total length of 512 bp, 41 sites (8%) were variable. RAG-1 had a length of 822 bp, and only 21 sites (2%) were variable.

Individual gene tree estimates

There was broad support that the three longtailed rattlesnake species formed an exclusive clade across BI trees estimated from individual loci, although there were several differences in topology between individual gene trees (figures not shown).

Nuclear genes had a low number of polymorphic sites and tended to provide lower phylogenetic resolution and support (see above). Longtailed rattlesnakes were recovered as monophyletic in all BI trees based on analyses of individual genes except for that based on the nuclear gene C-mos, which resulted in a polytomy that included the longtailed rattlesnakes, *C. horridus* and *C. molossus.* In the case of NT3, *C. ericsmithi* was nested within a cluster of *C. lannomi* samples, and *C. stejnegeri* was sister to this clade.

The relationships among the three species of longtailed rattlesnakes differed somewhat among BI phylogenetic estimates based on individual loci. A clade containing *C. lannomi* and *C. stejnegeri,* sister to *C. ericsmithi,* was inferred based on BI analysis of the mitochondrial loci ATP6_8 and ND4 (posterior probability [pp] = 1). In contrast, a clade containing *C. lannomi* and *C. ericsmithi* as the sister lineage to *C. stejnegeri* was inferred based on the mitochondrial cyt-b fragment and the nuclear fragments NT3 and RAG-1 (pp = 0.99, 1 and 0.73, respectively).

The phylogenetic placement of the longtailed rattlesnake clade within the phylogeny of the rattlesnakes was weakly and differentially resolved by individual gene BI estimates. The longtailed rattlesnake clade formed a polytomy with several other rattlesnake lineages based on ATP6_8. The ND4 BI tree recovered the longtailed clade as the sister-lineage to *C. horridus* plus *C. cerastes* (pp = 0.68), and cyt-b showed a different topology in which the longtailed clade formed the sister lineage to all other *Crotalus* species (pp = 0.98).

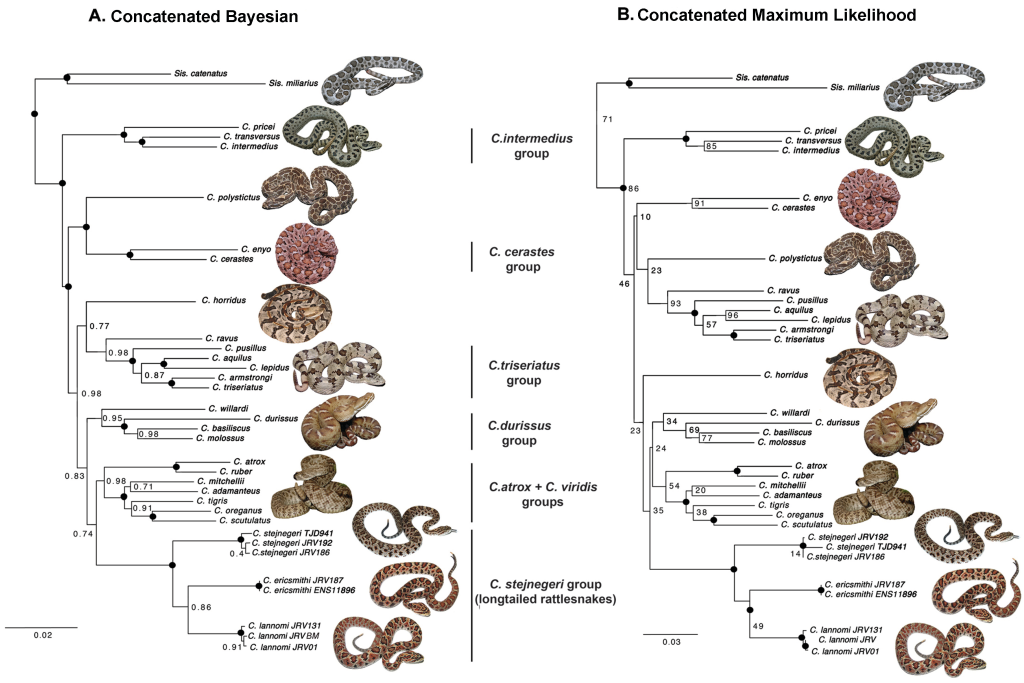


Figure 2 - Phylogenetic estimates based on concatenated data analyses. A. Majority-rule consensus tree from Bayesian phylogenetic tree estimates based on all genes concatenated, with bipartition posterior probabilities indicated by numbers or a filled circle if equal to 1.0. B. Maximum likelihood phylogenetic tree estimate based on all genes concatenated, with bipartition bootstrap values indicated by numbers or a filled circle if equal to 100%. Outgroups are omitted from figures.

Concatenated phylogenetic analyses

The concatenated nuclear dataset analyzed using BI (not shown) recovered C. stejnegeri as sister to a clade containing *C. lannomi* and *C. ericsmithi*, with a single representative of *C. ericsmithi* nested within a clade of three *C. lannomi* samples. The longtailed species were one of the only *Crotalus* clades with posterior support > 0.95 (the other groups being *C. tigris* + *C. oreganus* and *C. basiliscus* + *C. polystictus*), although their placement among other lineages of *Crotalus* was unresolved. The BI analysis of concatenated mitochondrial genes (not shown) also recovered the longtailed rattlesnakes as monophyletic, but with *C. ericsmithi* as sister to a clade comprising *C. lannomi* plus *C. stejnegeri* (pp = 1); this longtailed rattlesnake clade was inferred to be the sister group to the *C. durissus* + *C. atrox* + *C. viridis* groups (pp = 0.96).

When all genes where combined for BI and ML analyses, a slightly different topology was recovered (Fig. 2). The monophyly of the longtailed group was strongly supported in both BI and ML analyses (pp = 1; bootstrap support [bs] = 100%), with *C. stejnegeri* as the sister lineage to a clade containing *C. lannomi* and *C. ericsmithi*. In the BI estimate, the longtailed group was supported by 0.74 posterior probability as sister to the *C. atrox* + *C. viridis* groups, while the ML tree placed the longtailed rattlesnakes sister to a clade consisting of the *C. atrox*, *C. viridis* and *C. durissus* groups (like the BI tree of mitochondrial genes). Another difference between the BI and ML inferences was the position of *C. horridus*, which was the sister to the *C. triseriatus* group in the BI tree, but recovered as sister to a clade containing the *C. durissus, C. atrox*, *C. viridis, and* longtailed rattlesnakes in ML. Both BI and ML inferred that *C. enyo* and *C. cerastes* formed a clade (the *C. cerastes* group), but they differed in their placement of *C. polystictus*, which was sister to the *C. triseriatus* group in ML, and sister to the *C. cerastes* group in BI. In both analyses, the *C. intermedius* group was recovered as the sister group to all other species of *Crotalus*.

Species tree analysis

The species tree analysis of all loci using *BEAST recovered an exclusive longtailed rattlesnake species clade, with *C. stejnegeri* sister to *C. ericsmithi*+*C. lannomi*; this clade was recovered with strong support (pp = 0.99; Fig. 3). Contrary to the BI and ML analyses, the longtailed rattlesnakes were placed as sister to *C. horridus*, and this clade was the sister group to the *C. atrox* plus *C. viridis* groups. Unlike results from concatenated BI and ML analyses, species tree analyses implied that the *C. cerastes* group and *C. polystictus* formed a clade sister to the *C. durissus* group.

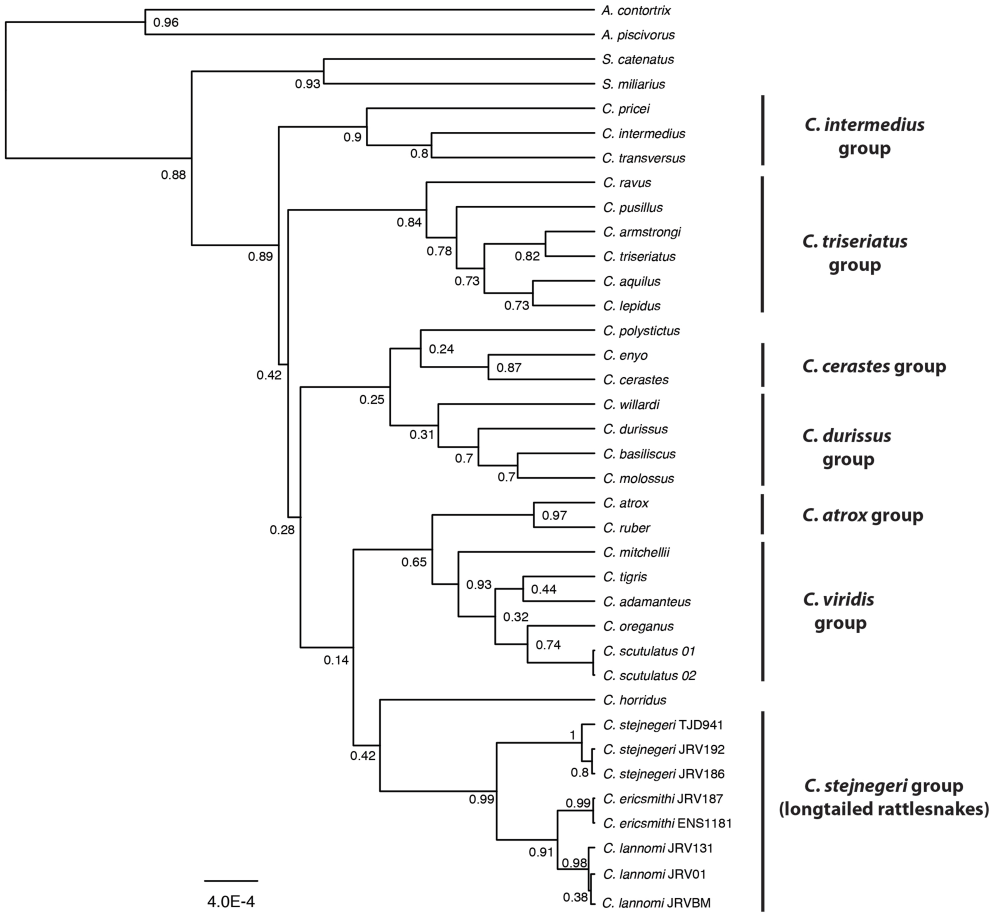


Figure 3 - Species tree estimate for the rattlesnakes based on analysis using *BEAST incorporating all six gene fragments (ATP6_8, C-mos, cyt-b, ND4, NT3 and RAG-1). Posterior probability values are given adjacent to respective nodes.

Allele Networks

We found no evidence for recombination within any of the three nuclear genes within the longtailed rattlesnake samples. For these longtailed rattlesnake samples, the six sequences of C-mos grouped into three distinct haplotypes, each haplotype was unique to each of the three species. In the case of NT3, the eight individuals grouped into 4 different haplotypes. *Crotalus lannomi* had two distinct haplotypes, each of which were homozygous in one individual, and heterozygous in a third individual. Samples of *C. ericsmithi* and *C. stejnegeri* were homozygous for a single variant unique to each species (Fig. 4). In sum, within the longtailed rattlesnake species, all nuclear variants observed are unique to a recognized species (as are all mitochondrial variants).

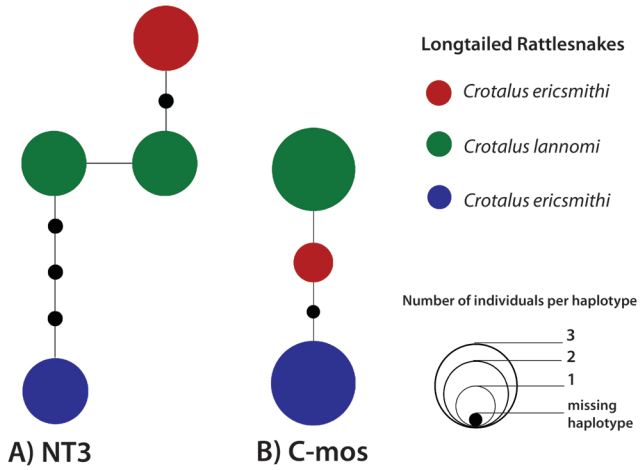


Figure 4 - Allele network for the variable nuclear genes (NT3 and C-mos) constructed for the longtailed rattlesnakes. All specimens of longtailed rattlesnakes shared the same RAG-1 haplotype, so it was excluded from this analysis.

Phylogenetic hypothesis test

Because different analyses resulted in different phylogenetic estimates, we tested two sets of hypotheses regarding the relationships of the longtailed rattlesnakes. Set A, which represents hypotheses regarding the placement of the longtailed

rattlesnakes among rattlesnakes: HA1) longtailed rattlesnakes sister to *C. atrox* + *C. viridis* groups – this topology was obtained from BI analysis of all genes; HA2) longtailed rattlesnakes sister to the *C. durissus* group – this topology was obtained in some of the BEAST runs; HA3) longtailed rattlesnakes sister to a clade containing the *C. durissus, C. atrox, and C. viridis* groups – this topology was recovered by the BI analysis of mitochondrial genes and RaxML analysis of all genes; HA4) longtailed rattlesnakes sister to *C. horridus* – recovered in the species trees analysis in *BEAST, although with very low support. Our second set (set B) focused on the branching order of the three longtailed rattlesnakes: HB1) *C. lannomi* sister to *C. stejnegeri* – recovered from BI analysis of individual ATP6_8 and ND4 genes. HB2) *C. lannomi* sister to *C. ericsmithi* – obtained from all other analyses. In tests of these hypotheses using Bayes factors [*bf*] based on the concatenated dataset, we found strong support (*bf* = 6.6 – 17.5) for the longtailed rattlesnakes as sister to the *C. atrox* plus *C. viridis* groups (HA1), but we found no notable support (*bf* = 1.7) for a particular branching order among the three longtailed rattlesnakes (Fig. 5).

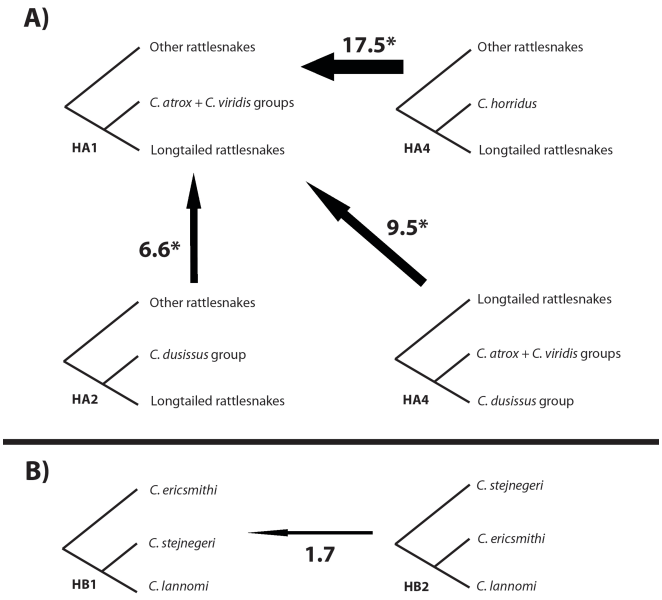


Figure 5 - Topological hypotheses tested for the placement and branching order of the longtailed rattlesnakes. HA 1-4: Hypothesis for which lineages are the sister group to the longtailed rattlesnake clade. HB 1-2: Hypotheses for the branching order among the three species of longtailed rattlesnakes. Arrows point towards the hypothesis that is favored by Bayes Factors. Numbers represent relative support based on Bayes factors (2ln [bf]) between topologies trees, which are considered as positive evidence for a particular topology if they range from two to six, strong evidence from six to ten, and as very strong evidence if > 10.

Divergence time estimates

Our divergence estimates are similar to previous studies of pitviper evolution (e.g. (Douglas et al. 2002; Daza et al. 2010; Bryson et al. 2011b), which is expected because many calibration points, and much sequence data are shared with these studies. Due to the lack of substantial intra-specific sampling, missing data, and strong support for the topology recovered in concatenated BI analyses (Fig. 5), we base our divergence time estimates on concatenated (non species tree) BI analysis. Based on the

divergence time analysis implemented in BEAST 2, we estimate that the split between the *C. intermedius* group and the rest of *Crotalus* occurred ~9.9 mya (7.8 to 12.3 mya, 95% highest posterior densities [HPD]). Following this event, most other major lineages of *Crotalus* (i.e., species groups) diverged in relatively rapid succession during the late Miocene, from ~9 to 6 mya (Fig. 6). Our estimates of the divergence dates among most *Crotalus* lineages are mostly similar to previous studies (e.g. (Douglas et al. 2006); (Bryson et al. 2011a; Bryson et al. 2011b); (Anderson and Greenbaum 2012)), with the exception of the divergence between *C. durissus* and *C. molossus*, as our estimate is substantially more recent than previous estimates (Wuster et al. 2005). We estimate that the ancestor of the longtailed rattlesnake group diverged from a common ancestor with the *C. atrox* + *C. viridis* group clade during the late Miocene, ~6.8 mya (5.1 to 8.6 mya, 95% HPD). The extant longtailed rattlesnake lineages are estimated to have split from one another during the Pliocene (Fig. 6), with the first division occurring when *C. stejnegeri* diverged from the other two longtailed species ~3.96 mya (2.5 to 5.46 mya, 95% HPD), followed by the divergence of *C. lannomi* from *C. ericsmithi* ~2.7 mya (1.6 to 4.1 mya, 95% HPD).

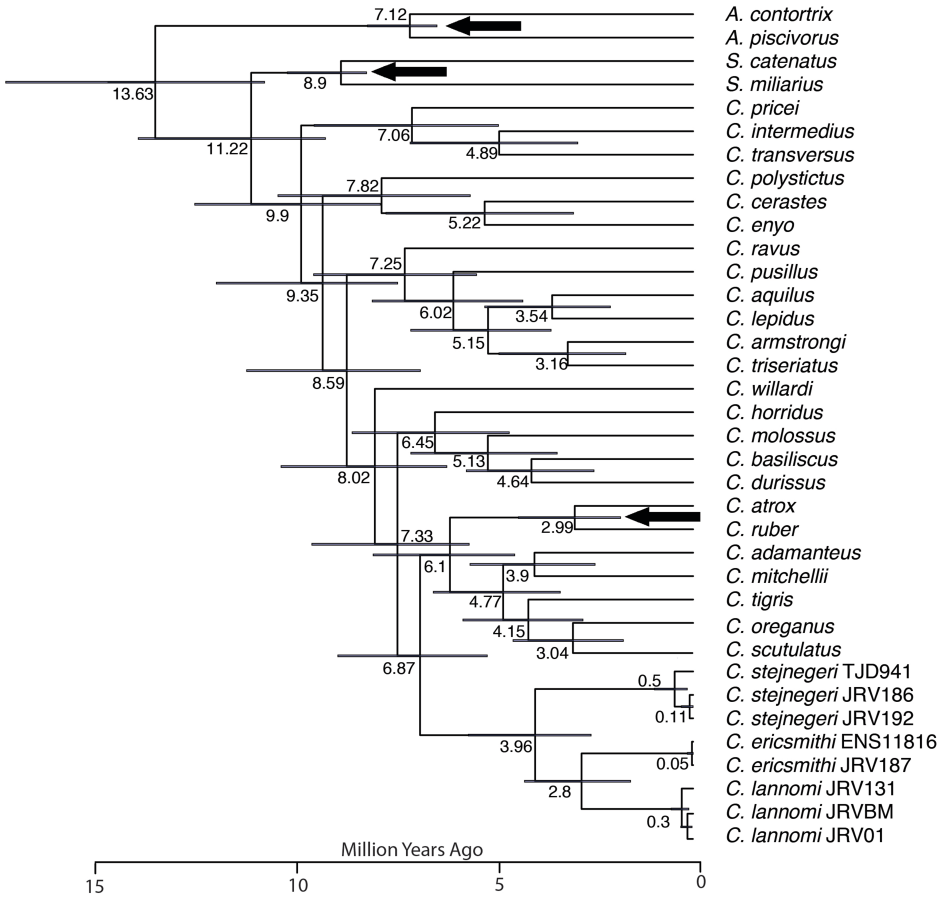


Figure 6 - Bayesian relaxed clock estimate of divergence times among rattlesnake lineages with 95% credibility intervals shown over nodes by shaded bars. Dark arrows represent calibration points used in the analysis.

Revision of skeletal material

Among the pitviper species examined for palatine teeth, the only species without teeth in the palatine bone is *C. polystictus,* and this trait was consistent across three specimens. *Crotalus stejnegeri* was reported by Klauber (1956) and Brattstrom (1964) to lack teeth on the palatine bone, but the specimen we examined (UTAR-10499) had three palatine teeth (Fig. 7).

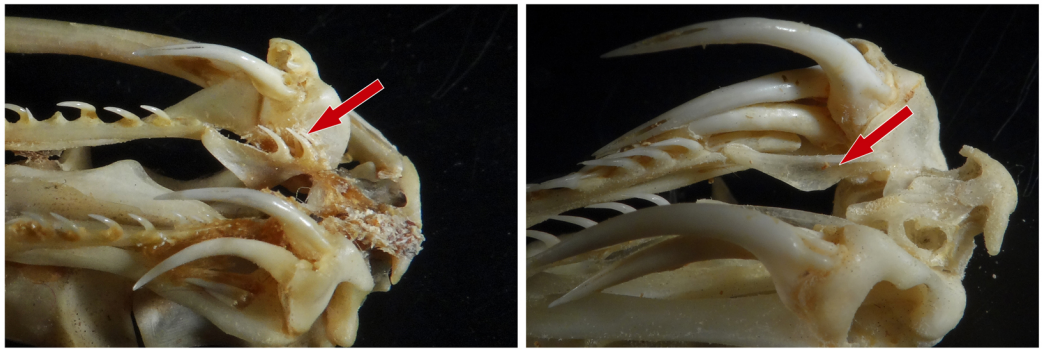


Figure 7 - Photographs of skulls of *Crotalus stejnegeri* (left) and *C. polystictus* (right). Red arrows point to the palatine bone. Notice the presence of palatine teeth in *C. stejnegeri* and their absence in *C. polystictus*.

Discussion

Monophyly and distinctiveness of the longtailed rattlesnakes

The importance of rattlesnakes transcends academic interests in many ways, including their medical importance and their central role in the imagery and folklore of North America (Greene and Cundall 2000). Furthermore, this group of 36 species is one of the most heavily studied lineages of reptiles, particularly when their relatively low diversity (equivalent to ~1% of all snake species) is considered. Among rattlesnake species, the longtailed rattlesnakes have remained the most enigmatic, largely because of the dearth of scientific material available for these species (e.g., a single specimen for *C. lannomi* for almost 50 years) and the recent discovery of *C. ericsmithi* (Campbell and Flores-Villela 2008). Thus, in the absence of sufficient comparative material, the origins, distinctiveness, and relationships among longtailed rattlesnakes have been much debated but insufficiently tested.

Our phylogenetic estimates provide unilateral evidence that the longtailed rattlesnakes form a well-supported monophyletic clade (Figs. 2–6). Most authors have considered the long tail of these species to be an ancestral character state (Gloyd 1940;

Klauber 1952; Tanner 1966; Klauber 1972), and therefore not a synapomorphy supporting the monophyly of the group (Campbell and Flores-Villela 2008). Our results instead indicate that the long tail condition is a shared derived character uniting these three species, as is the mediolateral compression of the hemipenial lobes (Jadin et al. 2010).

Although each of the three longtailed rattlesnake species share characteristics of their internal and external morphology (Jadin et al. 2010); (Reyes-Velasco et al. 2010); Reyes-Velasco, unpublished), we find each to constitute reciprocally monophyletic groups based on all mitochondrial gene analyses, analysis of the nuclear gene C-mos, and the species tree analysis of the combined data (Figs. 2-3). Furthermore, for nuclear genes that show variation in these three species (NT3 and C-mos), each species contains species-specific alleles and no alleles are shared among species (Fig. 4). We estimate that the three species have most likely diverged from one another during the Mid-Late Pliocene (Fig. 6). Based on our analyses, together with previous evidence for their morphological distinctiveness, there is broad agreement that these three species are indeed distinct.

Phylogenetic placement of the longtailed rattlesnakes

The long, attenuated tails and minute rattles characteristic of species of the longtailed rattlesnakes have led most researchers to conclude that these species where the sister group to all other *Crotalus* (Gloyd 1940; Klauber 1952; Tanner 1966; Klauber 1972). Based on morphological similarities, including high ventral counts, high number of dorsal scale rows, and a tendency toward subdivided head scales, Klauber (Klauber 1952) noted that *C. stejnegeri* more closely resembled *C. viridis* and *C. atrox* than other rattlesnakes. Longtailed rattlesnake species also, however, possess high numbers of spines on each hemipenial lobe, a trait that they shared with *C. polystictus* (Jadin et al. 2010). Further linking *C. polystictus* and *C. stejnegeri*, the absence of teeth on the palatine bone was considered a synapomorphy uniting these two species (Klauber 1956;

Brattstrom 1964; Klauber 1972), although LaDuc (2003) reported palatine teeth from a specimen of *C. stejnegeri* (UTA R-10499). We reexamined this specimen as well as various other rattlesnake species (including *Sistrurus catenatus, Crotalus aquilus, C. atrox, C. lepidus, C. molossus, C. pricei, C. stejnegeri, C. willardi*, and several *C. polystictus;* see Supplementary Table 4), and the lack of palatine teeth was found to be unique to *C. polystictus,* and the presence of palatine teeth in *C. stejnegeri* was confirmed (Fig. 7). Due to the lack of comparative skeletal material, we were not able to assess the presence of palatine teeth in *C. lannomi* and *C. ericsmithi*. The absence of teeth in the palatine bone is therefore an autapomorphy of *C. polystictus* and not a synapomorphy linking *C. polystictus* and *C. stejnegeri.*

The ML analysis of all genes placed the longtailed rattlesnakes as sister to a clade consisting of the *C. durissus* + (C. *atrox* and *C. viridis*) groups, but with little support (bs = 35%). Concatenated BI analysis estimated a sister relationship between the longtailed rattlesnakes and the *C. atrox* + *C. viridis* groups, but with relatively weak support (pp = 0.74). Our species tree inference from *BEAST resulted in the longtailed rattlesnakes placed as the sister to *C. horridus*, but with extremely low support (pp = 0.42), as was recovered at most other nodes of this tree (Fig. 3). Because we inferred multiple competing hypotheses for relationships of longtailed rattlesnakes across different phylogenetic methods, we tested these hypotheses using Bayes Factors implemented in MrBayes based on the concatenated data set. Our results strongly favored the sister relationship between the longtailed rattlesnakes and the *C. atrox* + *C. viridis* groups (Fig. 5), as was inferred by the BI concatenated analysis. The close relationship between the longtailed rattlesnakes and the *C. atrox* and *viridis* groups has never been explicitly inferred by phylogenetic analyses, although there are several similarities between these groups of rattlesnakes that others have noted (Klauber 1952). We find strong evidence countering previous hypotheses that the longtailed rattlesnakes are sister to all other *Crotalus*, and also against the hypothesis that they are close relatives of *C. polystictus*,

as previously suggested based on hemipenial characters (Jadin et al., 2010) and the presumed synapomorphy of the absence of palatine teeth that we confirm to have been incorrect (Fig. 7).

Insights into rattlesnake phylogeny

Estimating the phylogenetic placement of the longtailed rattlesnake clade within the context of rattlesnake phylogeny requires at least partial resolution of the phylogeny of rattlesnakes, which has historically been difficult. Although our sampling of *Crotalus* species was not exhaustive, we included multiple taxa from all major rattlesnake species groups, together with new data for other lineages, and recovered several well-supported clades within *Crotalus* (Fig. 2). Our phylogenetic results are largely congruent with many previous hypotheses (Castoe and Parkinson 2006; Bryson et al. 2011a; Bryson et al. 2011b), although there are some notable differences. Because our data and species coverage allow us to make inferences that were previously untenable, we briefly discuss salient findings below.

In contrast to other molecular studies (Murphy et al. 2002; Castoe and Parkinson 2006), our data provided support for the *C. intermedius* group as sister to all other species of *Crotalus* (combined data: pp ≥ 0.95, bs ≥ 70%; Fig. 2). We inferred that the next lineage to diverge from the remaining species of *Crotalus* is a clade containing *C. polystictus*, *C. enyo*, and *C. cerastes,* with these last two forming a clade. A close relationship between *C. enyo* and *C. cerastes* is not novel, and has been previously suggested by analyses of venom electrophoresis, skull morphology and molecular data (Minton 1956; Brattstrom 1964; Douglas et al. 2006). While support for the sister relationship between *C. enyo* and *C. cerastes* was consistently high in BI and ML concatenated analyses (pp ≥ 0.95, bs ≥ 0.90), the sister relationship between *C. polystictus* and the *C. cerastes* group was not supported by the ML analysis, which instead placed *C. polystictus* as sister to the *C. triseriatus* group with extremely low

support (bs = 23%). The instability of support values and topology suggests that the inclusion of *C. polystictus* within this clade is tentative, and may be an artifact of long-branch attraction (Bergsten 2005). *Crotalus enyo* had previously been assumed to be the northernmost member of the neotropical rattlesnake (*C. durissus*) group (Murphy et al. 2002; Castoe and Parkinson 2006), but our results strongly support the exclusion of *C. enyo* from this group. The inclusion of *C. enyo* in the *C. durissus* group seems to be based on previous use of a single sequence of cyt-b, which our analysis suggests represents a chimeric sequence (see Supplementary Table 3). Instead of a close relationship between *C. enyo* and *C. durissus*, our results find weak to moderate support for *C. willardi* as a basally-diverging member of the *C. durissus* group (Figs. 2-3).

Our results support an expanded definition of the *C. viridis* group that includes species of the former *C. mitchellii* group as well as *C. adamanteus*; this conclusion parallels that of previous studies (Castoe and Parkinson 2006). Although this clade is strongly supported in all of our analyses, the precise order of basal divergences within this clade remains poorly resolved (Figs. 2-3). The close phylogenetic affinity of *C. mitchellii* and *C. tigris* with the *C. viridis* group has been previously suggested on the basis of morphological data (Gloyd 1940; Klauber 1956). Although it has been assumed that the two “diamondback rattlesnakes” *C. adamanteus* and *C. atrox*, might be sister taxa, the accumulation of molecular data from this and other studies (Castoe and Parkinson 2006; Pyron et al. 2011) provide evidence against this.

Early morphological studies considered *C. horridus* to be closely related to *C. molossus* (Gloyd 1940; Klauber 1956; Brattstrom 1964). More recently, (Murphy et al. 2002) recovered *C. horridus* as sister to *C. viridis* plus *C. scutulatus*, and (Castoe and Parkinson 2006) placed *C. horridus* as a lineage roughly in the center of the *Crotalus* radiation. Our BI analysis of mtDNA sequences and ML analysis of all genes supported *C. horridus* as sister to a clade of “derived” rattlesnake species groups (*C. atrox*, *C. durissus*, *C. stejnegeri,* and C*. viridis* groups). This node, however, was not strongly

supported in our ML results (bs = 23%), similar to other previous studies (Castoe and Parkinson 2006). In contrast, the BI analysis of combined data placed *C. horridus* as the earliest diverging lineage within the *C. triseriatus* group with moderate support (pp = 0.77; Fig. 2), while the species tree analysis in *BEAST placed this species as sister to the longtailed rattlesnakes, but with almost no support (pp = 0.42; Fig. 3). Despite substantial progress, including contributions of this study, the phylogeny of the rattlesnakes is far from resolved, and the phylogenetic relationships of several rattlesnake taxa should be re-evaluated with additional loci and perhaps even additional sampling. Lineages that are particularly in question with regard to their placement on the rattlesnake tree include *C. horridus, C. polystictus* and *C. willardi,* as well as the *C. cerastes* group (*C. cerastes* and *C. enyo*).

Divergence and biogeography of the longtailed rattlesnakes

During the Pliocene, major volcanism occurred in what is now the boundary between the Mexican states of Jalisco and Nayarit, between the Río Grande de Santiago and Ameca rivers (Frey et al. 2007). This period of volcanic activity extended from 5 to 3 mya, which coincides with our estimates of the time that *C. stejnegeri* diverged form the ancestor of the two southern species of longtailed rattlesnakes. Regional changes in habitat distributions associated with these periods of volcanism may have split the putative ancestor of *C. stejnegeri* from the ancestor of *C. lannomi* + *C. ericsmithi* (Figs. 1, 6). On the other end, the Balsas Basin has been implicated as an important biogeographic barrier for other vertebrate groups, including snakes (Devitt 2006; Bryson et al. 2008), mammals (Amman and Bradley 2004) and birds (Navarro-Siguenza et al. 2008). At the heart of the Balsas Basin, the Rio Balsas is currently located at the border between the states of Michoacán and Guerrero and is a likely candidate for causing the divergence between ancestral lineages of *C. ericsmithi* and *C. lannomi* (Figs. 1, 6).

Longtailed rattlesnake species tend to occur at middle elevations in tropical deciduous and tropical oak forests (Campbell and Lamar 2004a; Campbell and Flores-Villela 2008; Reyes-Velasco et al. 2010). One of the most intriguing regions not yet thoroughly examined for the presence of these enigmatic snakes is the Sierra de Coalcomán, which is a small coastal mountain range in the state of Michoacán, West of the Rio Balsas. Although no longtailed rattlesnake species have been recorded from this locality, convincing reports from local residents indicate that a population of longtailed rattlesnake is likely to exist there. As additional collections are made in the region, it is therefore possible that yet another population of longtailed rattlesnakes will be discovered that may represent a new species, or possibly a population allocable to either *C. lannomi* (which is known from ca. 150 km away), or to *C. ericsmithi*, found farther to the southeast.

Conclusions

Our results provide new conclusive evidence for the distinctiveness, monophyly and phylogenetic placement of the longtailed rattlesnakes. A well-resolved phylogeny for the rattlesnakes has been elusive despite a substantial number of studies that have addressed this conspicuous group (e.g. (Parkinson 1999; Murphy et al. 2002; Castoe and Parkinson 2006; Pyron et al. 2011). By adding new data from the three most rare and enigmatic species of *Crotalus*, this study contributes important sampling for resolving *Crotalus* phylogeny. We also identified multiple instances where errors in GenBank submissions might have contributed to poor and conflicting resolution in previous studies. The fact remains, however, that although many studies have inferred rattlesnake phylogenies, most have essentially used a common set of data from a few mitochondrial and nuclear gene loci that (in some cases) have existed for more than a decade. We expect that definitive resolution of the phylogeny of rattlesnakes will ultimately require a

new influx of molecular data to resolve remaining questions about the relationships among major *Crotalus* lineages and species groups.

Chapter 2

Molecular systematics of coralsnakes of the *Micrurus diastema* species complex

Introduction

The highly venomous coralsnakes of the family Elapidae comprise a diverse radiation of more than 170 taxa distributed in Southeast Asia and the New World (McDiarmid et al. 1999; Campbell and Lamar 2004b; Castoe et al. 2007a). Coralsnakes are thought to have invaded the New World from Asia via a Beringian land bridge connecting Asia and North America during the late Oligocene (Kelly et al. 2009), similar to other major lineages of New World snakes (Holman 2000; Castoe et al. 2007a; Guo et al. 2012). Since their colonization of the New World, coralsnakes have diversified extensively across the Americas into three genera (*Micruroides, Micrurus* and *Leptomicrurus*) and approximately 80 species, and are currently distributed from Florida to Argentina (Campbell and Lamar 2004b). The genus *Micruroides* is composed of a single species (*M. euryxanthus*) and three subspecies, that occur in western North America (Campbell and Lamar 2004b), while the genus *Leptomicrurus* (sometimes considered a synonym of *Micrurus* (Slowinski 1995; Uetz and Jirí 2015) consists of four South American species. The genus *Micrurus* contains the majority of New World coralsnakes, with ~80 recognized species, many of which contain several recognized subspecies (Uetz and Jirí 2015). Despite the species diversity and broad distribution the genus *Micrurus*, the external morphology across species is highly conservative. This lack of external morphological variation has led to a taxonomy for species of *Micrurus* being highly dependent on color and color pattern variation (Roze 1967).

Although the relationships between the major lineages of coralsnakes and other elapids snakes has been relatively well-studied (Keogh 1998; Slowinski and Keogh 2000; Castoe et al. 2007a; Kelly et al. 2009), our understanding of the evolutionary relationships among members of the genus *Micrurus* are poorly known. The majority of

phylogenetic studies that have focused on *Micrurus* species have been based on external or internal morphological characters (e.g. color, scalation, immunological assays, hemipene morphology, etc.), and only a small number of limited studies have used molecular data to infer relationships (e.g. (Slowinski and Keogh 2000; Castoe et al. 2007a; Fry et al. 2010). Recent systematic accounts of the genus *Micrurus* have divided species into three main groups that are defined based on their color patterns: the monadal and bicolor group, the Central American triadal group, and the South American triadal group (Campbell and Lamar 2004b). Monadal coralsnakes are defined by a pattern of banding consisting of a single black ring followed by a yellow and a red ring. The bicolor group is composed of a few species that show a bicolor pattern of dark and pale rings, while the color pattern of both triadal groups consist of two black rings interspaced with pale colored rings, followed by a red ring. These major banding-pattern-based groups are often subdivided into smaller groups thought to represent clades of related species, for example, the *M. diastema, M. fulvius* and *M. nigrocinctus* species groups within the monadal coralsnakes (Lavin-Murcio and Dixon 2004; Castoe et al. 2012).

The *Micrurus diastema* species group of coralsnakes is composed of 13 species and multiple subspecies that range from the southern USA to Honduras, but the majority of the species and subspecies occur in southern Mexico and northern Central America (Campbell and Lamar 2004b). This species group currently lacks any formal taxonomic classification, however many of its members at one time were considered as conspecific or synonyms of *M. diastema* (Zweifel 1959; Roze 1967). There is considerable color variation among the members of this group, both among and within species. Some species show substantial variation in color and color pattern across their range, and sometimes within a population, while other species that are allopatric and presumably distantly related may possess very similar color patterns (see Figure 8). This variation in

color pattern is especially problematic when delimiting species because distinctions among species are based heavily on color pattern in *Micrurus*.

Color patterns are not only an important characteristic in delimiting currently recognized coralsnake species, it is also important in multiple types of mimicry systems and is likely under complex patterns of selection (Pfennig et al. 2001; Harper and Pfennig 2008). Coralsnakes are highly venomous and brightly colored, which has led to the conclusion that their coloration is aposematic (Brodie III 1993). An interesting aspect of coralsnake biology is the existence of

Figure 8 – Color variation in members of the *M. diastema* species complex from Mexico. *Left, M. distans oliveri*: top, Maruata, Michoacán; middle, Paticajo, Colima; bottom, Ixtlahuacán, Colima (C.I. Grünwald). *Right:* Top, *M. proximans*, Montitlán, Colima; middle, *M. sp.* Quesería, Colima; bottom, *M. browni,* Agua Fria, Colima (J.M. Jones). All specimens were found <150 km from one another.

distantly-related co-distributed snake species of non-venomous snakes that mimic coralsnake color patters, and thus presumably ward off predators that mistakenly avoid them because they misidentify them as venomous coralsnakes (Greene and McDiarmid 2005). Although some of the earliest references to mimicry involve coralsnakes (Wallace 1871), mimicry systems in snakes are still not well understood (Greene and McDiarmid 1981). Many studies of coralsnake mimicry systems focus on the intensity of predation to particular colors and patterns in clay models; these studies have shown that different colorations in snake clay models can greatly affect the rate of attacks by avian predators (Brodie III 1993; Brodie III and Janzen 1995), and that several avian species are innately predisposed to avoid certain patterns and colors on their prey, while others are not (Smith 1975, 1977; Brodie III 1993; Hinman et al. 1997; Sherbrooke et al. 2006). Historically, mimicry in coralsnakes was divided into two types: (I) Batesian mimicry, in which non-venomous or slightly venomous snakes mimic a highly venomous snake model (Bates 1862); and (II) Müllerian mimicry, were two allopatric venomous species resemble each other (Müller 1879). These two types of mimicry are now seen as the ends of a continuum, and not as two distinct types of mimicry (Greene and McDiarmid 2005). One of the greatest limitations preventing a thorough analysis of hypotheses related to the evolution of coral snake mimicry systems is the lack of a robust phylogeny and a stable taxonomy for the group. Among such hypotheses is the question of whether inter- and intra-specific color variation in coralsnakes might be driven by species interactions.

Here we use mitochondrial gene sequences and genome-wide SNPs to infer phylogenetic relationships, patterns of gene flow, and species boundaries within the *M. diastema* species complex. We use these data to address the following questions: 1) Is the *M. diastema* species complex a monophyletic group? 2) Does the current taxonomy reflect evolutionary relationships, and how many species should be recognized within this

species complex?, and 3) Does variation in color pattern reflect phylogenetic divergence and indicate species boundaries?

Materials and methods

Taxon sampling and DNA extraction

We obtained tissues from a total of 117 *Micrurus* coralsnakes from all species and subspecies of the *diastema* species complex (Fig. 9), as well as multiple taxa used as outgroups obtained from Genbank (Supplementary table 1). Tissues included samples of blood, liver, skin or shed skin preserved by snap freezing, lysis buffer or RNALater. Genomic DNA was isolated by one of four methods: using a Qiagen DNeasy extraction kit (Qiagen, Inc., Valencia, CA, USA), Zymo Research Genomic DNA Tissue MiniPrep kit (Zymo Research Corporation, Irvine, CA, USA), by standard phenol-choroform-isoamyl alcohol extraction, or with the use of AgenCourt Ampure XP DNA beads (Beckman Coulter, Inc., Irving, TX, USA).

Mitochondrial and nuclear locus amplification and sequencing

We used PCR to amplify a fragment of the mitochondrially-encoded NADH dehydrogenase subunit 4 (ND4), with the use of the primers ND4 and Leu (Arevalo et al. 1994). We also amplified and sequenced the nuclear recombination-activating gene (RAG-1) for a subset of the data (n = 9) in two overlapping fragments using the following primer sets: RAG-1-tc0225F (GCA GCT GTA ATG TCA CAA GTG C) and RAG-1-tc2000R (TTA CAA CAC AAC TCT GAA TTG GG), and RAG-1-tc1430F (TCA TCC AGC TGT TTG TTT GGC) and RAG-1-tc2700R (AAA GGT CCA TTA ATT CTC TGA GGG). PCR products were purified using AgenCourt AMPure XP beads (Beckman Coulter, Inc., Irving, TX, USA). We quantified the purified PCR products and later sequenced them in

both directions with the use of amplification primers and BigDye on an ABI 3730 capillary sequencer (Life Technologies, Grand Island, NY, USA).

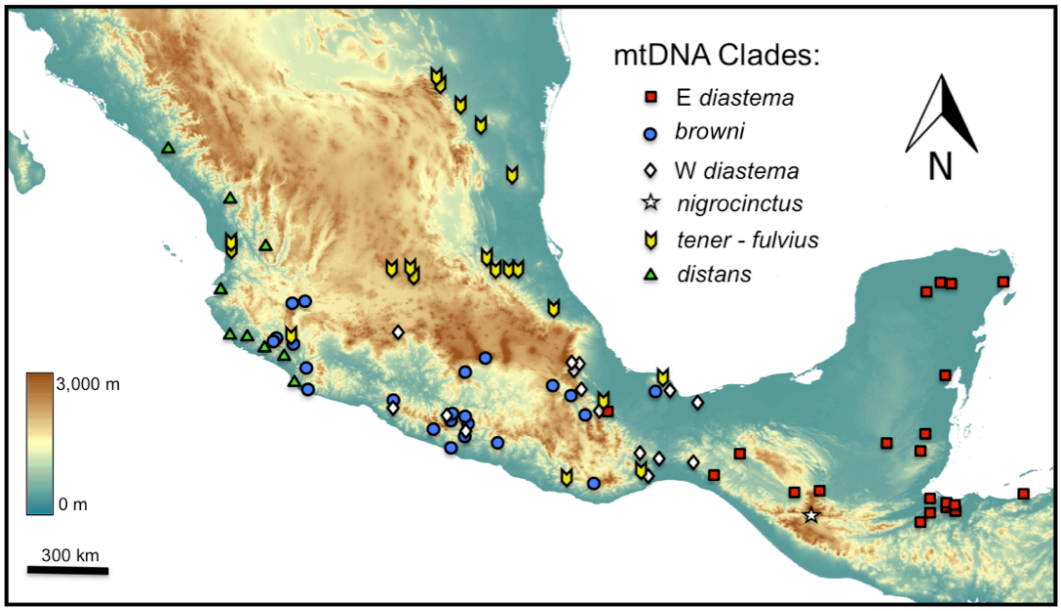


Figure 9 – Map showing localities of DNA samples used in this study.

ddRADseq library generation and sequencing

A subset of the DNA samples used for ND4 PCR were also used to generate double digest restriction associated DNA (ddRAD) libraries (n = 56; Supplementary Table 1) generally following the protocol of Peterson et al. (2012). We chose samples based on their placement on the mtDNA phylogeny in order to include representative samples of all putative species and as many mitochondrial clades as possible. We digested genomic DNA using a combination of rare and common cutting restriction enzymes: *SbfI* (8 bp recognition site) and Sau3AI (4 bp recognition site), respectively. We then ligated double-stranded indexed DNA adapters to the ends of digested fragments that also contained unique molecular identifiers (UMIs; eight consecutive N's upstream of the ligation site) using a mixture of digested DNA, adapters, T4 Ligase enzyme, and T4 Ligase Buffer (New England Biolabs, Ipswich, MA, USA). Ligations were performed on a thermalcycler at 16°C for one hour followed by a 65°C enzyme heat kill step for 10 mins. After adapter ligation, individual samples were pooled into pools of eight. We selected for a 440-540 bp

fragment size range using a Blue Pippin Prep (Sage Science, Beverly, MA, USA), and amplified size selected pools using PCR with primers including flow-cell binding sequences and an index specific to each sub-pool. A final sequencing library was constructed by re-pooling sub-pools in equimolar ratios based on molarity calculations from analysis on a Bioanalyzer DNA 7500 chip (Agilent, Santa Clara, CA, USA), and was sequenced using 100 bp paired-end reads on an Illumina HiSeq 2500.

mtDNA and RAG-1 sequence analysis

We edited the raw ND4 and RAG-1 gene sequence chromatograms using the program Geneious v6.1.6 (Biomatters Ltd., Auckland, NZ), and aligned the edited sequences with MUSCLE (Edgar 2004), with minimal manual adjustments to improve the alignment and to trim the 5' and 3' ends of all sequences in order to reduce columns with high levels of missing data. We decided to run all mitochondrial analyses with two different sets of data; one with all samples (127, including outgroups) but no complete coverage for all samples and a reduced dataset with less taxa (113, including outgroups) but with no missing data. The final RAG-1 alignment included nine samples for a total length of 2468 bases (See Supplementary Table 1 for reference numbers).

We estimated phylogenetic relationships among unique *Micrurus* haplotypes and outgroups using Bayesian phylogenetic inference in MrBayes v3.2.1. (Huelsenbeck and Ronquist 2001). For outgroups, we used a single representative of the Guerrero Longtailed rattlesnake, *Crotalus ericsmithi,* the Japanese coral snake, *Sinomicrurus japonicus,* the Sonoran coralsnake, *Micruroides euryxanthus* as well as 13 species of other *Micrurus* coralsnakes obtained from Genbank. We used the Bayesian Information Criterion (BIC) implemented in PartitionFinder v1.1.1 (Lanfear et al. 2012a) to select best-fit models: for ND4 we used K80 + Γ for 1st codon positions, F81 + Γ for 2nd codon positions, and GTR + Γ for 3rd codon positions. In the case of RAG-1, we used HKY for 1st and 2nd codon positions, and HKY + Invariant sites for 3rd codon positions. We used

these partitioned models for analyses of the individual genes in MrBayes, which consisted of four runs, each run for 10^7 generations with four chains (one cold and three heated), sampled every 1,000 generations. We confirmed that independent runs had converged based on overlap in likelihood and parameter estimates among runs, as well as effective sample size (ESS) and potential scale reduction factor value estimates (PSRF) values, which we evaluated in Tracer v1.5 (Drummond and Rambaut 2007). PSRF indicated that individual runs had converged by 10^5 generations, and thus we discarded the first 10^5 samples as burn-in. We generated a 50% majority rule consensus phylogram for each gene using combined estimates from post burn-in samples from the independent runs (Figs. 10 & 11). We also used the program Network v4.5.1.6 (Bandelt et al. 1999) to constructed a median-joining haplotype network to visualize relationships among unique haplotypes, with transitions weighted 2:1 over transversions (as recommended in the Network manual) and using the maximum parsimony option to reduce excess links among haplotypes from the resulting network.

Analysis of ddRADseq data

We processed the raw ddRADseq Illumina sequencing reads using the Stacks pipeline (Catchen et al. 2011; Catchen et al. 2013). Prior to running the pipeline, PCR clones were removed using the Stacks *clone_filter* program, which uses adapter UMIs to identify clones that are trimmed away using the FASTX-Toolkit trimmer (Hannon 2015). We then used the *process_radtags* function of Stacks to demultiplex samples by their unique barcodes, confirm the presence of restriction digest cut sites, and subsequently trim and discard reads with poor quality scores. Processed reads were aligned to the King Cobra (*Ophiophagus hannah*) genome (Vonk et al. 2013b) using the BWA mem algorithm (Li and Durbin 2009) with a mismatch penalty of 2, indel penalty of 3, and a minimum alignment score of 20. We used these mapping alignments in the reference-guided pipeline implemented within Stacks (Catchen et al. 2011; Catchen et al. 2013),

which includes Pstacks, Cstacks, and Sstacks, to generate and summarize SNP information for downstream analyses.

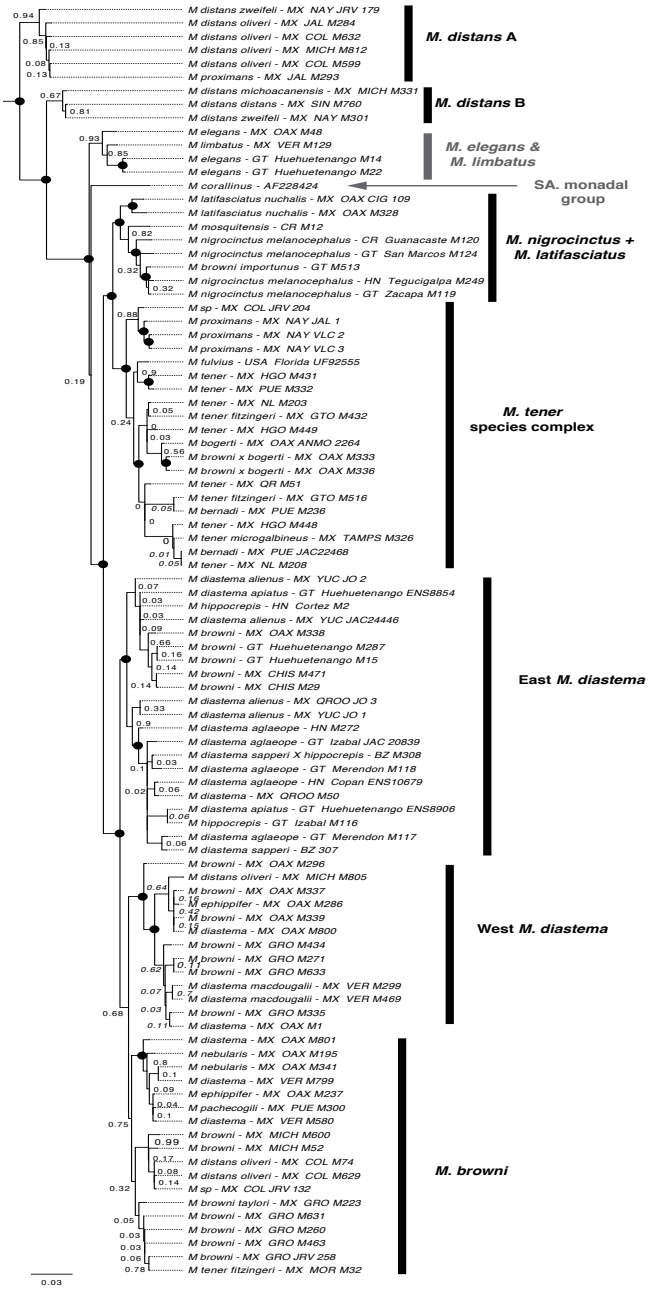


Figure 10 – Bayesian inference tree of the mitochondrial gene ND4. Samples are assigned to a taxon based on morphology and collector assignation. Black circles

represent nodes with PP >95%. Names on the right represent mtDNA clade assignations used in the rest of this study. Outgroups are not show.

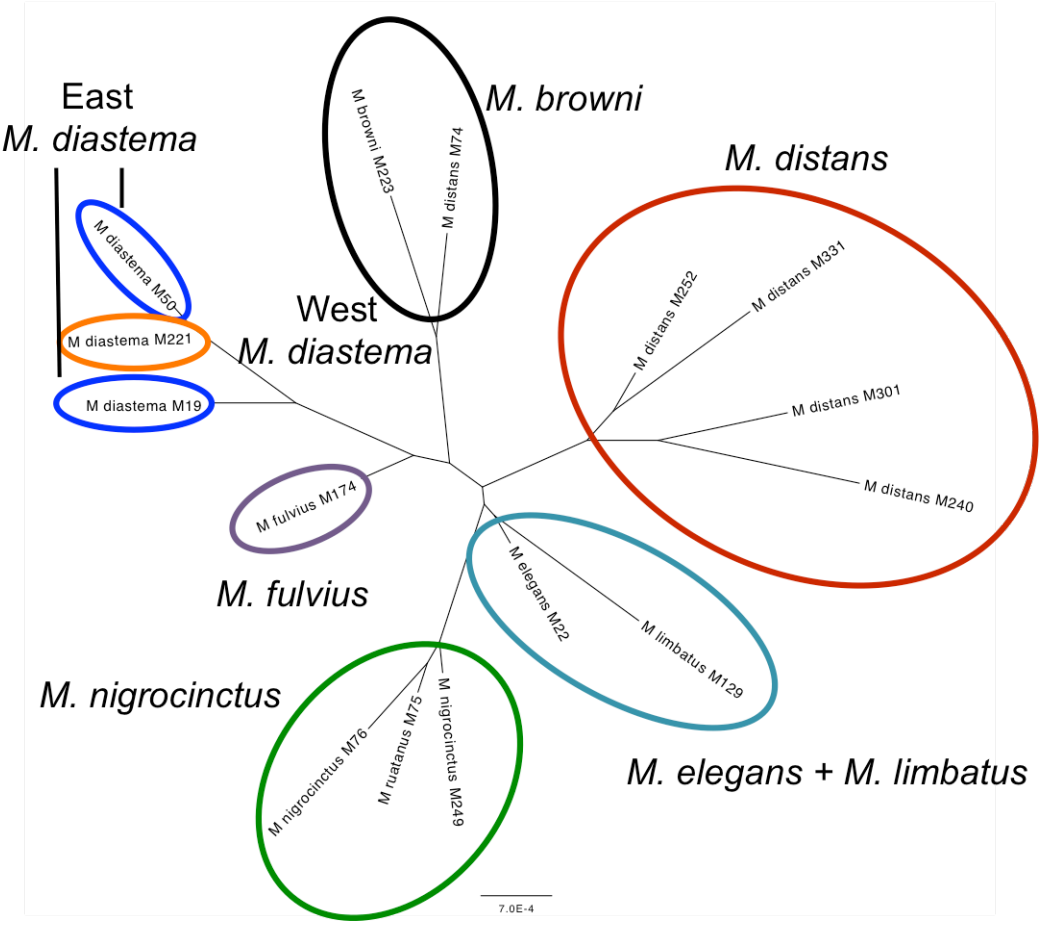


Figure 11 – Phylogenetic inference of the nuclear gene RAG-1. Names next to color circles represent the mtDNA clade were those samples belong to. From top right: red, *M. distans* clade; turquoise, *M. limbatus* + *M. elegans*; green, *M. nigrocinctus* clade; purple, *M. fulvius;* blue, east *M. diastema* clade*;* orange*,* west *M. diastema* clade; black, *M. browni* clade.

We used the *populations* program in Stacks to estimate a number of population genetic parameters and summarize genotypic information for further downstream analyses. For all analyses, we set *populations* thresholds for missing data (40%) and a minimum read depth per stack of 5x. *Populations* analyses of all 56 samples yielded no shared sites among all samples, thus we purged the RADseq data to include only samples with more than 100K reads (n=16).

We used the program *Structure* (Pritchard et al. 2000) to infer population structure and levels of admixture using the 20,416 SNPS obtained through our Stacks analyses. We estimated the allele frequency distribution parameter (λ) across a range of K values (1-4) under mixed ancestry and single population models (Fig. 12).

We applied principle component analysis (PCA) to identify the degree of genotypic clustering among all 16 individuals using the results generated from the *populations* analyses. Each individual was assigned a genotype for all loci (0 or 2 for homozygotes, 1 for heterozygotes), and PCA analysis was conducted in R using singular value decomposition for numerical accuracy.

We ran our RADseq data through the *pyRAD* v.3.0.1 pipeline (Eaton 2014) in order to obtain aligned loci to infer phylogenetic relationships among the *diastema* species complex. We choose the *pyRAD* pipeline over *Stacks* as *pyRAD* is specially designed to assemble data for phylogenomic studies that contain divergent taxa (Leaché et al. 2015). Sites with Phred quality scores <33 were changed into “N”s and if more than four sites per read had quality scores below this threshold they were discarded. We only used reads with coverage of >4 reads and a minimum sample of 50% of individuals for a final locus. We clustered the filtered reads in VSEARCH (Rogones) with a clustering threshold of 88% and aligned the final reads in *Muscle* (Edgar 2004). We estimated phylogenetic relationships for the contatenated RADseq data using a maximum likelihood approach implemented in RAxML (Stamatakis et al. 2008) with the GTRGAMMA model and 100 bootstraps (Fig. 13).

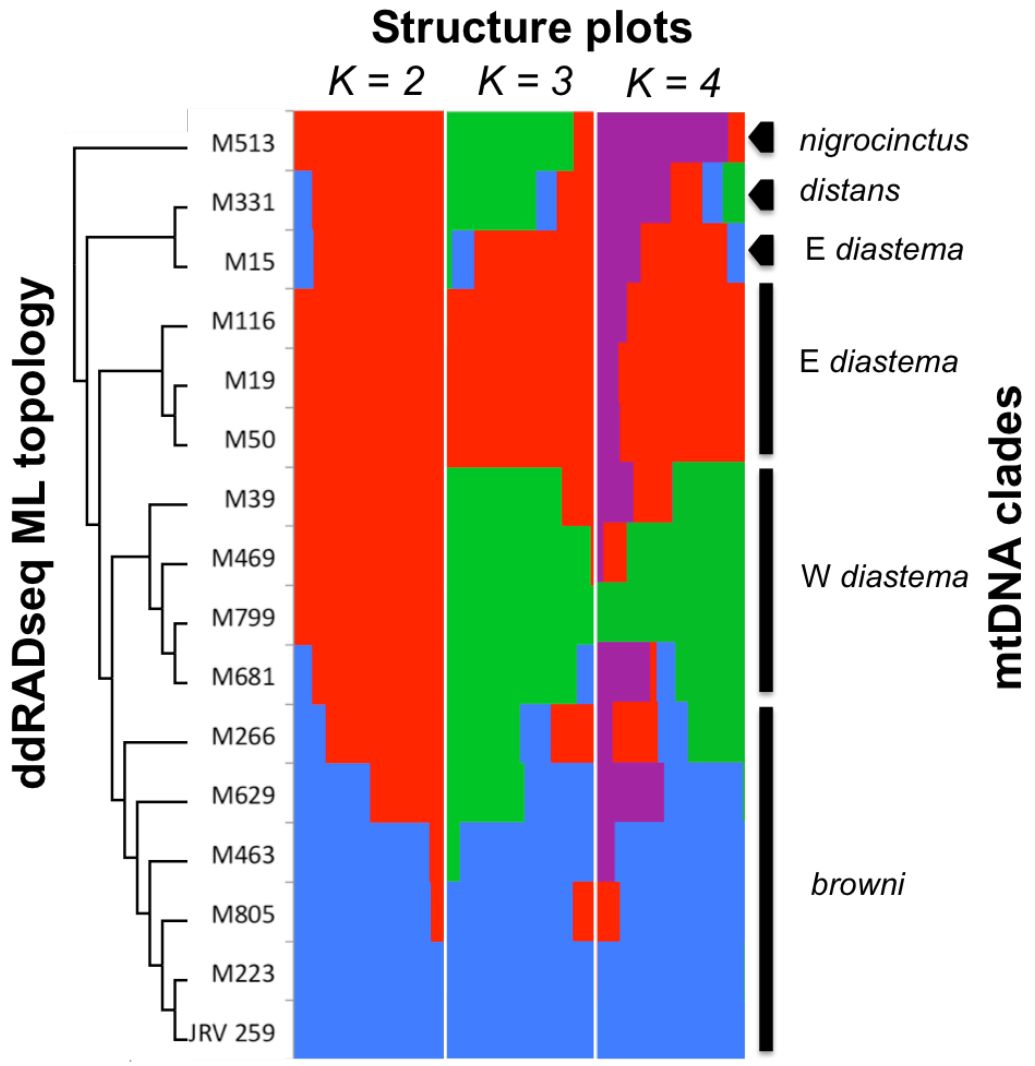


Figure 12 – Results of cluster analysis for *M. diastema* species complex samples from *Structure* estimated for different numbers of inferred populations (K = 2 – 4) based on 1,113 unlinked loci. The topology recovered in by ML phylogenetic inference of the ddRADseq data is shown at left, while mtDNA clades and clade names are show at right.

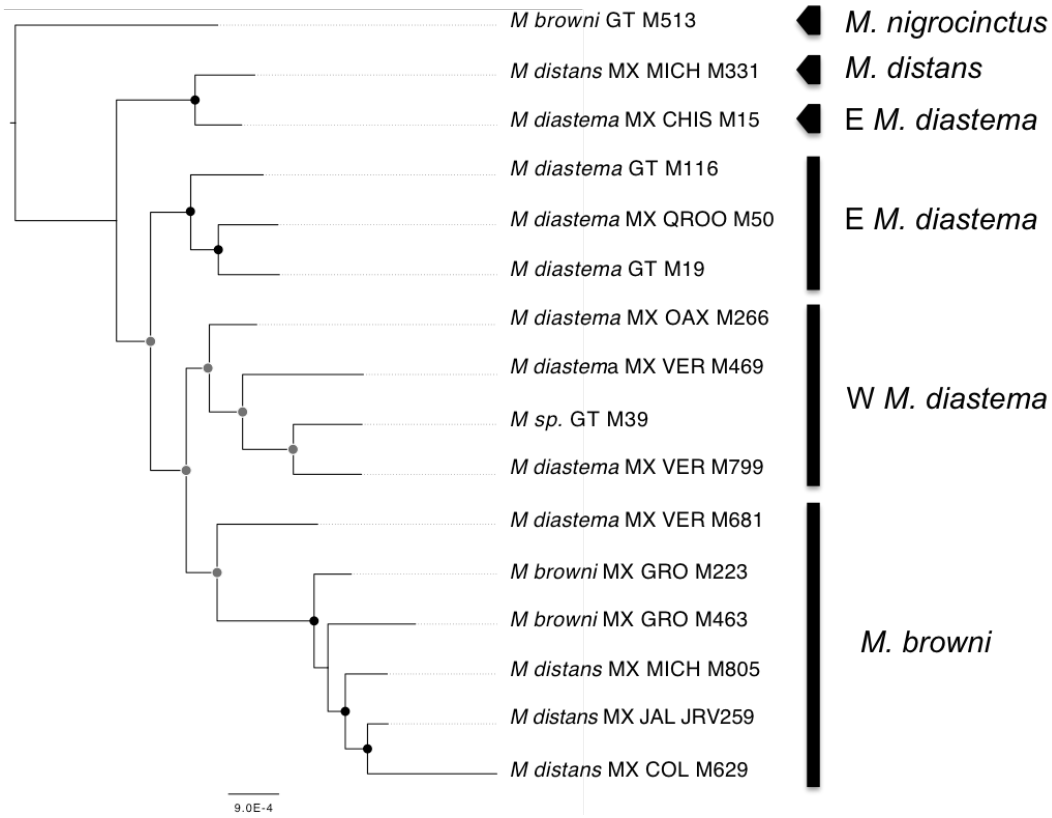


Figure 13 – Maximum likelihood Phylogenetic inference of the ddRADseq data. Names on the right represent mtDNA clade for each individual. Black circles represent a bootstrap support (BS) of >95, while gray circles are BS >50.

Non-molecular character analyses

In order to understand inter and intra-specific color variation in this group of coralsnakes we measured several coloration parameters in preserved specimens at the Amphibian and Reptile Diversity Research Center (ARDRC) at the University of Texas at Arlington. For each specimen we measured total length, tail length, number of black bands in body and tail, as well as size of the black yellow and red bands behind the head, at midbody and before the cloaca. We measured a total of 197 specimens that encompass all species of the *M. diastema* species complex. We gathered additional coloration data from specimens deposited in other museum collections in the US and Mexico. In this case, we only measured the body and tail length, as recorded the number of body bands in the body and tail. We obtained numerous additional records from the personal notes of Karl P. Schmidt and Janis A. Roze. Our final dataset included band counts and measurements for 491 specimens with species allocation and locality data (Fig. 14; Supplementary table 2). In order to assess for potential correlation between color pattern and climatic variables, we generated linear regression models in Rstudio (Racine 2012) for five bioclimatic variables (longitude, latitude, temperature, elevation, and precipitation) and banding patterns (Fig. 15). These bioclimatic variables were obtained from the WorldClim database (Hijmans et al. 2005).

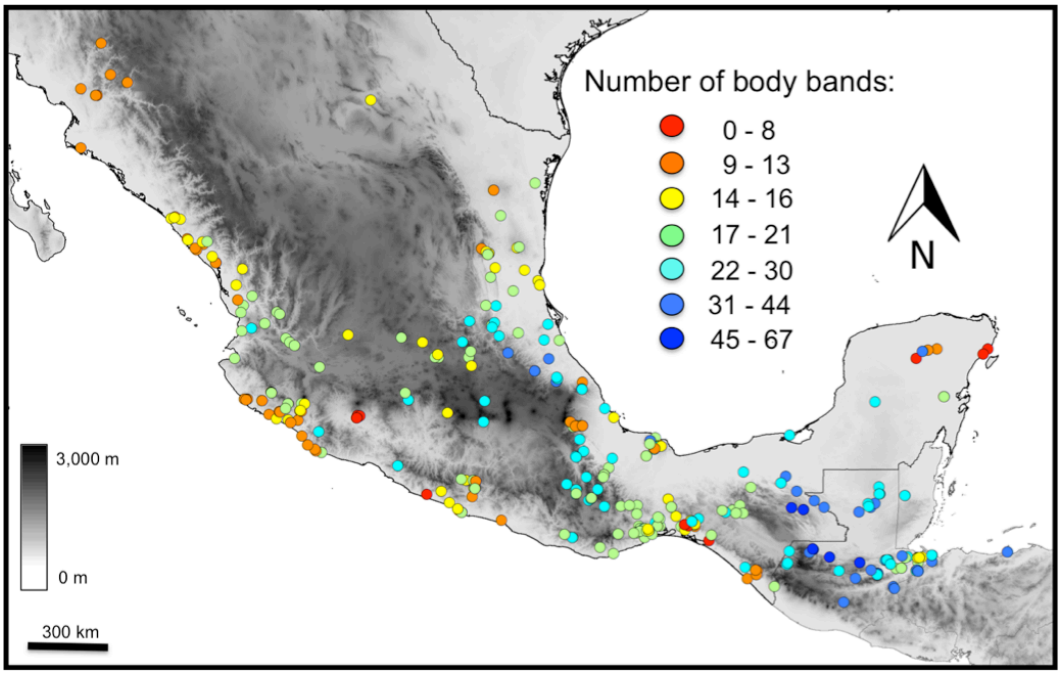


Figure 14 – Map showing localities of museum samples used in this study. Color circles represent number of body bands.

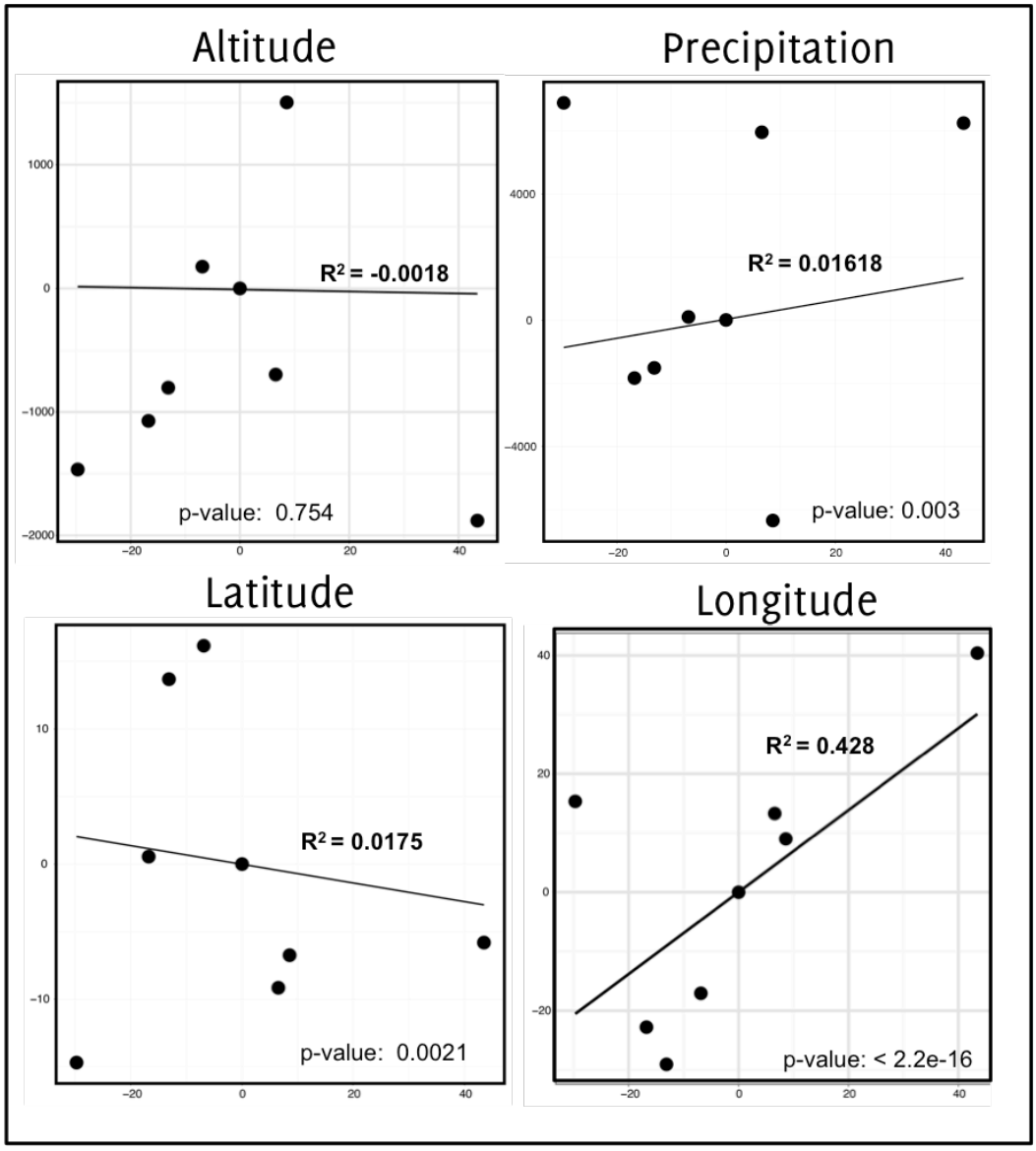


Figure 15 – Phylogenetically independent contrast analyses of body bands in members of the *M. diastema* species complex as a function of four geographic and climatic variables.

Results

mtDNA and nucDNA sequence characteristics and gene tree estimate

After cleanup, sequences of the mitochondrial gene ND4 measured between 196 and 709 base pairs. Our dataset that included all samples consisted of 629 aligned bases and ~15% missing data, while the dataset with no missing data was 269 bases long. In this dataset variable sites accounted for 30% of the sites (82 sites). Bayesian phylogenetic estimates for both datasets result in almost identical topologies. Figure 10 shows the topology obtained from the dataset with no missing data. Except for a few cases, the majority of the deeper nodes in our phylogeny showed high posterior probability (pp) support (>95%pp). We recover an early split of South American triadal coralsnakes (Fig. 1), sister to all other *Micrurus*. Not a single species of the *diastema* complex was recovered as monophyletic, with members of all species clustering into different clades (Figs. 10 &11). The second split in our phylogeny involved some members of *Micrurus distans,* which diverged early from the rest of coralsnakes . *Micrurus corallinus,* a member of the South and Central American monadal group was the next split in our phylogeny. We recovered the Middle American species *M. elegans* and *M. limbatus* to be sister taxa and the closest lineage to all remaining species. The rest of coralsnakes clustered into five main clades: Members of *M. nigrocinctus, M. mosquitensis, M. latifasciatus* and *M. browni importunus* formed a well supported clade; this group was sister to the *M. tener/fulvius* species complex, which included *M. fulvius, M. tener, M. bogerti, M. bernadi,* and *M. proximans*. Specimens of *Micrurus diastema diastema* formed a well-supported group along with individuals of *M. bogerti, M. browni* and *M. ephippifer.* The majority of *M. browni* formed a separate clade, which also included *M. pachecogili, M. nebularis* and several individuals of *M. ephippifer, M. bogerti, M. diastema,* and *M. distans.* The remaining clade consisted of all samples east of the Isthmus of Tehuantepec, including all the remaining *M. diastema* subspecies, along with

M. hippocrepis and samples of *M. browni* from Chiapas and Guatemala, to the exclussion of *M. browni importunus* (Figs. 10 &11).

Our RAG-1 tree contained only 15 samples, despite the low sampling, the six clades that where recovered are congruent with the mtDNA clades (Fig. 12), the only exception is that in the RAG-1 analysis, the three *M. diastema* individuals are recovered as monophyletic (vs. paraphyletic in the mtDNA topology).

ddRADseq results

A total of ~34 million raw reads were obtained across the 16 individuals used in the nuclear SNP analyses. Mapping reads to the King Cobra genome generated a total of ~20 million sequence alignments that were used in the Stacks pipeline to assemble 7,911 near loci and 20,416 SNPs under our RADseq filtering thresholds of 40% missing data and 5x read depth per locus.

Structure analyses – We estimated a $K = 2$ as the optimal model of population clustering using the Evanno method, implemented in Structure Harvester (Evanno et al. 2005; Earl and Vonholdt 2012). A comparison of population assignments of K from 2 to 4 are given in figure 12 for comparison. All these different models show some level of admixture between all species included the analyses, from low admixture in $K = 2$ to intermediate in $K = 4$. None of the currently recognized species in the *diastema* complex were completely distinguishable based on the structure plots alone, however, they seem to partially correspond to the clades obtained from the ddRADseq ML analysis.

Phylogenetic estimates – Our final alignment for the RADseq data was 450,251 base pairs long for a total of 16 individuals, with ~40% of missing data (Fig. 13). Despite the small sample size and large amount of missing data, the RADseq ML tree recovered a well-supported topology that largely congruent with the mtDNA phylogeny. We recovered an east *diastema* clade, a west *diastema* clade and a *browni* clade, all of which are consistent with the mtDNA topology. *Micrurus browni importunus* (M513) was

sister to all other samples in the analyses. The only topological discrepancy between the RADseq and mtDNA topologies is that an individual of *M. distans* from Michoacán, MX (M331) and a *M. browni* from Chiapas, MX (M15) are recovered as sister to one another, while in the mtDNA analysis they are distantly related.

Principal Component Analyses – We conducted two separate PCA analyses, one that included all samples of *M. diastema* (Fig 16; n =17), and another analysis that excluded an individual of *M. diastema* (M669) that was highly divergent from the rest (Fig. 17; n = 16). Principal component analyses indicate concordance with both of our population structure analyses and phylogenetic clustering (Figs 16 & 17), with the exception of two cases: M331 (*M. distans*) and M15 (*M. browni*), which are sister in the RADseq ML analysis do not cluster together in the PCA, while M463 (*M. browni*) and M469 (*M. diastema*) cluster near each other in the PCA, but are not closely related in any other analyses.

Non-molecular character analyses

Our analysis of non-molecular characters shows that there is a great amount of variation in color patterns across all species, and members of the same species did not cluster together in the PCA analysis of coloration (figure not show). There seems to be an important correspondence between geography and banding pattern, despite evolutionary relationships (Figs. 14 & 15). Some of the individuals with the highest number of body bands are found in areas with very high precipitation, for example, on the easter versant of the Sierra Madre Oriental and in the Atlantic highlands of Chiapas and Guatemala, while some of the individuals with the least number of bands are found in the dry areas of the northern Yucatan Peninsula, the Balsas Basin and the northwestern coast of Mexico. Phyllogenentically Independet Contrast identified significant relationships between in three of the four environmental and geographic variables (longitude, latitude, and

precipitation, all p-values < .001; Fig. 15). Multiple R-squared values indicate that these three variables explain a considerable proportion of the variation in coloration, even after corrected for phylogeny.

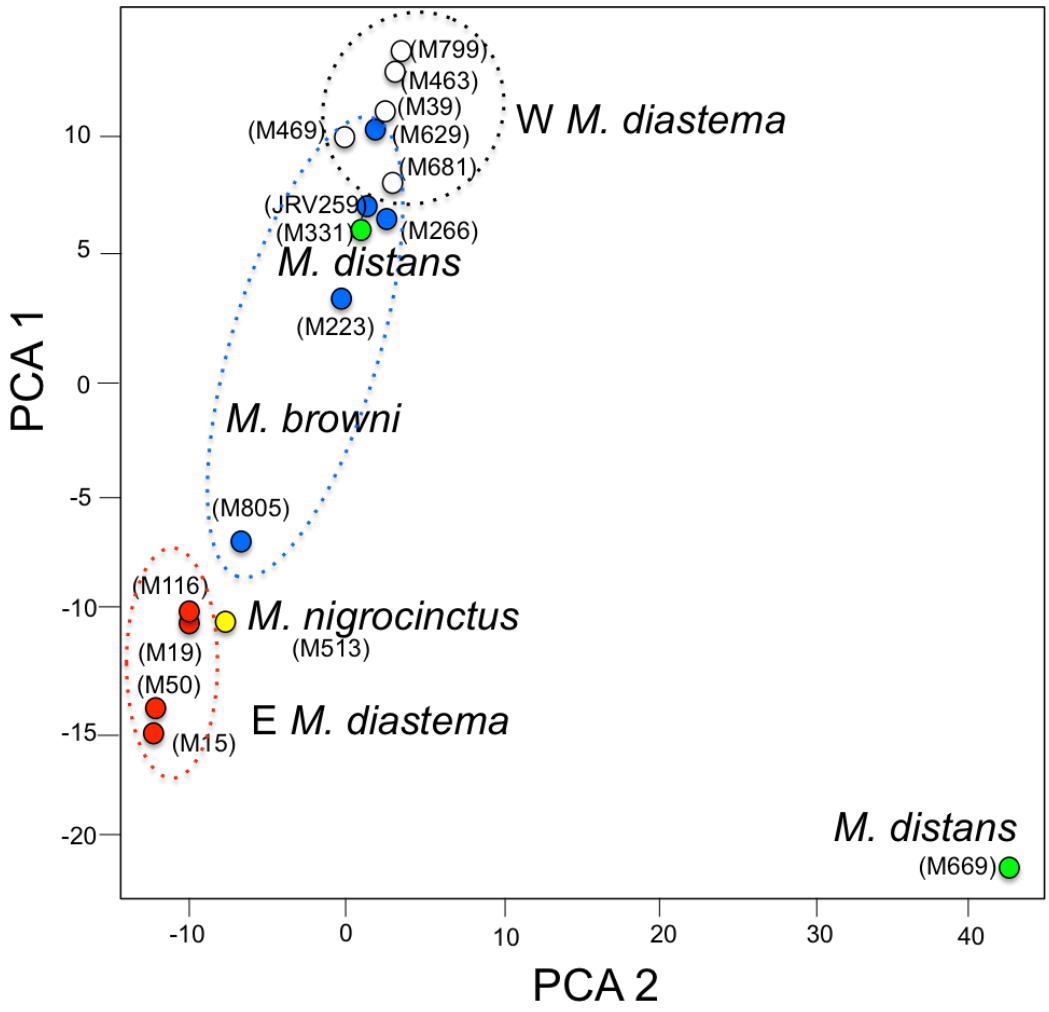


Figure 16 – PCA plot of 1115 loci shared across 17 individuals of the *M. diastema* species complex of *Micrurus*. Names next to dots represent reference number for each individual. Names next to dotted circles indicate mtDNA clade that the samples belong to. Circles are colored according to mtDNA clade

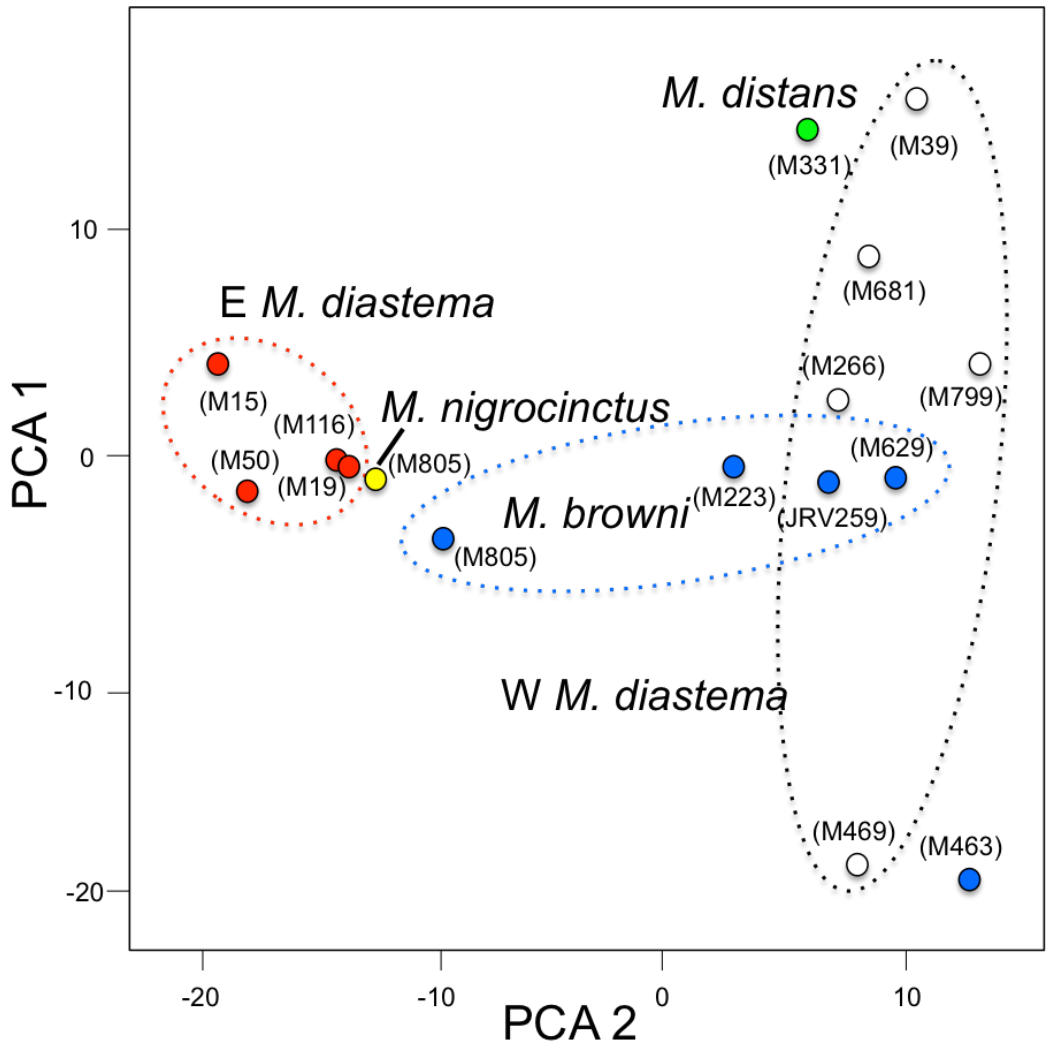


Figure 17 – PCA plot of 1115 loci shared across 16 individuals of the *M. diastema* species complex of *Micrurus,* excluding an individual of *M. distans (*M669) that was very divergent from the rest. Names next to dots represent reference number for each individual. Names next to dotted circles indicate mtDNA clade that the samples belong to. Circles are colored according to mtDNA clade.

Discussion

New World coralsnakes represent a highly diverse, widely distributed and highly venomous group of snakes that also represent an important model for studying mimicry systems. Despite these characteristics, there is remarkably little known about their phylogenetic relationships and molecular data has only been applied to study limited sets of species or higher-level relationships among lineages (Slowinski and Keogh 2000; Castoe et al. 2007a). *Micrurus* is the most speciose genus of New World coralsnakes, with 84 described species distributed from the southern USA to southern South America (McDiarmid et al. 1999; Campbell and Lamar 2004b; Uetz and Jirí 2015). High morphological conservatism within this genus has led to a taxonomy that relies heavily on diagnostic color patterns, although it is well known that these color patterns are central to various types of mimicry systems (Greene and McDiarmid 1981; Brodie III 1993; Pfennig et al. 2001), which may lead to convergence in color patterns among lineages.

Our phylogenetic inferences from mitochondrial and nuclear gene sequences, as well as nuclear SNPs, are largely in agreement regarding relationships among *diastema* group lineages. All datasets inferred that a group of *M. distans* from northeastern Mexico are distantly related to all other Central American monadal coralsnakes (Figs. 9-11), and PCA analysis of nuclear SNPs (Figs 16 &17) indicates that such *M. distans* samples are highly divergent from all other *diastema* group samples. Mitochondrial gene-based inferences identified two ladderized clades of *M. distans* (labeled as “distans A” and “distans B”; Fig. 10) that diverged early from remaining monadal coralsnakes. Nuclear gene sequence and SNP datasets, however, were not able to test the hypothesis of there being two clades versus a single clade because of insufficient taxon sampling (all nuclear sampling included individuals from the “distans B” clade). Regardless, all datasets, including all nuclear data, strongly agree that *M. distans* from northeastern Mexico are

very distantly related to other members of the *diastema* species complex and thus should no longer be included within the *diastema* species group.

All of our phylogenetic analyses agree in identifying multiple distinct phylogenetic lineages within the *diastema* species complex. Phylogenies inferred from the nuclear gene RAG-1 indicated major clades: 1) a *M. diastema* clade, and 2) a *M. browni* clade (Fig. 11). Mitochondrial gene analyses indicated three main clades, which differ from the nuclear gene inference by splitting up the *M. diastema* clade into an East *diastema* and a West *diastema* clade (Fig. 2). The phylogeny based on nuclear SNPs agrees well with both of these inferences, and supports each of the three mitochondrial clades. The SNP-based phylogeny also shows evidence for a fourth, rather enigmatic and unexpected clade that includes samples of *M. diastema* and *M. distans* from widely separated localitites of the Pacific coast of Mexico (Chiapas and Michocán, respectively). Cluster-based analysis of SNPs across the *diastema* group using Structure (Fig. 12) shows these two individuals as being comprised of substantially different allelic content. Similarly, PCA analysis results of nuclear SNPs also place these individuals very far apart from each other (Figs. 16 &17), collectively raising substantial doubt about the findings in the SNP-based ML phylogenetic tree (Fig. 13). In conclusion, we interpret the results as showing evidence for three major lineages within the *M. diastema* species complex.

Our cluster analyses of SNP data suggest the existence of a considerable degree of gene flow between all species studied, and these patterns of apparent introgression are consistent with the geographical proximity of lineages and populations (Fig. 12). Members of *M. browni* from Colima, Jalisco, Michoacán and Guerrero cluster together and share a small number of loci with other species. The degree of loci that are shared is partially concordant with a northwest-southeast pattern, as the southeastern most individual of this clade (M266, from Oaxaca) shares the most loci with members of *M. diastema* from south-central Mexico. Many loci are shared between the eastern and western mtDNA clades of *diastema,* however, at higher *K* values, a more clear patter

shows population differentiation between these two clades, which is also in agreement with geographic location.

Considering that the current taxonomy of New World coralsnakes is largely based on coloration and color patterns, our finding of extensive disagreement between the current taxonomy and the phylogeny indicates that color pattern is not a reliable character for systematics of coralsnakes. In particular, there are reasons to expect that color pattern in coralsnakes may be particularly plastic, and under complex patterns of selection due to its central role in complex Batesian and Muellerian mimicry systems (Greene and McDiarmid 1981; Brodie III 1993; Pfennig et al. 2001). Thus, color pattern characters, such as the numbers of bands per snake, may instead more strongly covary with features that correlate with predation pressures, rather than phylogeny. To explore this further we conducted analyses of the numbers of bands in relation to geography (Fig. 14) and to particular bioclimatic variables (Fig. 15). These analyses show longitude and precipitation covarying most tightly with band number, with individuals inhabiting more eastern and more humid environments having a higher number of body bands (Figs. 14 &15).

It is currently unknown if other environmental conditions that might be correlated with precipitation (for example, the diversity of snake predators in a given area) might be responsible for the variation in color patterns. One of the most significant examples of color pattern vs. habitat differences occurs in members of E *M. diastema.* Despite the low genetic diversity between populations, members of this clade vary from no bands or very few bands in the dry Yucatan Peninsula, to more than 60 bands in the humid highlands of Chiapas and Guatemala (Fraser 1973). There is also considerable variation between individuals at the same locality, however, at any given place the majority of individuals do not deviate more than a few bands from one another, perhaps indicating strong selection for that particular color pattern (see supplementary table 2).

Taxonomic implications

Our molecular phylogenetic inferences show an extensive degree of discordance with taxonomy, indicating that the current taxonomy is inadequate for these highly venomous coralsnake species. We estimate that the *M. diastema* species complex is not monophyletic, and is comprised of several lineages, some of which are only distantly related to *M. diastema.* Based on our mtDNA and nuclear SNPs datasets, we believe that seven species of coralsnakes in the *diastema* species complex are recognizable in Mexico and Central America. *Micrurus distans* appears be composed of two distinct taxa (*distans* A and *distans* B in out mtDNA phylogeny, fig. 9), Unfortunately it not possible with the data at hand to adequately diagnose these two taxa because they are not concordant with subspecific designations and their ranges appear to overlap in western Mexico. Several species that occur in northeastern and southern Mexico, including *M. bernadi*, *M. tamaulipensis,* and *M. bogerti* are conspecific with *M. tener.* At the same time, none of the currently recognized subspecies of *M. tener* form monophyletic groups, so they should be synonymized with *M. tener*. Despite being closely related to *M. tener*, *M. proximans* from Jalisco, Colima and Nayarit form a monophyletic clade that is the sister lineage to all other *tener-fulvius* group lineages, and thus seems to warrant unique specific recognition. In the case of *M. diastema*, there is strong evidence to suggest that this species is not monophyletic, and is instead composed of two taxa that are not very closely related to one another. Individuals of *M. diastema* that occur mostly east of the isthmus of Tehuantepec form a monophyletic group, which is composed of the subspecies *M. diastema alienus, M.d. sapperi, M. d. apiatus, M. aglaeope* and some individuals of *M. d. affinis* that occur west of the isthmus. *Micrurus hippocrepis* and *M. browni* from Chiapas and Guatemala also belong to this group. The name *M. diastema apiatus* Jan (1858) has priority for this group, so we suggest the use of the new combination *M. apiatu*s Jan. Individuals of *Micrurus diastema* that occur west of the isthmus of Tehuantepec represent a distinct taxon which includes the subspecies *M. diastema diastema, M. d. macdougalli, M. d. affinis,* as well as some members of *M.*

ephippifer from Oaxaca and *M. browni* that range as far west as Guerrero (Figs. 9 & 10). The name *M. diastema* Duméril, Bibron & Duméril (1854) is applicable to this group. We found that the endemic *Micrurus nebularis* and *M. pachecogili* are nested within *M. browni*, as well as members of *M. ephippifer* from the highlands of Oaxaca and some *M. distans oliveri* from Jalisco, Colima and Michoacán, while the subspecies *M. b. importunus* is a junior synonym of *M. nigrocinctus*.

Our study points out the difficulty in delimiting species in this group of morphologically conserved and chromatically diverse group of snakes. Based on our results, we would expect that other coralsnake lineages are equally chaotic in terms of mismatches between phylogeny and taxonomy. Previous studies have shown that venom content, antivenom efficacy, and envenomation symptoms are all tightly correlated with phylogeny (de Roodt et al. 2004). Thus, in addition to clarifying the taxonomy of these intriguing and brightly colored model species for studying mimicry systems, an improved understanding of coralsnake phylogeny and systematics will also be important for treatment of coralsnake envenomation and for understanding patterns of variation in coralsnake venom.

Chapter 3

Using the burmese python genome to understand the evolution of snake venom systems

Introduction

Snake venoms and their evolutionary origins have received substantial attention over the past several decades (Vidal 2002; Fry et al. 2006; Casewell et al. 2014), including the evolutionary processes that have led to the toxic effects of these proteins (Casewell et al. 2013). A dominant hypothesis for the evolutionary origins of most venom toxin families involves the duplication of non-toxic genes, with subsequent neofunctionalization of gene copies to adaptively modify the structure and function of these proteins (Ivanov and Ivanov 1979; Ivanov 1981; Fujimi et al. 2003; Fry 2005; Fry et al. 2006; Tamiya and Fujimi 2006; Fry et al. 2009; Kini and Chinnasamy 2010; Casewell et al. 2012). Recent genome-scale resolution of this phenomenon has confirmed many of these assertions, indicating that in some cases the process of toxin gene duplication can result in expansive multi-locus venom gene families, as observed in the king cobra genome (Vonk et al. 2013a). Such duplication, neofunctionalization and recruitment events appear to have occurred multiple times throughout the evolution of snakes, including multiple parallel expansion events of particular gene families in different snake lineages (Casewell et al. 2012).

There are more than twenty gene families that are traditionally considered to be “venom toxins” in squamate reptiles due primarily to their detection in venom gland secretions, and in some species, evidence for the toxicity of some of these venom components (Mackessy 2002; Mackessy et al. 2006; Mackessy 2010b). The detection of expression of genes related to these “venom toxins” in venom glands or other oral glands in squamate reptiles has further become an accepted proxy for labeling such genes as “venom toxins” and the labeling of such species as “venomous” (Fry et al. 2009; Fry et al. 2010; Fry et al. 2013). Several studies, however, have shown evidence that venom

genes or their homologs are expressed in tissues other than the venom or accessory venom gland of snakes and other venomous vertebrates (Rádis-Baptista et al. 2003; Whittington et al. 2008; Hargreaves et al. 2014), which calls this practice into question. Despite these inferences, there have been no comprehensive expression analyses of such “venom toxin” gene families across a broad diversity of snake organs and tissues. Thus, the degree to which venom genes or venom homolog expression in oral glands may be either a physiological default or an adaptive feature indicative of their functional role in oral secretions remains an important yet insufficiently studied question.

While most previous studies have focused on either gene duplication or patterns of molecular evolution of snake venoms (Fry 2005; Fox and Serrano 2008; Casewell et al. 2013), no previous studies have focused on the role that gene expression might play specifically in this venom gene recruitment process. The genes that have been targeted for recruitment into venoms appear to share certain common attributes, which support the hypothesis that successful recruitment may be linked to functional constraints of the recruited proteins (Alape-Girón et al. 1999; Fry et al. 2009). Successful recruitment of genes as venom toxins hypothetically requires a transition in which nascent venom proteins must be targeted for gene expression in specific tissues (i.e., the venom glands). Therefore understanding the evolution of expression of such genes is an essential but largely absent component for understanding their functionality, origins, and the constraints that have shaped venom repertoires. Gene expression in the venom glands of snakes has been evaluated in a number of studies (Junqueira-de-Azevedo and Ho 2002; Pahari et al. 2007; Doley et al. 2008; Fry et al. 2013; Margres et al. 2013), but due to the relative scarcity of comparative expression data for other snake tissues, venom gland gene expression is rarely viewed in the broader context of expression across diverse tissues (e.g. (Hargreaves et al. 2014)). It therefore remains unknown whether certain protein expression characteristics might favor their recruitment as venom toxins, or if their expression profiles are not a relevant factor influencing recruitment.

There is uncertainty and debate over the origins of venom systems in squamate reptiles, with a common view being that a core venom system evolved a single time in the common ancestor of snakes and a clade of lizards, referred to collectively as the Toxicofera (Fry et al. 2006). This hypothesis remains controversial largely due to disagreement about what, indeed, constitutes a “venom toxin” (Terrat and Ducancel 2013) as well as a lack of apparent venom homolog expression and function in multiple large clades of Toxicoferan lizards (Fry et al. 2010; Fry et al. 2013). A functional definition for venom would be that it is a specialized glandular secretion which causes deleterious effects to a recipient organism when injected; this secretion is typically protein-rich and may consist of many different molecules or toxins, often representing a specialized trophic adaptation which facilitates prey handling (Mackessy 2002). However, there is continued debate of details of this definition (Nelsen et al. 2014). Current evidence indicates that a massive radiation of snakes with highly toxic venoms probably evolved after the divergence between the python and caenophidian snakes, which include, elapids, colubrids, lamprophiids and viperids (Vidal 2002; Fry and Wüster 2004). Accordingly, recent genomic evidence from the king cobra demonstrates that many toxic venom gene families have experienced substantial duplication and divergence in the cobra relative to the python (Vonk et al. 2013a). Collectively, these data indicate that the Burmese python (Python molurus bivittatus) may provide a system in which to estimate patterns of gene expression prior to the expansion of highly toxic venom genes in caenophidian snakes, particularly in the highly venomous colubroid snakes (Fig. 18). The genome and genomic resources of the non-venomous Burmese python (Castoe et al. 2013), thereby offer a unique opportunity to study patterns of expression for genes recruited into the snake venom system within the context of a complete set of snake genes and a large set of gene expression data from diverse python tissues and organs.

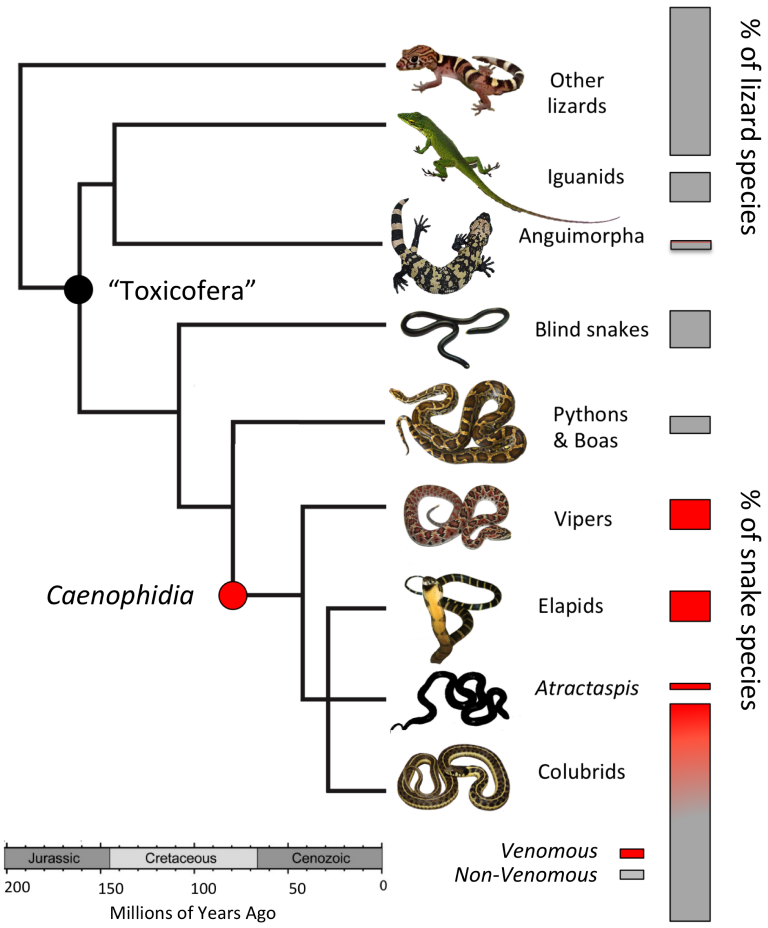


Figure 18 Phylogenetic tree showing lizard and snake relationships and the distribution of venomous species. The black circle refers to the "Toxicofera", which includes all snakes and some lizards, and the red circle represents the Caenophidia, which contains all known deadly venomous snakes. The percentage of venomous colubrid snakes is an approximation.

In this study we use the python genome and tissue-specific expression data to investigate the origins of venom genes in highly venomous caenophidian snakes and to assess the validity of defining genes as 'venom toxins' based solely on evidence of gene expression detected in the oral glands of squamates. As a first step toward addressing these goals, we conducted thorough analyses to identify the relationships between python genes and known venom genes from caenophidian snakes and other squamate reptiles, and we provide new evidence for the orthology and patterns of gene expansion in snake venom gene families. We used these estimates of gene orthology together with python gene expression data to address two related questions: 1) Are there inherent characteristics of gene expression for venom gene homologs that may have predisposed them for recruitment as venoms? 2) Are venom gene homologs uniquely expressed or particularly abundant in python oral glands, such as the rictal gland?

Materials and methods

BLAST analyses to identify python gene homologs of known venom genes.

We studied a total of 24 venom gene families (Mackessy 2002, 2010b, a), which we obtained examples of from GenBank (Tables 1 and S2). These 24 venom gene families represent the vast majority of known squamate venoms, and the only ones with available DNA sequences. To identify homologous genes in other lineages, we blasted each venom gene to the complete protein coding sequences (CDSs) of the human, anole lizard, Burmese python and king cobra using tblastx. CDS files were obtained from Ensembl (Flicek et al. 2014) and from recently published snake genomes (Castoe et al. 2013; Vonk et al. 2013a). From each blast search, we retained the top three hits for each taxon based on its E-value (E-value < 1e-05), and the top three hits based on bit scores (bit score > 70). If neither criterion was met, we retained the highest E-value hit and the gene with the highest bit score for each queried species. To increase phylogenetic

resolution, we included additional sequences from several other vertebrate species from GenBank, and sampling used previously (Vonk et al. 2013a).

Phylogenetic analysis to identify gene homologs.

We conducted first-pass alignments of translated amino acid sequences using Muscle (Edgar 2004). Once aligned, sequences were converted to nucleotides, and nucleotide-level alignments were used for all subsequent analyses. We estimated best-fit models of nucleotide evolution using PartitionFinder (Lanfear et al. 2012b). We inferred phylogenies in MrBayes version 3.2.1 (Ronquist et al. 2012). For each gene we ran two simultaneous analyses of 10^7 generations, and sampled the chain every 10^3 generations. We confirmed mixing and convergence using Tracer V.1.5 (Rambaut and Drummond 2007), and discarded the first 10% of all runs as burn-in. After first-pass analysis, we identified non-homologous sequences as those with extremely long branches and very low posterior support (<50%), and these sequences were removed from alignments, alignments were re-optimized, and we estimated new phylogenetic trees based on these revised alignments.

Analysis of gene expression data from the python.

We used all gene expression data available for the Burmese python (Castoe et al. 2013). Where available, expression data from multiple individuals were combined per tissue for all analyses. We normalized read counts using TMM normalization in edgeR (Robinson et al. 2010), and converted read counts to counts per million (CPM). We used our phylogenetic estimates for each of the 24 venom gene families to identify venom gene homologs in the python (Table 1), and we use the term homolog to refer to multiple situations, including evidence of orthology (including 1:1 orthology) and other instances where our best estimate is based on a blast-based hit. We categorized patterns of gene expression in several ways and compared these patterns between python venom

homologs and the complete python gene set. We assigned all python genes to one of seven different log-scale categories based on their normalized expression levels in a given tissue: (1) CPM = 0; (2) CPM = 0-1; (3) CPM = 1-10; (4) CPM = 10-100; (5) CPM = 100-1,000; (6) CPM = 1,000-10,000; and (7) CPM = >10,000. We compared the pattern of expression levels between venom gene homologs and all other python genes using a Fisher's exact test. For each gene we also calculated the mean and variance in expression level across all tissues combined and tested for differences between venom homologs and all genes using Fisher's exact tests (Table S5). Because it is unclear what level of gene expression might be biologically relevant, we used multiple thresholds of CPM read counts for "presence" of a gene in a given tissue: (1) CPM >1; (2) CPM >10; (3) CPM >100; (4) CPM >1,000; and (5) CPM > 10,000 (Fig. 11). Significant differences between venom homologs and all other genes were tested using Fisher's exact tests.

Results

Estimates of python gene homology to known venom genes.

We were able to confidently identify the homologous gene (or genes) in the python for 20 out of the 24 venom gene families analyzed (Table 1 and S1). We identified a single orthologous gene in the python for 15 of the venom gene families, while two homologs were found for cystatin, metalloproteinase, phospholipase A_2 (PLA_2), serine proteinase and veficolin. In the case of PLA_2, however, we found two separate clades of venom genes, each with a single ortholog in the python. Our analyses resulted in the identification of a total of 25 homologs for 20 gene families (Figs. S1-S20). Phylogenetic inferences of orthology of python venom homologs in relation to known venom genes were strongly supported for 19 gene families (>95% posterior probability; Figs. S1-S20). Only the python orthologs for exendin had posterior support below this threshold, with 92% posterior support. In bradykinin potentiating peptide/natriuretic peptide (BPP) and sarafotoxin, orthologous sequences could not be confidently identified by phylogenetic

analyses; these genes appear to have many domain insertions and deletions yielding poor alignments and it is known that sarafotoxin presents a unique structure which is very distinct from its putative ancestral endothelin protein (Takasaki et al. 1992; Ducancel et al. 1993). The other genes for which a python homolog could not be inferred with confidence from phylogenetic analyses were crotamine and waprin. Several studies have found homologous sequences for these genes in non-venomous reptiles with either low posterior support or when no reptilian outgroups were included, which we believe may result in a biased inference of gene relationships (Fry 2005; Fry et al. 2006; Vonk et al. 2013a). Given an absence of quality alignments for these four gene families, we instead used protein similarity (based on the best tblastx hit) to estimate the probable homolog in the python for subsequent analyses. In total, further analyses therefore included 29 gene homologs for 24 gene families. Due to the controversy surrounding resolution of what qualities define a protein as a venom toxin, we also repeated all analyses including only venom protein families known to have well-defined toxic and/or cytotoxic properties (Table S2). In this case only four gene families were included: 3FTs, metalloproteinase, serine proteinase and PLA2.

Table 1 - Venom gene families used in this study and the number of orthologs estimated in the python and other snake genomes. Orthologs in the python with an asterisk represent venom genes where homology could not be inferred by gene trees. Gene numbers are based on the following citations: python (this study), cobra (Vonk et al. 2013a), vipers (Casewell et al. 2009; Casewell et al. 2014), rattlesnake (Pahari et al. 2007). Gene numbers for the Cobra are based on the complete genome sequence; estimates for vipers and the rattlesnake are based on venom gland transcriptome data and may represent a lower bound.

		Caenophidian Snakes		
Venom Gene Family	**Python**	**Cobra**	**Vipers**	**Rattlesnake**
3 Finger Toxin	1	25	-	5
5' Nucleotidase	1	1	1	-
Acetylcholinesterase	1	0	-	-
AVIToxin	1	1	-	-
BPP	1*	1	>1	1
C-type lectin	1	11	>6	3
Cobra Venom Factor	1	5	-	-
CRiSP	1	6	1	2
Crotamine/Crotasine	1*	-	-	-
Cystatin	2	2	2	-
Exendin	1	1	-	-
Exonuclease	1	0	-	-
Hyaluronidase	1	0	1	-
LAAO	1	1	1	1
Metalloproteinase	2	8	>11	6
Nerve Growth Factor	1	1	1	1
Phosphodiesterase	1	0	1	1
PLA_2 I (Viperids)	1	1	>2	1
PLA_2 II (Elapids)	1	8	-	-
Sarafotoxin	1*	-	-	-
Serine Proteinase	2	5	>3	12
Veficolin	2	2	-	-
VEGF	1	1	1	2
Vespryn	1	1	-	-
Waprin	1*	-	-	1

Comparison of expression profiles of python venom homologs across tissues.

Twenty of the 29 venom homologs identified in the python show at least some level of gene expression in the python rictal gland (Fig. 19A-B). Four venom homologs (3FTs, C-type lectin, veficolin I, and vespryn) show their highest levels of expression in the rictal gland. Of these, C-type lectin is expressed at levels that are orders of magnitude higher in the rictal gland than in any other tissues surveyed (1,000-10,000 CPM), while 3FTs, vespryn and veficolin orthologs are expressed at intermediate to high levels (100-1,000 CPM). All of the venom homologs that show expression in the rictal gland, however, show some level of expression in other python tissues. Two venom homologs, 5' nucleotidase and cobra venom factor, show very high levels of expression in the liver (1,000-10,000 CPM) and phosphodiesterase is found expressed at similar levels in the small intestine. Five venom homologs are expressed at intermediate to high levels across all of the sampled tissues (Fig. 19): 5' nucleotidase, exonuclease, metalloproteinase A, phosphodiesterase and PLA_2 I. Eighteen of the 29 homologs are expressed in at least half of all tissues samples, but only four of them are expressed at medium to high levels (100-1,000 CPM) in most tissues (Fig. 19). In contrast, 10 python venom orthologs are expressed in only seven tissues or less and at low levels (<100 CPM). Thus, although the majority of venom homologs are expressed in the rictal gland, other tissues demonstrate similar or higher levels of expression of these same genes, and the brain, small intestine and kidney had more venom homologs being expressed than the rictal gland (Fig. 19).

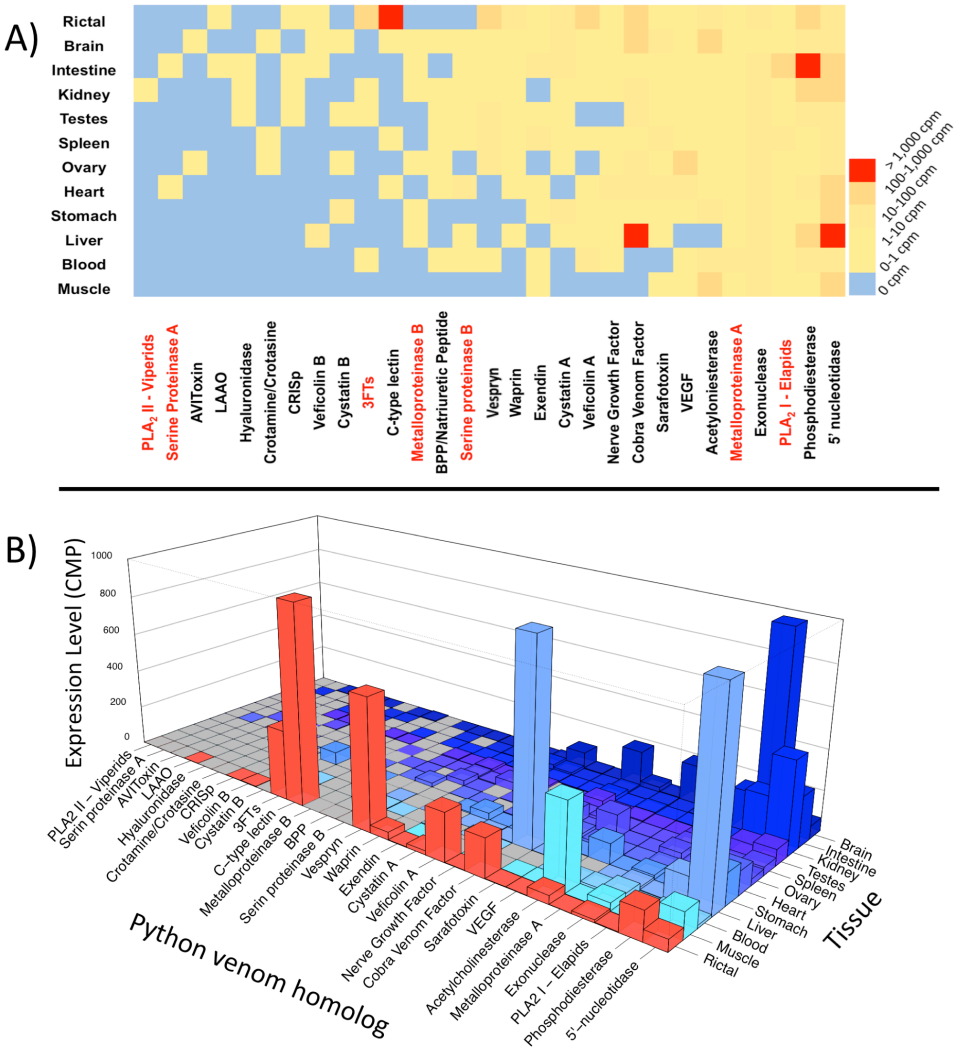


Figure 19 - Expression profiles for python venom gene homologs across tissues. A) Heatmap of gene expression profiles shown as counts per million (CPM) on a $\log_{10}$-scale. Names of genes with known toxicity are in red. B) Python venom gene homolog expression with expression levels are shown in CPM. Note that the Y-axis (expression level) is truncated to 1,000 CPM. Abbreviations include: 3FTs = 3-finger toxins; BPP = Bradykinin potentiating peptide/natriuretic peptide; CRISp = Cysteine rich secretory protein; CVF = Cobra Venom Factor; LAAO = L-amino acid oxidase; NGF = Nerve Growth Factor; VEGF = Vascular endothelial growth factor.

Statistical enrichment analysis of python venom gene homolog expression.

Comparison of expression patterns between all other python genes versus python venom homologs indicates that python venom homolog expression is statistically different from the patterns observed for all other genes. Venom homologs tend to be expressed at lower levels (0-1 CPM) more frequently than expected, and are less commonly expressed at intermediate levels (10-100 CPM; Fig. 20A). Very similar patterns of deviation from the complete set of genes are observed when only genes with known cytotoxic activity are compared to all other python genes (Fig. 20B).

To address the question of whether venom homologs tend to be expressed in more or in fewer tissues compared to all python genes, we used multiple expression levels as cutoff values for “presence” in a tissue because it is unclear what level of expression might be physiologically relevant. At the lowest threshold for presence (>1 CPM), venom homologs were enriched for higher frequencies of presence in a single tissue, and their presence was substantially under-represented in many tissues (Fig. 21A). The trend of venom homologs to be present at greater than expected frequencies in a single tissue was also found at higher thresholds of >10 CPM, >100 CPM (Figs. 21B-C), and >1,000 CPM (data not shown). Last we asked if the variation in venom homolog expression across tissues was significantly different than that of all python genes, and found that python venom homologs tended to show greater variation in expression levels across tissues, based on the standard error in expression levels across tissues (Fig. 21D).

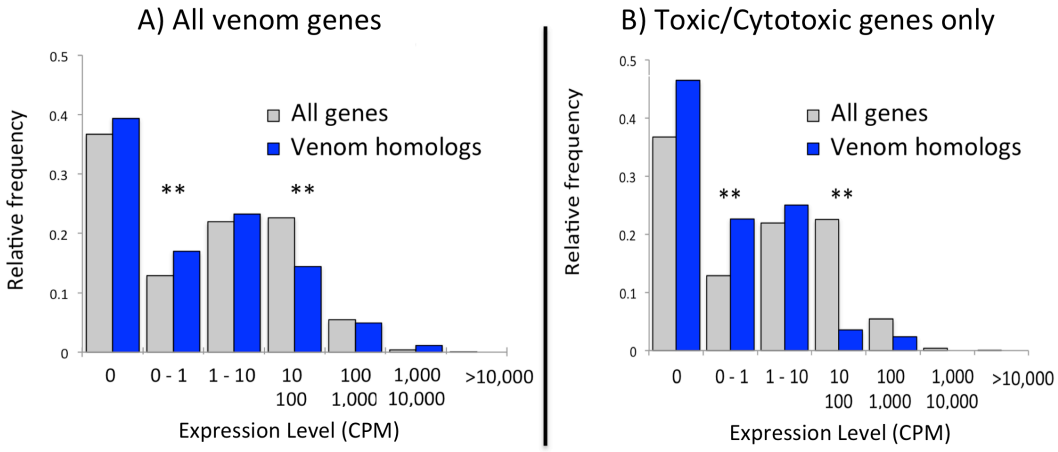


Figure 20 - Relative frequencies of genes observed at different expression levels calculated across all tissues. Results are shown for (A) all venom gene homologs, and (B) venom gene homologs that are known to be cytotoxic only. Asterisks represent expression-level bins where the difference between venom homologs and all genes is statistically significant (Fisher's exact test, p-value <0.05).

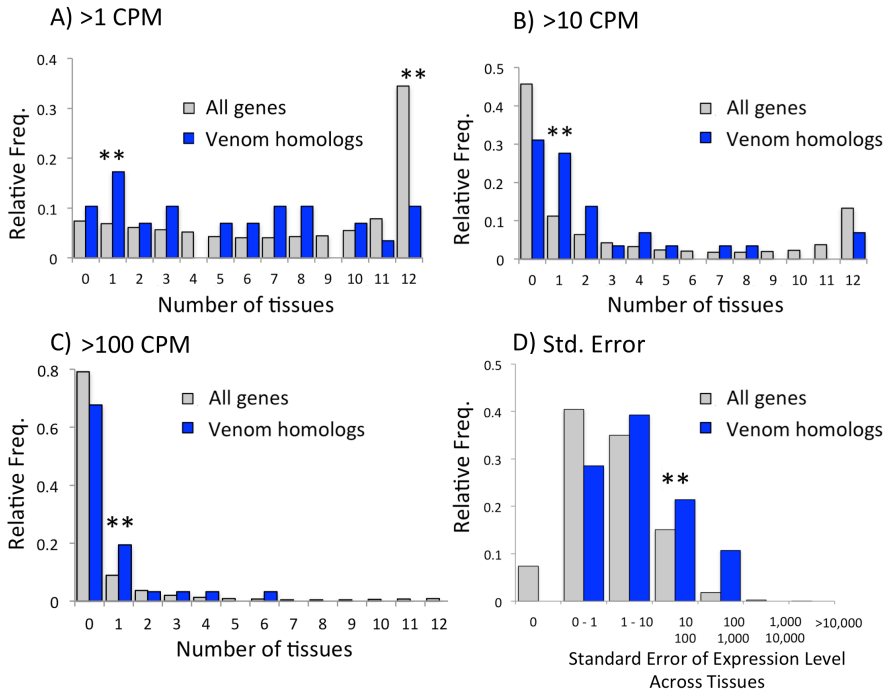


Figure 21 The numbers of tissues in which genes are expressed and variation in expression across tissues. In A-C, different CPM values are used in different panels as thresholds for the 'presence' of a gene being expressed in a given tissue; (A) threshold = >1 CPM, (B) threshold = >10 CPM, and (C) threshold = >100. Asterisks represent bins were the difference between venom homologs and all genes is statistically significant (Fisher's exact test, p-value <0.05). (D) Comparison of standard error in expression level across tissues for all genes and venom gene homologs. Asterisks represent bins where the difference between venom homologs and all genes is statistically significant (Fisher's exact test, p-value <0.05)

Discussion

Our findings provide broad evidence that there are one or two venom gene orthologs present per venom gene family in the python genome. These gene families appear to have undergone varying degrees of duplication and diversification in highly venomous caenophidian snake lineages (including elapids, viperids, and others) and in several cases, result in large multi-locus gene families that encode many related toxins. The python belongs to a lineage that is the sister group to the caenophidian snakes, which appears to have diverged from caenophidian snakes prior to the expansion and diversification of major venom gene families (Table 1; Fig. 18). These findings have two important ramifications. First, they suggest that regardless of when venom systems may have initially evolved in squamate reptiles, either a single time in the ancestor of the Toxicofera (Fry et al. 2006) or independently in caenophidian snakes and lizards (Kochva 1978), substantial venom gene family expansion and diversification is unlikely to have occurred in snakes prior to the caenophidian lineage (Casewell et al. 2012; Vonk et al. 2013a). The availability of additional genomes from basally-diverging snake lineages (e.g., blindsnakes) would be valuable to test this hypothesis further, as it is possible that instead the python secondarily lost many copies of venom genes that were duplicated early in snake or toxicoferan evolution. However, this alternative hypothesis seems unlikely, as it would require that the python would have independently lost numerouscopies of at least 7 different venom gene families (see table 1). Second, our results indicate that the python provides a reasonable and valuable approximation of ancestral gene expression patterns prior to major venom gene recruitment in caenophidian snakes. Thus, patterns of venom gene homolog expression in the python may provide evidence for biases in the processes of venom gene recruitment in caenophidian snakes related to patterns of expression of ancestral venom gene homologs.

With the increasing availability of transcriptome sequencing, it has become common for researchers to sequence the transcriptome of venom glands or other oral glands of squamate reptiles and other venomous taxa (Casewell et al. 2009; Whittington et al. 2010; Fry et al. 2013; Vonk et al. 2013a). Based on such data, it has also become common to identify transcripts of genes with sequence similarity to known venom toxins, to define these as “venom toxin” transcripts, and in some cases even classify a particular species as ‘venomous’ (Fry et al. 2009; Fry et al. 2010; Fry et al. 2013). Here we compared gene expression of python venom homologs in the rictal gland, an oral gland, to that of other python tissues. We find that although the rictal gland does indeed show expression of many venom homologs, these homologs are also expressed at comparable levels in many other tissues. In some limited cases, such venom homologs are expressed at remarkably high levels in particular organs or tissues (Fig. 19). For example, brain, liver and intestinal tissue all show moderate to high levels for several venom homologs.

Our results, including multiple examples of venom homolog expression across many tissues, argue against the adaptive and functional relevance of simply observing such transcripts in a given tissue, as has also been argued recently by Hargreaves et al. (2014). Expression patterns in the rictal gland (Fig. 19) are intriguing, particularly with regards to 3FTx and C-type lectin orthologs, which at first glance appear to be consistent with previous reports of venom production in some Australian pythons (Fry et al., 2013). Interpreting this data under the Toxicofera hypothesis would suggest that the high amplification of such genes in the rictal gland might be an artifact of the shared evolutionary history of the venom system with other toxicoferans, with the python ‘venom system’ presumably atrophying following a switch to using constriction for prey capture (Fry et al. 2013). However, it is important to note that even these levels observed in the python rectal gland are not particularly unusual compared to expression patterns of other toxin orthologs in various non-gland tissues (Fig. 19). Additionally, in the absence of

functional activity data, caution is required when attempting to extrapolate from protein toxin family (e.g., 3FTx) identification to biological activity, as many toxin family members have diverse actions, which are difficult to correlate with structure. For example, proteins with the canonical 3FTx fold and highly conserved disulfides have pharmacological activities as diverse as neurotoxins and anticoagulants (Heyborne and Mackessy 2013) to salamander pheromones (Palmer et al. 2007) and regulators of limb regeneration (Garza-Garcia et al. 2009). Thus, using such data singularly from an oral gland and reaching the conclusion that venom homolog expression represents evidence of "venom toxin" production, or "venomousness" of a species, would be base-less without additional evidence for a functional role of such gene products.

Our results indicate that the probability of successful recruitment of a particular gene for use in caenophidian venom systems may have been biased by the ancestral expression pattern of that gene. Compared to all other python gene expression profiles, python venom homologs tend to be expressed at lower levels overall, expressed at moderate-high levels in fewer tissues, and show among the highest variation in expression level across tissues. These python venom homologs also tend to have higher expression in a single tissue and tend not be expressed in all tissues.

In highly venomous caenophidian snakes, recent studies have shown that highly toxic venom proteins are expressed at moderate to high levels in the venom gland and low-moderate levels in the accessory venom glands (Vonk et al. 2013a), but there is only limited data on their expression levels in other tissues. What is known about their expression in diverse tissues pertains only to their presence/absence (Hargreaves et al. 2014), which substantially limits insight into their relative biological activity in those tissues, particularly since we find here that python venom homologs may be expressed at levels that span more than four orders of magnitude across tissues. Many caenophidian venom toxins are known to be cytotoxic (Lee 1972), to the extent that they are difficult to study in expression vectors (Brenes et al. 2010); within the caenophidian venom gland,

redundant mechanisms maintain these venom toxins in a competent but inactive state (Mackessy and Baxter 2006). The expression of such genes at high (biologically active) levels in other non-venom-related tissues would thus likely be deleterious. These data collectively indicate that during the evolutionary recruitment of such venom toxins in caenophidian snakes, the evolution of venom protein toxicity and higher levels of "venom toxin" expression in the venom system must have been coordinated with an increase in the degree to which such a toxin's expression is confined to venom system tissues. For most venom gene families in caenophidian snakes, this process also appears to be coupled with gene duplication and neofunctionalization via accelerated point mutation (Nakashima et al. 1995; Deshimaru et al. 1996; Kordiš and Gubenšek 2000), accelerated segment switch in exons (Doley et al. 2008; Doley et al. 2009) and other mechanisms, resulting in a diversity of functionalities housed within a conserved protein scaffold. This functional diversification has been well documented for most of the potent functional toxins of caenophidian snake venoms, including 3FTs, PLA_2s, serine proteinases and metalloproteinases (Lynch 2007; Vaiyapuri et al. 2011; Brust et al. 2013; Sunagar et al. 2013).

Based on biases in the regulatory characteristics we have identified in venom homologs in the python, we propose a step-wise model for how proto-venom genes with such regulatory characteristics might have originally been recruited into snake venom systems. We refer to this model as the step-wise intermediate nearly neutral evolutionary recruitment (SINNER) model. This model has three main steps which may or may not involve gene duplication: 1) expression of proto-venom genes in oral secretory glands at low levels, which is favored as a default by regulatory architecture favoring low near constitutive expression, 2) switching of tissue-specific higher expression levels to target oral/venom glands and 3) reduction in expression levels in non-venom-related tissues that is driven by the degree of toxicity to the tissue itself. In this model, the evolution of toxicity (i.e., neofunctionalization) would be constrained by two factors: the functional

requirements and expression levels of the protein in non-venom tissues. Gene duplication would release the first of these two constraints, allowing the evolution of reduced expression in non-venom tissues, and thus allowing the evolution of greater toxicity to prey. This SINNER model therefore implies the existence of a nearly neutral intermediate phase during which the pace of evolution of the toxicity of a venom gene product is balanced by the tissue-specificity and magnitude of its expression, and it accounts for variation in the evolution of toxicity of such venom homologs in various lineages due to differential patterns of drift and selection. The SINNER model thus successfully predicts that a large number of different gene families may exist in venom systems and possess members with different toxicity and expression levels in different lineages, as the expression of those gene families in the venom gland (or any tissue) is a physiological default to some extent. Also, different genes may occupy one of an infinite number of steps along the continuum of the recruitment model's nearly neutral intermediate landscape due to both selection and drift. For example, even though three-finger toxins do not constitute the main components of viperid venom, they are still expressed in viperid venom glands (Pahari et al. 2007); on the other hand, metalloproteinases and serine proteinases, both important components of viper venom, but not of elapid venom, are still expressed but at low levels in elapid venom glands (Correa-Netto et al. 2011; Jiang et al. 2011; Margres et al. 2013).

Some venom genes are also known to produce multiple splice variants (Ducancel et al. 1993; Cousin et al. 1998; Siigur et al. 2001), and it is relevant to consider how these alternative transcripts may contribute to evolution under the SINNER model. If alternative splicing were capable of producing toxic and non-toxic peptides from the same gene, this would decrease the relative role of gene duplication, and would also increase the number of evolutionarily labile features that could act to shift venom toxin expression towards venom gland specificity. Specifically in the case of alternative splice variants, evolution could act on siRNAs and spliceosomal components, in addition to

promoter/enhancer/repressor regulatory elements, to accomplish venom gland targeting of toxic peptides; thus, alternative splicing may act to increase the evolvability and rate of progression of genes across the continuum of the SINNER model.

While we have developed the SINNER model of gene functional recruitment based on snake venom genes, many aspects of this model may apply equally well to other instances of evolutionary co-option of genes that involve duplication and sub-/neo-functionalization. Particularly when there is selection for novel tissue function (e.g., salivary-to-venom gland function), genes that are essentially constitutively expressed in many tissues at low levels and at higher levels in a small number of tissues may be important 'raw material' for shifting tissue function via co-option of these genes in a variety of biological circumstances. It is likely that the SINNER model of gene co-option and recruitment may also fit the evolution of venom systems in other animals, and comparative analysis of gene expression across diverse tissues and venomous and non-venomous sister lineages will be important for evaluating the explanatory power of this model in these systems. One prediction of the SINNER model is that a venom repertoire should contain a diverse collection of gene families, some of which are expressed as a physiological default, and some will be intermediate on the spectrum between high secretion level, venom system specificity, and toxicity, and thus will not be particularly toxic. For example, even though three-finger toxins do not constitute the main toxic components of viperid venom, they are still expressed in viperid venom glands (Pahari et al. 2007). Similarly, metalloproteinases and serine proteinases, both important functional components of viper venom, but not of elapid venom, are expressed at low levels in elapid venom glands (Correa-Netto et al. 2011; Jiang et al. 2011; Margres et al. 2013). Some of the most common venom components include CRISp, waprin/kunitz, hyaluronidases, serine proteases and PLA_2, among many others (Fry et al. 2009), and even thought the same venom protein families can be found across venoms of several animal phyla, their unique patterns of expression, functionality and toxicity can vary

considerably among species (e.g. (Kreil 1995; Ma et al. 2010; Whittington et al. 2010; Ruder et al. 2013; Undheim et al. 2014), which is consistent with predictions from the SINNER model.

Similar to our study, a recent study also found evidence for the presence of venom homologs and known venom genes in diverse tissues of non-venomous and venomous snakes, respectively (Hargreaves et al. 2014). Based on these data, the authors argue for a shift in the otherwise broadly-accepted model of venom gene duplication and recruitment, and suggesting instead that this processed be viewed as "restriction" rather than recruitment because venom genes do not appear to be targeted *de novo* to venom glands but instead are "restricted" to venom systems over evolutionary time. Their conclusions do share some aspects of our SINNER model in that venom genes are not likely *de novo* targeted to the venom gland, but instead undergo a spectral evolutionary transition towards venom gland-specific targeting. Analysis of next-generation RNAseq data to measure expression is so highly sensitive to extremely rarely expressed transcripts, however, that their use of a "presence-absence" detection of venom-related transcrips is potentially misleading and is capable of detecting transcripts far below the levels at which they will produce physiologically relevant biologically active proteins. Thus future work examining organism-wide patterns of venom gene expression should carefully consider the relative frequencies of venom homologs in the context of estimating patterns of expression across tissues to differentiate between biologically relevant expression levels and extremely rare transcripts due, for example, to slightly 'leaky' promoters.

As additional genomic and transcriptomic information becomes available for snakes, particularly different lineages of highly venomous caenophidian snakes as well as in more basally-diverging lineages of snakes (e.g., Scolecophidian blindsnakes) and toxicoferan lizards, it will be interesting to further test the SINNER model for snake venom gene recruitment, and the hypothesis that venom gene expansion occurred "late",

in the caenophidian lineage. Such diverse sampling across the toxicoferan tree is ultimately required to more definitively determine how evolution has shaped tissue expression patterns of venom homologs in the development of squamate venom systems. The SINNER model, our data from the python, other evidence for venom homolog expression in multiple non-venom gland tissues in other venomous and non-venomous snakes (Rádis-Baptista et al. 2003; Whittington et al. 2008; Hargreaves et al. 2014), and the lack of evidence for toxicity or function of multiple venom components relevant for prey capture (Lavin MF 2010; Ahmed et al. 2012; Fry et al. 2012), collectively suggest that a strict and static definition of a gene family as representing “venom toxins” is inaccurate. Instead, these data indicate that a set of venom gland (or other oral gland) secretions may represent a collection of proteins that span the full continuum of stages in the evolution of toxicity and functionality as venoms, some of which may be present largely due to random processes rather than selection for function as venom. Thus, the definition of proteins as “venom toxins” based solely on homology in the absence of functional evidence of toxic effects on prey (or other functional advantages for prey handling) may be misleading. Accordingly, our results indicate the need for a critical re-evaluation of the criteria required to consider a protein a “venom toxin” across the tree of life, not only in snakes. We suggest that such criteria should incorporate more direct evidence for the toxicity or function of such proteins in prey handlin

Appendix A

Supplementary Material for Chapter 1

Supplementary Table 1. Specimen voucher information, locality information, GenBank accession number and references for sequences used in this study. Dash line indicated missing loci for that particular specimen. Except for the species *Crotalus lannomi, C. stejnegeri, and C. ericsmithi*, sequences from different voucher specimens were combined for each taxon.

Species	Voucher	Country	State	ATP6,8	Cyt-b	ND4	C-mos	NT3	RAG-1	Reference
Agkistrodon contortrix	Actx3	USA	Texas	FJ417879.1	-	-	-	-	-	Douglas et al., 2009
-	ROM18230	Unknown	Unknown	-	AF259154.1	-	-	-	-	Murphy et al., 2002
-	HWG2218	USA	Texas	-	-	AF156577.1	-	-	-	Parkinson et al., 2000
-	LSUH0607	Unknown	Unknown	-	-	-	-	JN703027.1	EU402833.1	Weins et al. 2008; 2012
Agkistrodon piscivorus	Unknown	Unknown	Unknown	FJ417905.1	-	-	-	-	-	Douglas et al., 2009
-	LSUMZ-17943	Unknown	Unknown	-	DQ523161.1	DQ523161.1	-	-	-	Jiang, et al. 2007
-	CAS214406	USA	Florida	-	-	-	AF471096.1	-	-	Lawson et al., 2005
Crotalus adamanteus	UTEP18472	USA	Florida	-	-	JN620959.1	JN620891.1	-	-	Anderson & Greenbaum,

-	CLP-4	USA	Florida	-	AY223605.1	-	-	-	-	2012 Parkinson, 1999
-	JMM 619	USA	Mississippi	KF410269	-	-	-	-	KF410328	This study
Crotalus aquilus	ROM 18117	México	SnL. Potosi	-	AF259162.1	HQ257878.1	-	-	-	Murphy et al., 2002
-	UTA: JMM 618	México	México	KF410270	-	-	-	-	-	This study
Crotalus armstrongi	ROM 47025	México	Colima	-	-	HQ257812.1	-	-	-	Bryson et al., 2011b
-	ROM 47013	México	Jalisco	HQ257745.1	-	-	-	-	-	Bryson et al., 2011b
Crotalus atrox	UTEP: 19051	USA	Arizona	-	-	JN620960.1	-	-	-	Anderson & Greenbaum, 2012
-	UTEP: 19849	USA	Texas	JN620852.1	-	-	-	-	JN621002.1	Anderson & Greenbaum, 2012
-	ROM 18144	USA	California	-	AF259186.1	-	-	-	-	Murphy et al., 2002
-	UTA: JAC26689	México	Durango	-	-	-	KF410299	KF410312	-	This study
Crotalus basiliscus	ROM 18188	México	Nayarit	-	AF259174.1	-	-	-	-	Murphy et al., 2002
-	822	Méxi	Nay	-	-		-	-	-	Wüst

		co	arit			AY704894.1				er et al., 2005
-	UTA: JAC 30443	México	Colima	-	-	-	-	KF410313	-	This study
Crotalus cerastes	Unknown	USA	Arizona	DQ493803.1	-	-	-	-	-	Douglas et al., 2006
-	UTA: CJF-5316	USA	Arizona	-	KF410282	KF410288	KF410300	KF410314	KF410329	This study
Crotalus durissus	ROM-FC 265	Venezuela	Unknown	-	AF259178.1	-	-	-	-	Murphy et al., 2002
-	775	Venezuela	D.F.	-	-	AY704875.1	-	-	-	Wüster et al., 2005
Crotalus enyo	UTA: JMM 601	México	Baja California	KF410271	-	-	-	-	-	This study
Crotalus ericsmithi	UTA: ENS 11816	México	Guerrero	-	-	KF410289	-	-	-	This study
-	MZFC: 26942 (JRV-187)	México	Guerrero	KF410272	KF410284	KF410290	KF410301	KF410315	KF410330	This study
Crotalus horridus	UTEP 19288	USA	Texas	-	-	JN870207.1	-	-	JN621004.1	Anderson & Greenbaum, 2012
-	UTA: JMM 501	USA	Texas	KF410273	KF410286	-	KF410302	KF410316	-	This study
Crotalus intermedius	ROM 47124	México	Veracruz	-	-	JN022852.1	-	-	-	Bryson et al., 2011a

-	ROM 47122	México	Guerrero	JN022803.1	-	-	-	-	-	Bryson et al., 2011a
-	ROM 18164	México	Veracruz	-	AF259168.1	-	-	-	-	Murphy et al., 2002
Crotalus lannomi	MZFC: JRV 01	México	Colima	KF410274	KF410287	KF410291	-	KF410317	-	This study
-	MZFC: 26949 (JRV-131)	México	Colima	KF410275	KF410280	KF410292	KF410303	KF410318	KF410331	This study
-	JRV BM	México	Colima	-	-	KF410293	KF410304	KF410319	-	This study
Crotalus lepidus	UTEP:19237	USA	Texas	JN620857.1	-	JN620966.1	JN620898.1	-	-	Anderson & Greenbaum, 2012
-	ROM 18128	México	Chihuahua	-	AF259160.1	-	-	-	-	Murphy et al., 2002
Crotalus mitchellii	UTEP:17628	USA	California	JN620858.1	-	JN620967.1	JN620899.1	-	JN621006.1	Anderson & Greenbaum, 2012
-	ROM 18178	USA	California	-	AF259180.1	-	-	-	-	Murphy et al., 2002
Crotalus molossus	UTEP 19982	USA	New Mexico	JN620881.1	JN620840.1	JN620990.1	JN620919.1	-	JN621025.1	Anderson & Greenbaum, 2012

Crotalus oreganus	ROM 19656	USA	California	-	AF259183.1	AF194152.1	-	-	-	Murphy et al., 2002
-	UTEP 17519	USA	California	JN620855.1	-	-	JN620896.1	-	JN621008.1	Anderson & Greenbaum, 2012
Crotalus polystictus	UTA: JAC 26918	México	Jalisco	KF410276	KF410281	KF410294	KF410305	KF410320	KF410332	This study
Crotalus pricei	UTA: JAC 29184	México	Chihuahua	-	-	-	KF410306	KF410321	-	This study
-	ROM 47091	México	Nuevo Leon	-	-	JN022878.1	-	-	-	Bryson et al., 2011a
-	ROM-FC 2144	México	Nuevo Leon		AF259167.1	-	-	-	-	Murphy et al., 2002
-	ROM 47057	México	Jalisco	JN022839.1	-	-	-	-	-	Bryson et al., 2011a
Crotalus pusillus	ROM 47055	México	Michoacán	HQ257696.1	-	-	-	-	-	Bryson et al., 2011b
-	ROM-FC 271	México	Michoacán	-	AF259159.1	HQ257880.1	-	-	-	Murphy et al., 2002; Bryson et al, 2011b
Crotalus ravus	ROM 47049	México	Oaxaca	HQ257700.1	-	-	-	-	-	Bryson et al.,

-	UTA live collection	México	Puebla	-	AY223609.1	AY223647.1	-	-	-	2011b Parkinson, 1999; Parkinson et al., 2002
Crotalus ruber	ROM 18197	USA	California	-	AF259191.1	HQ316632.1	-	-	-	Murphy et al., 2002
-	CRa10	USA	California	DQ493801.1	-	-	-	-	-	Douglas et al., 2006
Crotalus scutulatus	ROM 18210	USA	Arizona	-	AF259184.1	-	-	-	-	Murphy et al., 2002
-	UTEP (CRH) 153	USA	New Mexico	-	-	AF194167.1	-	-	-	Pook et al, 2000.
-	UTA: JAC 29076	México	Chihuahua	-	-	-	-	KF410322	-	This study
Crotalus stejnegeri	MZFC: 26986 (JRV-186)	México	Sinaloa	-	KF410285	KF410295	tKF410307	-	-	This study
-	MZFC: JRV 192	México	Sinaloa	KF410277	KF410283	KF410296	KF410308	KF410323	KF410333	This study
-	MZFC: TDJ 941	México	Sinaloa	-	-	KF410297	KF410309	KF410324	-	This study
Crotalus *tigris*	UTEP:18442	USA	Arizona	JN620859.1	JN620818.1	JN620968.1	JN620900.1	-	-	Anderson & Greenbaum, 2012

-	CLP169	USA	Arizona	-	-	-	-	GQ334665.1	-	Parkinson, 1999; Daza et al., 2009
-	KWS252	USA	Arizona	-	-	-	-	-	KF410334	This study
Crotalus transversus	ROM45249	México	Morelos	JN022789.1	-	JN022840.1	-	-	-	Bryson et al., 2011a
-	KZ-shed skin	México	Unknown	-	AF259169.1	-	-	-	-	Murphy et al., 2002
Crotalus triseriatus	ROM42412	México	D.F.	HQ257656.1	-	-	-	-	-	Murphy et al., 2002
-	ROM18120	México	D.F.	-	AF259164.1	HQ257879.1	-	-	-	Murphy et al., 2002; Bryson et al, 2011b
-	UTA: JAC27451	México	México	-	-	-	KF410310	KF410325	-	This study
Crotalus willardi	W4011	México	Sonora	KF410278	KF410279	KF410298	-	KF410326	KF410335	This study
Gloydius sp.	Unknown	China	Unknown	EU913477.1	EU913477.1	EU913477.1	-	-	-	Chen and Fu, unpublished
-	Unknown	Unknown	Unknown	-	-	-	-	-	AY662614.1	Townsend et al., 2004
-	Unknown	China	Hebei	-	-	-	JQ687523.	-	-	Guo et al.,

							1			2012
Sistrurus catenatus	ROM 41217	Canada	Ontario	HQ257638.1	-	-	-	-	-	Bryson et al., 2011b
-	Moody-502	USA	Texas	-	AY223610.1	AY223648	-	-	-	Parkinson, 1999
-	RLG-369	USA	Texas	-	-	-	KF410311	KF410327	-	This study
Sistrurus miliarius	ROM 18232	USA	Florida	HQ257639.1	-	HQ257760.1	-	-	-	Bryson et al., 2011b
-	UTA-Live	USA	Florida	-	AY223611	-	-	-	-	Parkinson et al., 2002

Supplementary Table 2. Primer names and sequences used for the amplification and sequencing of gene fragments in this study. Sequence length and number of variable sites at each loci are also given.

Locus	Primer	Reference	Primer sequence (5' -3')
ATP6,8	*CmitchF*	Meik et al. 2012	CCGGTCTGAACTCAGATCACGT
	CMR2	Meik et al. 2012	CCGGTCTGAACTCAGATCACGT
Cyt-b	*ATRCB3*	Parkinson et al. 2002	TGAGAAGTTTTCYGGGTCRTT
	GLUDG	Palumbi, 1996	TGACTTGAARAACCAYCGTTG
ND4	*ND4*	Arévalo et al. 1994	CACCTATGACTACCAAAAGCTCATGTAGAAGC
	Leu	Arévalo et al. 1994	CATTACTTTTACTTGGATTTGCACCA
C-mos	*S77*	Cox et al. 2012	CATGGACTGGGATCAGTTATG
	S78	Cox et al. 2012	CCTTGGGTGTGATTTTCTCACCT
NT3	*NT3-F3*	Noonan & Chippindale, 2006	ATATTTCTGGCTTTTCTCTGTGGC
	NT3-R4	Noonan & Chippindale, 2006	GCGTTTCATAAAAATATTGTTTGACCGG
RAG-1	*RAG1_F*	This study	CAGCTGYAGCCARTACCATAAAAT
	RAG1_R	This study	CTTTCTAGCAAAATTTCCATTCAT

Supplementary table 3. Descriptions of evidence for putative errors identified in existing GenBank submissions of *Crotalus* species.

Genebank #	Species	Gene	Source of error
AF057272.1	*C. atrox*	16s	Equal to AF057271, a *Crotalus molossus*. BLASTn places it as *Porthidium*
AF057271.1	*C. molossus*	16s	Equal to AF057272, a *Crotalus atrox*. BLASTn places it as *Porthidium*
AF259175.1	*C. enyo*	Cyt-b	Possible chimeric sequence. BLASTn of half the sequence places it as *C. durissus*
HM631837.1	*C. horridus*	ND4	Sequence with many amino acid changes not shared with any other *C. horridus* or *Crotalus*
AF259243.1	*C. molossus*	12s	Locality listed as “USA: California, Imperial Co.” in GenBank, but listed as “Mexico: Veracruz” in original publication
AF259129.1	*C. polystictus*	16s	Field number for the specimen, ROM-FC 263, is used twice, for this specimen and for a *C. pricei.*
AF259237.1	*C. pricei*	12s	Same as above
HQ257775.1	*C. armstrongi*	ND4	Sequence identical to HQ257880.1, a *C. pusillus*
AF259172.1	*C. willardi*	Cyt-b	Locality in GenBank given as “Mexico: Sonora”, but listed as “Cochise Co., Arizona” in original publication
AF259135.1	*C. molossus*	*16s*	Feld number given in GenBank for this specimen, ROM 18178, is assigned to *C. mitchelli* in original publication

Supplementary Table 4. Collection number and locality data for skeletal preparations examined in this study.

Species	Locality	Catalogue number	Teeth in the Palatine bone
Agkistrodon piscivorus	Oklahoma, USA	UTAR-34945	Present
Crotalus atrox	Texas, USA	UTAR-40712	Present
Crotalus durissus	No data	UTAR-45028	Present
Crotalus horridus	USA	UTAR-40474	Present
Crotalus lepidus	No data	UTAR-40483	Present
Crotalus molossus	Oaxaca, México	UTAR-14512	Present
Crotalus polystictus	Jalisco, México	UTAR-8270	Absent
Crotalus polystictus	Jalisco, México	UTAR-12583	Absent
Crotalus polystictus	Jalisco, México	UTAR-40482	Absent
Crotalus pricei	No data	UTAR-7432	Present
Crotalus stejnegeri	Sinaloa, México	UTAR-10499	Present
Crotalus triseriatus	Jalisco, México	UTAR-6257	Present
Crotalus willardi	No data	UTAR-40529	Present
Sistrurus catenatus	Texas, USA	UTAR-8730	Present

Appendix B

Supplementary Material for Chapter 2

Supplementary Table 1. Collection number and locality data for skeletal preparations examined in this study.

Genus	species	Reference #	Country	State/Province	Locality
Crotalus	*ericsmithi*	ENS 11816	Mexico	Guerrero	La Laguna
Micruroides	*euryxanthus*	AF217823	-	-	-
Micrurus	*altirostris*	AF228429	-	-	-
Micrurus	*baliocoryphus*	AF228433	-	-	-
Micrurus	*bernadi*	M236	Mexico	Puebla	Puebla
Micrurus	*bernadi*	M201	Mexico	Puebla	Cuetzalan
Micrurus	*bogerti*	M680	Mexico	Oaxaca	Metates
Micrurus	*brasilensis*	AF228427	-	-	-
Micrurus	*browni browni*	M287	Guatemala	Huehuetenango	-
Micrurus	*browni browni*	M15	Guatemala	Huehuetenango	Nentón
Micrurus	*browni browni*	M471	Mexico	Chiapas	Tierra y Libertad
Micrurus	*browni browni*	M29	Mexico	Chiapas	Tuxtla Gutierrez
Micrurus	*browni browni*	M339	Mexico	Oaxaca	Cerro Baul
Micrurus	*browni browni*	M338	Mexico	Oaxaca	Cerro Baul
Micrurus	*browni browni*	M337	Mexico	Oaxaca	Cerro Baul
Micrurus	*browni browni*	M296	Mexico	Oaxaca	Sierra Madre del Sur
Micrurus	*browni importunus*	M513	Guatemala	-	-
Micrurus	*browni taylori*	M633	Mexico	Guerrero	Vallecitos
Micrurus	*browni taylori*	M631	Mexico	Guerrero	Chichihualco
Micrurus	*browni taylori*	M463	Mexico	Guerrero	AguadeObispo
Micrurus	*browni taylori*	M434	Mexico	Guerrero	Filo De Caballo
Micrurus	*browni taylori*	M335	Mexico	Guerrero	San Vicente
Micrurus	*browni taylori*	M271	Mexico	Guerrero	Tierra Colorada
Micrurus	*browni taylori*	M260	Mexico	Guerrero	El Pazclar

Micrurus	*browni taylori*	M223	Mexico	Guerrero	Acapulco
Micrurus	*browni taylori*	JRV 258	Mexico	Guerrero	Cacahuamilpa
Micrurus	*browni taylori*	M600	Mexico	Michoacan	Aquila
Micrurus	*browni x bogerti*	M336	Mexico	Oaxaca	Santa Catarina Juquila
Micrurus	*browni x bogerti*	M333	Mexico	Oaxaca	Santa Catarina Juquila
Micrurus	*corallinus*	AF228424	-	-	-
Micrurus	*decoratus*	AF228441	-	-	-
Micrurus	*diastema*	M801	Mexico	Oaxaca	Cerro Baul
Micrurus	*diastema*	M800	Mexico	Oaxaca	Cerro Baul
Micrurus	*diastema*	M266	Mexico	Oaxaca	-
Micrurus	*diastema*	M681	Mexico	Veracruz	El Fortín
Micrurus	*diastema affinis*	M1	Mexico	Oaxaca	Valle Nacional
Micrurus	*diastema aglaeope*	M798	Guatemala	Izabal	-
Micrurus	*diastema aglaeope*	M118	Guatemala	Merendon	-
Micrurus	*diastema aglaeope*	M117	Guatemala	Merendon	-
Micrurus	*diastema aglaeope*	M272	Honduras	-	-
Micrurus	*diastema aglaeope*	M251	Honduras	Copán	-
Micrurus	*diastema alienus*	M50	Mexico	Quintana Roo	-
Micrurus	*diastema alienus*	JRV 260	Mexico	Quintana Roo	Puerto Morelos
Micrurus	*diastema alienus*	M295	Mexico	Yucatan	-
Micrurus	*diastema alienus*	JRV 262	Mexico	Yucatan	Oxkuzcab
Micrurus	*diastema alienus*	JRV 261	Mexico	Yucatan	Oxkuzcab
Micrurus	*diastema apiatus*	M19	Guatemala	Huehuetenango	-
Micrurus	*diastema apiatus*	ENS 8854	Guatemala	Huehuetenango	-
Micrurus	*diastema diastema*	M799	Mexico	Veracruz	Zongolica

Micrurus	*diastema diastema*	M580	Mexico	Veracruz	Los Tuxtlas
Micrurus	*diastema macdougalii*	M469	Mexico	Veracruz	Las Choapas
Micrurus	*diastema macdougalii*	M299	Mexico	Veracruz	Las Choapas
Micrurus	*diastema sapperi*	M307	Belize	Cayo District	-
Micrurus	*diastema sapperi X hippocrepis*	M308	Belize	Cayo District	-
Micrurus	*distans distans*	M760	Mexico	Sinaloa	San Ignacio
Micrurus	*distans distans*	M240	Mexico	Sonora	Alamos
Micrurus	*distans michoacanensis*	M331	Mexico	Michoacan	Ixtapilla
Micrurus	*distans oliveri*	M74	Mexico	Colima	Vicinity of Colima
Micrurus	*distans oliveri*	M632	Mexico	Colima	Ixtlahuacan
Micrurus	*distans oliveri*	M629	Mexico	Colima	AguaFria
Micrurus	*distans oliveri*	M599	Mexico	Colima	Manzanillo
Micrurus	*distans oliveri*	M599	Mexico	Colima	Manzanillo
Micrurus	*distans oliveri*	M599	Mexico	Colima	Paticajo
Micrurus	*distans oliveri*	M284	Mexico	Jalisco	Chamela
Micrurus	*distans oliveri*	M812	Mexico	Michoacan	Coahuayana
Micrurus	*distans oliveri*	M812	Mexico	Michoacan	Coahuayana
Micrurus	*distans oliveri*	M812	Mexico	Michoacan	Coahuayana
Micrurus	*distans oliveri*	M805	Mexico	Michoacan	Morelia
Micrurus	*distans oliveri*	M52	Mexico	Michoacan	Coalcoman
Micrurus	*distans zweifeli*	M301	Mexico	Nayarit	SierraPajaritos
Micrurus	*distans zweifeli*	JRV 179	Mexico	Nayarit	Huajicori
Micrurus	*elegans*	M22	Guatemala	Huehuetenango	-
Micrurus	*elegans*	M14	Guatemala	Huehuetenango	-
Micrurus	*elegans*	M48	Mexico	Oaxaca	Sierra Juarez
Micrurus	*ephippifer*	M286	Mexico	Oaxaca	Mixtequilla
Micrurus	*ephippifer*	M237	Mexico	Oaxaca	Nizanda

Micrurus	*frontalis*	AF228425	-	-	-
Micrurus	*fulvius*	UF92555	USA	Florida	-
Micrurus	*hemprichi*	AF228442	-	-	-
Micrurus	*hippocrepis*	M116	Guatemala	Izabal	Montañas del Mico
Micrurus	*hippocrepis*	M2	Honduras	Cortez	-
Micrurus	*ibiboca*	AF228440	-	-	-
Micrurus	*latifasciatus nuchalis*	M328	Mexico	Oaxaca	Zanatepec
Micrurus	*latifasciatus nuchalis*	CIG 109	Mexico	Oaxaca	Tapanatepec
Micrurus	*lemiscatus*	AF228435	-	-	-
Micrurus	*limbatus*	M129	Mexico	Veracruz	LosTuxtlas
Micrurus	*mosquitensis*	M12	Costa Rica	-	-
Micrurus	*nebularis*	M341	Mexico	Oaxaca	San J. Bautista Atlatlahuaca
Micrurus	*nebularis*	M195	Mexico	Oaxaca	Teotitlan
Micrurus	*nigrocinctus nigrocinctus*	M76	-	-	-
Micrurus	*nigrocinctus melanocephalus*	M120	Costa Rica	Guanacaste	-
Micrurus	*nigrocinctus melanocephalus*	M124	Guatemala	San Marcos	-
Micrurus	*nigrocinctus melanocephalus*	M119	Guatemala	Zacapa	-
Micrurus	*nigrocinctus melanocephalus*	M249	Honduras	Francisco Morazan	Tegucigalpa
Micrurus	*pachecogili*	M300	Mexico	Puebla	Zapotitlan de Salinas
Micrurus	*proximans*	M293	Mexico	Jalisco	Puerto Vallarta
Micrurus	*proximans*	JRV 265	Mexico	Nayarit	Pintadeño
Micrurus	*proximans*	JRV 264	Mexico	Nayarit	Cuarenteño
Micrurus	*proximans*	JRV 263	Mexico	Nayarit	Cuarenteño
Micrurus	*pyrrhocryptus*	AF228434	-	-	-

Micrurus	*ruatanus*	M75	Honduras	Roatan	-
Micrurus	*sp*	JRV 204	Mexico	Colima	Montitlan
Micrurus	*sp*	JRV 132	Mexico	Colima	Minatitlan
Micrurus	*spixii*	AF228443	-	-	-
Micrurus	*surinamensis*	AF228444	-	-	-
Micrurus	*tener*	M449	Mexico	Hidalgo	Tehuetlan
Micrurus	*tener*	M448	Mexico	Hidalgo	Tehuetlan
Micrurus	*tener*	M431	Mexico	Hidalgo	Coyolapa
Micrurus	*tener*	M208	Mexico	Nuevo Leon	Santiago
Micrurus	*tener*	M203	Mexico	Nuevo Leon	Santiago
Micrurus	*tener*	M332	Mexico	Puebla	-
Micrurus	*tener*	M51	Mexico	Queretaro	Neblinas
Micrurus	*tener*	M326	Mexico	Tamaulipas	-
Micrurus	*tener fitzingeri*	M516	Mexico	Guanajuato	San Miguel de Allende
Micrurus	*tener fitzingeri*	M432	Mexico	Guanajuato	Guanajuato
Micrurus	*tener fitzingeri*	M32	Mexico	Morelos	Tepoztlán
Micrurus	*tener microgalbineus*	M326	Mexico	Tamaulipas	Acuña
Sinomicrurus	*japonicus*	AY058971	-	-	-

Supplementary Table 2. Color pattern, morphological measurements and locality data for specimens revised in chapter 2.

Genus	Species	Lat	Long	body bands	tail bands	T. length	Tail_length	Museum
Micrurus	bernadi	20.956116	-97.406336	21				Other
Micrurus	bernadi	21.347224	-97.683333	26	8			Other
Micrurus	browni	15.423912	-89.079897	22	7	371.531	57.963	UTACV
Micrurus	browni	15.517484	-89.363392	24	9	215.959	32.781	UTACV
Micrurus	browni	15.810391	-86.502763	30	8	589.541	83.618	UTACV
Micrurus	browni	16.086949	-93.761312	8	3			Other
Micrurus	browni	16.583335	-98.81667	11				Other
Micrurus	browni	17.0091	-100.087812	14	4			Other
Micrurus	browni	19.301233	-104.067077	17	5			Other
Micrurus	browni	19.410339	-103.643941	16	6			MZFC
Micrurus	browni	19.412008	-103.604103	23	5			Other
Micrurus	browni	19.508877	-101.083676	28	5			INIREN A
Micrurus	browni	19.676839	19.676839	20	5			INIREN A
Micrurus	browni	16.554495	16.554495	17	5			MZFC
Micrurus	browni	17.1596	-99.5307	12	5			MZFC

			72					
Micru		15.1795	-90.2042			680.9		
rus	diastema	54	34	40	9	32	76.43	UTACV
Micru		15.2528	-89.6698			528.4		
rus	diastema	44	80	36	11	31	95.135	UTACV
Micru		15.2619	-89.6716			391.9		
rus	diastema	40	11	29	7	28	40.384	UTACV
Micru		15.2891	-89.6657			650.8		
rus	diastema	85	12	41	9	57	69.794	UTACV
Micru		15.2984	-89.6806			581.1		
rus	diastema	24	38	29	8	6	71.151	UTACV
Micru		15.2991	-89.6655			551.9		
rus	diastema	63	74	46	11	21	60.896	UTACV
Micru								
rus	diastema	15.345	-88.678	32	11			UTACV
Micru								
rus	diastema	15.360	-88.723	33	6			UTACV
Micru								
rus	diastema	15.360	-88.723	23	9			UTACV
Micru		15.4128	-89.2164			561.4		
rus	diastema	15	51	20	9	89	91.477	UTACV
Micru		15.4239	-89.0798			511.9		
rus	diastema	12	97	20	8	11	101.96	UTACV
Micru		15.4239	-89.0798			438.9		
rus	diastema	12	97	21	7	99	51.072	UTACV
Micru		15.4239	-89.0798			492.7		
rus	diastema	12	97	17	6	91	91.895	UTACV
Micru		15.4239	-89.0798			537.4		
rus	diastema	12	97	21	7	4	104.284	UTACV
Micru		15.5146	-91.8696			665.9		
rus	diastema	61	4	26	5	57	66.629	UTACV
Micru		15.5553	-88.6747			825.6		
rus	diastema	25	96	18	6	34	96.439	UTACV
Micru		15.5862				678.5		
rus	diastema	7	-88.3366	34	7	99	78.444	UTACV
Micru		15.6170	-			579.6		
rus	diastema	34	89.4811	23	8	34	82.592	UTACV

			50					
Micru		15.6199	-89.4479			362.7		
rus	diastema	37	24	35	9	21	51.968	UTACV
			-					
Micru		15.6237	89.4121			455.9		
rus	diastema	86	69	31	9	96	55.503	UTACV
			-					
Micru		15.6237	89.4121			464.3		
rus	diastema	86	69	39	9	12	54.592	UTACV
			-					
Micru		15.6237	89.4121			602.2		
rus	diastema	86	69	32	5	9	57.409	UTACV
			-					
Micru		15.6237	89.4121			410.4		
rus	diastema	86	69	27	8	23	48.431	UTACV
			-					
Micru		15.6237	89.4121			346.3		
rus	diastema	86	69	30	11	35	57.867	UTACV
			-					
Micru		15.6237	89.4121			543.9		
rus	diastema	86	69	36	9	25	63.573	UTACV
			-					
Micru		15.6237	89.4121			509.3		
rus	diastema	86	69	34	11	78	96.88	UTACV
			-					
Micru		15.6237	89.4121			514.3		
rus	diastema	86	69	40	7	11	58.265	UTACV
			-					
Micru		15.6237	89.4121			406.0		
rus	diastema	86	69	35	7	98	48.861	UTACV
			-					
Micru		15.6237	89.4121			442.2		
rus	diastema	86	69	29	10	05	76.03	UTACV
			-					
Micru		15.6237	89.4121			491.6		
rus	diastema	86	69	24	8	58	86.458	UTACV
			-					
Micru		15.6237	89.4121			465.3		
rus	diastema	86	69	33	7	2	58.549	UTACV
			-					
Micru		15.6237	89.4121			534.5		
rus	diastema	86	69	32	11	32	105.407	UTACV
			-					
Micru		15.6237	89.4121			589.9		
rus	diastema	86	69	36	8	02	64.439	UTACV
			-					
Micru		15.6237	89.4121			519.0		
rus	diastema	86	69	30	11	96	87.871	UTACV
Micru	diastema	15.6237	-	28	8	493.3	87.236	UTACV

rus		86	89.412169			2		
Micrurus	diastema	15.623786	-89.412169	39	11	527.577	83.254	UTACV
Micrurus	diastema	15.623786	-89.412169	35	12	456.461	73.466	UTACV
Micrurus	diastema	15.623786	-89.412169	28	9	450.142	73.993	UTACV
Micrurus	diastema	15.623786	-89.412169	35	9	603.035	67.497	UTACV
Micrurus	diastema	15.623786	-89.412169	26	8	465.463	85.978	UTACV
Micrurus	diastema	15.623786	-89.412169	30	7	488.32	63.883	UTACV
Micrurus	diastema	15.623786	-89.412169	26	6	329.867	38.131	UTACV
Micrurus	diastema	15.623786	-89.412169	32	10	402.281	65.63	UTACV
Micrurus	diastema	15.623786	-89.412169	39	9	524.947	63.039	UTACV
Micrurus	diastema	15.623786	-89.412169	31	9	483.133	83.394	UTACV
Micrurus	diastema	15.623786	-89.412169	31	11	352.372	55.058	UTACV
Micrurus	diastema	15.623786	-89.412169	33	7	754.96	87.697	UTACV
Micrurus	diastema	15.623786	-89.412169	37	9	609.623	68.743	UTACV
Micrurus	diastema	15.623786	-89.412169	32	13	529.886	99.501	UTACV
Micrurus	diastema	15.623786	-89.412169	26	10	445.483	69.99	UTACV
Micrurus	diastema	15.675097	-88.6875	2	5	576.383	121.184	UTACV

			92					
Micru rus	diastema	15.6852 99	-88.6451 60	14	4	620.17	64.609	UTACV
Micru rus	diastema	15.6929 89	-88.5785 59	15	5	274.011	42.345	UTACV
Micru rus	diastema	15.7038 47	-88.9329 91	24	9	537.85	97.905	UTACV
Micru rus	diastema	15.8346 7	-91.8163 3	24	5	655.257	75.759	UTACV
Micru rus	diastema	15.8728 30	-91.2321 70	51	11	505.512	65.373	UTACV
Micru rus	diastema	15.8728 30	-91.2321 70	67	11	594.815	65.836	UTACV
Micru rus	diastema	15.8851 2	-91.2406 7	53	10	614.942	71.809	UTACV
Micru rus	diastema	15.8851 2	-91.2406 7	56	11	462.643	75.248	UTACV
Micru rus	diastema	17.3823 73	-92.7487 61	19	7			Other
Micru rus	diastema	17.4876 37	-92.0175 18	25				Other
Micru rus	diastema	19.9330 42	-96.8513 36	13	7			Other
Micru rus	diastema	20.6180 81	-87.0943 07	0				Other
Micru rus	diastema	20.7476 01	-86.9795 48	0	2			Other
Micru rus	diastema	15.1694 72	-92.8347 35	12	3	910.496	108.822	JAC
Micru rus	diastema	15.360	-88.723	31	8			UTACV
Micru rus	diastema	15.3735 36	-89.8290 31	35	9	549.283	72.692	UTACV
Micru	diastema	15.377	-88.702	29	10			UTACV

rus								
Micrurus	diastema	15.561818	-90.104877	45	14	358.975	61.255	UTACV
Micrurus	diastema	15.616667	-89.450000	31	6	666.597	78.759	UTACV
Micrurus	diastema	15.623786	-89.412169	27	7	554.101	67.267	UTACV
Micrurus	diastema	15.623786	-89.412169	30	7	665.123	83.339	UTACV
Micrurus	diastema	15.623786	-89.412169	24	7	337.796	59.339	UTACV
Micrurus	diastema	15.623786	-89.412169	23	5	570.852	57.171	UTACV
Micrurus	diastema	15.623786	-89.412169	34	9	711.606	80.043	UTACV
Micrurus	diastema	15.801284	-91.314681	38	10	580.575	106.121	UTACV
Micrurus	diastema	15.810391	-86.502763	36	10	544.225	70.101	UTACV
Micrurus	distans	18.272583	-103.326993	12	4			JRV
Micrurus	distans	18.307502	-103.353 42	12	5			JRV
Micrurus	distans	18.626828	-103.671882	10	4			INIRENA
Micrurus	distans	19.012851	-103.764992	9	5			Other
Micrurus	distans	19.421042	-103.678921	20	7			JRV
Micrurus	distans	19.483528	-104.642999	11	5			INIRENA
Micrurus	distans	19.483529	-104.643009	13	6			Other

Micrurus	distans	19.523813	-105.036951	11	6			Other
Micrurus	distans	21.082541	-102.54455	15	5			Other
Micrurus	distans	21.642088	-104.276526	17	4			Other
Micrurus	distans	23.288044	-106.067219	11	4			Other
Micrurus	hippocrepis	15.547828	-91.850323	24	7	417.764	64.28	UTACV
Micrurus	hippocrepis	15.627929	-89.430580	27	10	563.159	104.47	UTACV
Micrurus	hippocrepis	15.675097	-88.687592	18	4	477.66	52.527	UTACV
Micrurus	hippocrepis	15.675097	-88.687592	12	6	443.968	79.661	UTACV
Micrurus	hippocrepis	15.675869	-88.987706	20	8	304.25	48.244	UTACV
Micrurus	hippocrepis	15.675967	-88.686073	16	4	352.476	35.946	UTACV
Micrurus	hippocrepis	15.678532	-88.682089	15	6	547.505	88.635	UTACV
Micrurus	hippocrepis	15.679616	-88.682960	15	6	509.66	84.921	UTACV
Micrurus	hippocrepis	15.685299	-88.645160	15	4	211.146	19.888	UTACV
Micrurus	hippocrepis	15.685299	-88.645160	10	6	285.879	46.655	UTACV
Micrurus	hippocrepis	15.685299	-88.645160	8	4	476.523	59.923	UTACV
Micrurus	hippocrepis	15.685299	-88.645160	7	4	629.019	72.837	UTACV
Micrurus	hippocrepis	15.6852	-	15	7	405.4	71.63	UTACV

rus	s	99	88.645160			68		
Micrurus	hippocrepis	15.685299	-88.645160	13	6	441.46	75.943	UTACV
Micrurus	hippocrepis	15.685299	-88.645160	13	5	490.485	71.3	UTACV
Micrurus	hippocrepis	15.685299	-88.645160	14	3	488.052	43.022	UTACV
Micrurus	hippocrepis	15.68575	-88.644391	15	4	711.403	74.257	UTACV
Micrurus	latifasciatus	16.451578	-94.273662	8	3			Other
Micrurus	latifasciatus var nuchalis	14.599855	-90.495765	30	6	734.617	76.615	UTACV
Micrurus	limbatus	18.374398	-95.015085	15	4	523	73	TA&M
Micrurus	limbatus	18.374398	-95.015085	12	4	513	68	TA&M
Micrurus	limbatus	18.39001	-95.01139	15	3	470	44	TA&M
Micrurus	limbatus	18.438554	-95.022962	12	3	165	20	UMMZ
Micrurus	limbatus	18.52	-95.2	41	4	175	10	UMMZ
Micrurus	limbatus	18.52	-95.2	37	4	585	50	UMMZ
Micrurus	limbatus	18.585186	-95.074498	18	4	310		IBUNAM
Micrurus	nigrocinctus	14.97487	-92.17815	19	4	207	20	AMNH
Micrurus	nigrocinctus	15.262488	-92.61626	12	3			UIMNH
Micrurus	nigrocinctus	15.355452	-92.59504	12	3	207	20	USNM
Micrurus	nigrocinctus	15.38	-92.63	12	4	629	93	USNM

rus	s							
Micru rus	proximans	19.682842	-104.418866	21	5			BYU
Micru rus	proximans	20.511277	-105.31322	20	6			LACM
Micru rus	proximans	22.028303	-104.8791	21	5			UAZ
Micru rus	tener	19.410339	-103.643941	16	6			MZFC
Micru rus	tener	20.543722	-99.614925	19	4			MZFC
Micru rus	tener	20.907594	-100.742055	16	4			Other
Micru rus	tener	21.060783	-98.880196	25				TCWC
Micru rus	tener	21.794494	-98.950251	24	4			UMMZ
Micru rus	tener	21.804219	-99.197056	17	6			Other
Micru rus	tener	22.40277	-97.92421	14	4			USNM
Micru rus	tener	22.40277	-97.92421	13	3			Other
Micru rus	tener	22.658332	-98.253057	15	5	342	35	MCZ
Micru rus	tener	22.717649	-98.970944	16	4			TCWC
Micru rus	tener	23.037822	-99.15062	14	3	618	98	UM
Micru rus	tener	23.037822	-99.15062	19	3	578	51	UMMZ
Micru rus	tener	23.037822	-99.15062	16	3	703	67	UMMZ
Micru	tener	23.0378	-	12	3	435	58	UMMZ

rus		22	99.15062					
Micrurus	tener	23.037822	-99.15062	12	3	604	87	UMMZ
Micrurus	tener	23.037822	-99.15062	13	4	695	103	UMMZ
Micrurus	tener	23.037822	-99.15062	18	5	713	114	UMMZ
Micrurus	tener	23.067207	-99.124812	14	2	604	55	UMMZ
Micrurus	tener	23.067207	-99.124812	15	4	493	51	UMMZ
Micrurus	tener	23.099722	-99.1925	14	4	240	33	UMMZ
Micrurus	tener	23.175925	-99.303552	13	5			Other
Micrurus	tener	23.201777	-98.437327	16	5	364	52	UMMZ
Micrurus	tener	23.222185	-98.385914	17	4	754	79	UMMZ
Micrurus	tener	23.982951	-98.835285	20	4			MCZ
Micrurus	tener	24.601131	-99.01344	18	3	63.5	5.3	TCWC
Micrurus	tener	24.601131	-99.01344	16	5	63.6	8.8	TCWC
Micrurus	tener	24.601131	-99.01344	13	4	46.6	5.8	TCWC
Micrurus	tener	24.77	-98	20	5	703	95	CUMV
Micrurus	tener	26.787083	-102.000722	14	4			MZFC
Micrurus	tener	23.067207	-99.124812	14	4	551	79	UMMZ
Micru	tener	20.3305	-	15	5			MSU

rus	fitzingeri	36	99.553179					
Micrurus	tener fitzingeri	20.530556	-100.354722	19	4			Minton(?)
Micrurus	tener fitzingeri	20.54	-100.44	20	3			UMMZ
Micrurus	tener fitzingeri	20.606811	-100.36841	16	3			TCWC(?) "Hardy"
Micrurus	tener fitzingeri	20.606811	-100.36841	19	5			UMMZ
Micrurus	tener fitzingeri	21.019014	-101.257359	24	4	850	74	USNM
Micrurus	tener fitzingeri	21.019014	-101.257359	20	4	767	65	USNM
Micrurus	tener fitzingeri	21.42055556	-99.60083333	22	5	538	25	TCWC
Micrurus	tener microgalbineus	21.384173	-98.99217	24	5	630	60	LSUMZ
Micrurus	tener microgalbineus	22.15	-98.53333	20	4			KU
Micrurus	tener microgalbineus	22.490455	-99.083791	19	5			TAY-SM
Micrurus		15.886330	-91.245500	59	11	645.672	79.826	UTACV
Micrurus		16.24013	-97.28312	20	5	522.151	69.362	JAC
Micrurus		16.24013	-97.28312	19	6	357.098	60.007	JAC
Micrurus		16.375976	-95.259938	15	4	937.808	112.269	JAC
Micrurus		16.48333	-94.363034	8	3	632.848	127.848	JAC
Micrurus		16.489977	-94.11061	14	5	410.973	70.925	UTACV

Micrurus	16.553995	-94.183326	21	6	380.8	85.02	JAC
Micrurus	16.553995	-94.183326	14	4	403.778	71.505	JAC
Micrurus	16.553995	-94.183326	17	4	531.167	72.961	JAC
Micrurus	16.555083	-94.183826	22	5	839.366	96.185	UTACV
Micrurus	16.555083	-94.183826	25	6	753.681	97.967	UTACV
Micrurus	16.555083	-94.183826	24	5	750.696	82.134	UTACV
Micrurus	16.555083	-94.183826	26	7	679.448	109.786	UTACV
Micrurus	16.555083	-94.183826	23	7	559.179	87.386	UTACV
Micrurus	16.555083	-94.183826	26	6	618.582	74.851	UTACV
Micrurus	16.555083	-94.183826	26	5	889.692	97.674	UTACV
Micrurus	16.555083	-94.183826	25	8	568.089	82.131	UTACV
Micrurus	16.6589	-95.0109	21	5	439.874	53.176	UTACV
Micrurus	16.734034	16.734034	19	9			UTACV
Micrurus	16.8636	-99.8825	14	5	590.752	104.02	UTACV
Micrurus	16.8636	-99.8825	13	3	449.637	54.83	MZFC
Micrurus	16.902691	-89.835859	31	7			UTACV
Micrurus	16.917779	-89.884476	24	8			UTACV
Micrurus	16.981112	-89.9015	28	6			UTACV

Micrurus	16.998481	-74-89.726909	31	8			UTACV
Micrurus	17.058119	-91.202015	34	8			UTACV
Micrurus	17.22	-97.01	23	6	569.109	85.668	JAC
Micrurus	17.224865	-89.614027	26	8			UTACV
Micrurus	17.224865	-89.614027	29	11			UTACV
Micrurus	17.224865	-89.614027	25	6			UTACV
Micrurus	17.23	-97	21	4	543.153	65.648	UTACV
Micrurus	17.29061	-100.27955	14	3	658.14	84.75	JAC
Micrurus	17.360233	-99.467235	17	4	1041.419	103.718	JAC
Micrurus	17.58757	-99.80267	18	4	666.58	75.885	UTACV
Micrurus	17.631981	-96.342408	21	11	446.173	70.744	UTACV
Micrurus	17.632889	-96.756861	22	7	626.603	107.003	LCM
Micrurus	17.692934	-96.326906	21	10	576.765	99.724	UTACV
Micrurus	17.692934	-96.326906	20	6	209.45	21.297	UTACV
Micrurus	17.692934	-96.326906	18	8	469.298	75.049	UTACV
Micrurus	17.692934	-96.326906	26	7	531.591	57.084	UTACV
Micrurus	17.692934	-96.326906	19	6	495.384	51.383	UTACV

Micru rus	17.6929 34	-96.3269 06	13	7	353.8 27	51.239	UTACV
Micru rus	17.6929 34	-96.3269 06	21	9			UTACV
Micru rus	17.8602 06	-96.2099 03	19	6	590.2 89	98.979	UTACV
Micru rus	17.9046 43	-101.343 091	25	6	526.0 8	67.77	JAC
Micru rus	18.12	-97	27	8	197.6 47	30.262	UTACV
Micru rus	18.1382 83	-97.0176 5	25	7	236.5 5	38.874	UTACV
Micru rus	18.3108	-95.1048	13	7			ENS
Micru rus	18.3108	-95.1048	17	6	552.9 69	66.133	UTACV
Micru rus	18.3108	-95.1048	27	7	557.0 88	67.198	UTACV
Micru rus	18.3108	-95.1048	19	6	709.0 7	82.931	UTACV
Micru rus	18.3315 27	-95.0927 79	14	3	562	50	HWG-R
Micru rus	18.3315 27	-95.0927 79	13	7	293.1 59	45.632	ENS
Micru rus	18.3315 27	-95.0927 79	22	4	529.5 89	68.696	UTACV
Micru rus	18.3720 09	-95.0031 02	22	3			UTACV
Micru rus	18.3880 56	-94.94	16	3	509.1 25	44.312	UTACV
Micru rus	18.3880 56	-94.94	13	4	430.6 57	72.297	UTACV
Micru rus	18.3880 56	-94.94	16	3	378.5 7	40.537	UTACV
Micru rus	18.4160 1	-103.532 06	10	4			JAC
Micru rus	18.8494 57	-97.0666 37	18	5	727.2 42	71.408	UTACV
Micru	18.8743	-	15	3	511.6	46.349	ENS

rus		9	96.90494			19		
Micrurus		19.301233	-104.067077	19	5	401.02	47.899	UTACV
Micrurus		19.510904	-105.035696	12	6	432.697	75.448	UTACV
Micrurus		20.020327	-97.457428	28	9	535.826	97.947	UTACV
Micrurus		20.063	-97.475	25	11	435.037	69.693	MZFC
Micrurus		20.509108	-97.66876	33	7	933.907	86.948	UTACV
Micrurus		20.51215	-88.7182 7	8	N/A	400.776	N/A	JAC
Micrurus		20.71136	-88.43466	10	4			JAC
Micrurus		20.71136	-88.43466	10	4	551.237	106.46	JAC
Micrurus		20.7424	-88.21232	13	5	446.208	74.014	JAC
Micrurus		20.916944	-99.766944	24	4	507.71	49.02	UTACV
Micrurus		21.59	-104.21	19	7	483.4	81.75	JAC
Micrurus		27.067697	-109.044177	11	4	665.109	88.429	UTACV
Micrurus	bernadi	20.211915	-98.011547	25	6	278	28	AMNH
Micrurus	bernadi	20.211915	-98.011547	32	6	617	61	AMNH
Micrurus	bernadi	20.211915	-98.011547	24	5	278	28	AMNH
Micrurus	bernadi	20.211915	-98.011547	32	6	247	22	MCZ
Micrurus	bernadi	20.211915	-98.0115	44	12	504	66	UMMZ

Micrurus	bernadi	20.648525	-98.65767	41	8	826	85	AP
Micrurus	bernadi	20.96	-97.78	27	6	507	50	UMMZ
Micrurus	bogerti	15.772789	-96.097052	19	6	344	50	AMNH
Micrurus	bogerti	16.333402	-95.229467	18	4	730	76	MCZ
Micrurus	bogerti	16.432136	-94.10619	13	3	587	65	AMNH
Micrurus	bogerti	16.483333	-94.35	13	3	770	91	UIMNH
Micrurus	browni	15.34812	-92.59961	20	5			?
Micrurus	browni	15.361933	-90.601654	22	5			UMMZ
Micrurus	browni	15.44	-92.89	29	6	700	72	UCM
Micrurus	browni	16.169121	-97.092287	26	8	407	56	UCM
Micrurus	browni	16.169121	-97.092287	20	7	595	89	UCM
Micrurus	browni	16.169121	-97.092287	23	8	666	103	UCM
Micrurus	browni	16.169121	-97.092287	18	8	613	92	UCM
Micrurus	browni	16.169121	-97.092287	23	7	595	86	UCM
Micrurus	browni	16.169121	-97.092287	24	6	595	84	UCM
Micrurus	browni	16.169121	-97.092287	19	7	605	93	UCM
Micrurus	browni	16.169121	-97.092287	27	8	715	72	UCM

Micrurus	browni	16.169121	-97.092287	20	6	563	86	UCM
Micrurus	browni	16.169121	-97.092287	24	8	479	67	UCM
Micrurus	browni	16.169121	-97.092287	21	7	597	92	UCM
Micrurus	browni	16.169121	-97.092287	22	7	545		UCM
Micrurus	browni	16.169121	-97.092287	20	6	537	85	UCM
Micrurus	browni	16.169121	-97.092287	22	6			UCM
Micrurus	browni	16.483333	-94.35	26	8	835	123	UIMNH
Micrurus	browni	16.483333	-94.35	23	7	402	57	UIMNH
Micrurus	browni	16.483333	-94.35	24	7	526	72	UIMNH
Micrurus	browni	16.53	-95.08	23	7	380	47	UIMNH
Micrurus	browni	16.55517	-94.183539	23	7	702	105	AMNH
Micrurus	browni	16.55517	-94.183539	23	7	640	96	AMNH
Micrurus	browni	16.61936	-94.018004	23	6			UIMNH
Micrurus	browni	16.758058	-93.107916	23	4			MZTG
Micrurus	browni	16.758058	-93.107916	20	6			UIMNH
Micrurus	browni	16.75857	-92.940576	19	6	641	90	JFC
Micrurus	browni	16.779479	-93.362608	22	6			MZTG
Micrurus	browni	16.78	-99.79	10	5	469	85	IBUNAM
Micrurus	browni	16.78	-99.79	10	4	417	64	IBUNA

rus								M
Micru rus	browni	16.78	-99.79	11	4	346	42	IBUNAM
Micru rus	browni	16.78	-99.79	11		452		IBUNAM
Micru rus	browni	16.78	-99.79	11	4	363	46	IBUNAM
Micru rus	browni	16.78	-99.79	11	4	170	24	IBUNAM
Micru rus	browni	16.78	-99.79	12	4	360	51	IBUNAM
Micru rus	browni	16.78	-99.79	12	5	441	55	IBUNAM
Micru rus	browni	16.78	-99.79	12	4	416	54	IBUNAM
Micru rus	browni	16.78	-99.79	12	3	428	51	IBUNAM
Micru rus	browni	16.78	-99.79	12	4	435	51	IBUNAM
Micru rus	browni	16.78	-99.79	12	4	515	61	IBUNAM
Micru rus	browni	16.78	-99.79	12	5	495	90	IBUNAM
Micru rus	browni	16.78	-99.79	13	5	570	90	IBUNAM
Micru rus	browni	16.78	-99.79	13	5	540	81	IBUNAM
Micru rus	browni	16.78	-99.79	13	6	256	35	IBUNAM
Micru rus	browni	16.78	-99.79	13	4	504	57	IBUNAM
Micru rus	browni	16.78	-99.79	13	5	457	73	IBUNAM
Micru rus	browni	16.78	-99.79	13	5	518	65	IBUNAM
Micru rus	browni	16.78	-99.79	13	5	546	62	IBUNAM
Micru rus	browni	16.78	-99.79	13	4	462	53	IBUNAM
Micru rus	browni	16.78	-99.79	14		478		IBUNAM
Micru rus	browni	16.78	-99.79	14	4	545	59	IBUNAM
Micru rus	browni	16.78	-99.79	14	4	470	65	IBUNAM
Micru rus	browni	16.78	-99.79	14	4	517	65	IBUNAM
Micru rus	browni	16.78	-99.79	14	4	463	55	IBUNAM
Micru	browni	16.78	-99.79	15	4	508	56	IBUNA

rus								M
Micrurus	browni	16.78	-99.79	17	4	240	21	IBUNAM
Micrurus	browni	16.824084	-93.437314	22	5			USNM
Micrurus	browni	16.824084	-93.437314	20	7			USNM
Micrurus	browni	16.826234	, -99.854159	11	4	799	121	CNHM
Micrurus	browni	16.83	-93.19	21	6	560	79	UMMZ
Micrurus	browni	16.945585	-93.014196	20	7			MZGT
Micrurus	browni	17.36	-99.47	14	5	593	92	TCWC
Micrurus	browni	17.36	-99.47	13	5	476	69	TCWC
Micrurus	browni	17.36	-99.47	13	5	452	65	TCWC
Micrurus	browni	17.36	-99.47	23	6	468	52	TCWC
Micrurus	browni	17.36	-99.47	13	5	550	78	TCWC
Micrurus	browni	17.36667	-99.48333	13	5	6654	105	TCWC
Micrurus	browni	17.36667	-99.48333	16	6	464	66	TCWC
Micrurus	browni	17.48	-99.47	16	5	808	117	TCWC
Micrurus	browni	17.556283	-99.688018	15	6			BM
Micrurus	browni	17.55694444	-99.50083333	22	6	705	96	AMNH
Micrurus	browni	17.55694444	-99.50083333	22	6	705	96	AMNH
Micrurus	browni	17.55694444	-99.50083333	21	8	557	81	AMNH
Micrurus	browni	17.55694444	-99.50083333	21	8	557	81	AMNH

Micrurus	browni	17.556944444	-99.500833333	19	7	689	97	CNHM
Micrurus	browni	17.556944444	-99.500833333	24	5	828	91	MCZ
Micrurus	browni	17.556944444	-99.500833333	23	5	579	65	MCZ
Micrurus	browni	17.556944444	-99.500833333	17	5	602	67	MCZ
Micrurus	browni	17.556944444	-99.500833333	18	6			MVZCU
Micrurus	browni	17.556944444	-99.500833333	17	6			MVZCU
Micrurus	browni	17.556944444	-99.500833333	20	7			MVZCU
Micrurus	browni	17.556944444	-99.500833333	18	7	679	98	UMMZ
Micrurus	browni	17.556944444	-99.500833333	17	5	594	83	UMMZ
Micrurus	browni	17.556944444	-99.500833333	14	5	421	56	UMMZ
Micrurus	browni	17.556944444	-99.500833333	22	6	231	29	UMMZ
Micrurus	browni	18.973076	-99.251881	22	6			USNM
Micrurus	browni	19.194463	-100.132813	22	4			KU
Micrurus	browni	19.42121	19.42121	23	5	700	70	BM
Micrurus	distans	18.9565	-103.949	12	5	649	103	AMNH
Micrurus	distans	18.9565	-103.949	13	5	451	51	UMMZ
Micrurus	distans	19.018249	-104.001271	14	4	761	90	UMMZ
Micrurus	distans	19.049501	-104.294	14	4	950	110	AMNH

			215					
Micru rus	distans	19.1567	-104.469	13	6	670	106	AMNH
Micru rus	distans	19.2328 93	-103.677 923	12	5	797	116	AMNH
Micru rus	distans	19.2359 7	-103.722 976	12	3	370	58	MCZ
Micru rus	distans	19.53	-104.99	12	5	415	60	IBUNA M
Micru rus	distans	19.53	-105.01	12	6	540	85	IBUNA M
Micru rus	distans	19.53	-105.07	12	6	572	86	IBUNA M
Micru rus	distans	19.53	-105.07	11	4	425	60	IBUNA M
Micru rus	distans	20.9102 76	-103.969 51	20	4	810	80	USNM
Micru rus	distans	20.9967 9	-104.061	19	4	249	25	EA Liner
Micru rus	distans	21.7964 19	-105.097 9	17	6	454	64	LBSC
Micru rus	distans	22.8343 38	-105.768 596	14	5	648	94	AMNH
Micru rus	distans	22.8343 38	-105.768 596	12	5	881	111	AMNH
Micru rus	distans	23.15	-106.233 333	13	5	965	137	BM
Micru rus	distans distans	19.0713 7	-102.344 328	9	3	484	51	FA Shanno n
Micru rus	distans distans	21.9333 3333	-105.233 3333	13	6	730	105	AMNH
Micru rus	distans distans	22.291	-105.271 8	14	6	743	76	LBSC
Micru rus	distans distans	22.8343 38	-105.768 596	12	4	882	90	AMNH
Micru rus	distans distans	22.8343 38	-105.768 596	14	3	965	92	AMNH
Micru	distans	22.8343	-	12	5	881	111	AMNH

rus	distans	38	105.768596					
Micrurus	distans	22.991008	-105.851278	14	5	805	91	UKMNH (KU?)
Micrurus	distans	23.193228	-106.26829	12	5	737	105	U Arizona
Micrurus	distans	23.34057	-106.0858	14	6	630	95	KU
Micrurus	distans	23.34754	-105.97392	17		790		USC
Micrurus	distans	23.39213	-106.4409	15	6			FA Shannon
Micrurus	distans	23.41642	-106.4567	12	5	645	95	USC
Micrurus	distans	23.8933	-106.634	14	4			FA Shannon
Micrurus	distans	23.9171	-106.85058	14	4	685	96	USC
Micrurus	distans	23.93293	-106.77402	15	4	820	90	USC
Micrurus	distans	23.95129	-106.75727	13	5	486	70	USC
Micrurus	distans	25.6246	-109.0533	12	3	870	86	AMNH
Micrurus	distans	26.9	-108.65	12	4	780	108	MVZCU
Micrurus	distans	26.9	-108.68333	12				MVZCU
Micrurus	distans	26.9	-108.68333	12	3	449	47	SDSNH (?)
Micrurus	distans	26.9	-108.68333	12	3	1000	96	SDSNH (?)
Micrurus	distans	26.9	-108.68333	12	3	825	90	SDSNH (?)
Micru	distans	26.9	-	12	3	1075	103	UCLA

rus	distans		108.68333					
Micrurus	distans	27.210507	27.210507	11	5			KU
Micrurus	distans	27.404167	-108.322222	12	5			USNM
Micrurus	distans	28.167835	-108.553513	12	3	1015	102	U Arizona
Micrurus	distans michoacanensis	17.21367	-100.639527	7	3	865	88	Guanajuato
Micrurus	distans michoacanensis	19.07137	-102.344328	7	3	471	49	FA Shannon
Micrurus	distans michoacanensis	19.07137	-102.344328	6	3	522	71	FMNH
Micrurus	distans michoacanensis	19.07137	-102.344328	8	2	811	82	FMNH
Micrurus	distans michoacanensis	19.096468	-102.289644	7	3			UMMZ
Micrurus	distans michoacanensis	19.11507	-102.26375	7	3	510	68	UH (?)
Micrurus	distans michoacanensis	19.130969	-102.32084	7	3	646	92	USNM
Micrurus	distans zweifeli	21.36417	-104.578577	19	6			CAS
Micrurus	ephippifer	16.168183	-95.331148	20	6	361	51	AMNH
Micrurus	ephippifer	16.168183	-95.331148	18				UIMNH
Micrurus	ephippifer	16.244503	-95.152099	20	5	588	66	USNM
Micrurus	ephippifer	16.2507	-95.9817	21	4			UIMNH
Micrurus	ephippifer	16.26667	-95.6	18	4			UIMNH
Micrurus	ephippifer	16.278894	-95.5378	18	6	339	46	AMNH
Micru	ephippifer	16.3028	-	21	5	470	53	UIMNH

rus		31	95.50685					
Micrurus	ephippifer	16.302831	-95.50685	19	6	523	76	USNM
Micrurus	ephippifer	16.316667	-95.45	18	6	266	36	UIMNH
Micrurus	ephippifer	16.31667	-95.45	17	4	735	90	UCM
Micrurus	ephippifer	16.31667	-95.45	18	6	550	80	UCM
Micrurus	ephippifer	16.3194	-95.2836	17	6			UIMNH
Micrurus	ephippifer	16.32	-95.24	18	6			USNM
Micrurus	ephippifer	16.3244	-95.23888	22	5	168	79	AMNH
Micrurus	ephippifer	16.33	-95.24	19	4			UIMNH
Micrurus	ephippifer	16.33	-95.24	16	6	340	48	UIMNH
Micrurus	ephippifer	16.333402	-95.229467	22	2	926		AMNH
Micrurus	ephippifer	16.333402	-95.229467	15	4			MCZ
Micrurus	ephippifer	16.333402	-95.229467	20	5			PM
Micrurus	ephippifer	16.333402	-95.229467	23	5	523	56	USNM
Micrurus	ephippifer	16.349748	-94.35405	16	4	806	84	UCM
Micrurus	ephippifer	16.383333	-95.45	21	5	800	75	UIMNH
Micrurus	ephippifer	16.38578	-95.323433	15	5	317	40	AMNH
Micrurus	ephippifer	16.38578	-95.323433	19	5			UIMNH
Micrurus	ephippifer	16.39	-95.35	20	4	461		AMNH
Micrurus	ephippifer	16.39	-95.35	22	5	335	34	AMNH
Micru	ephippifer	16.4333	-94.45	17	6			UIMNH

rus								
Micrurus	ephippifer	16.43994	-95.4827	18	6	722	106	USNM
Micrurus	ephippifer	16.441947	-95.029572	16	9	217	31	UCM
Micrurus	ephippifer	16.441947	-95.029572	18	8	200	25	UCM
Micrurus	ephippifer	16.441947	-95.029572	14	8	615	98	UCM
Micrurus	ephippifer	16.441947	-95.029572	20	8	395	44	UCM
Micrurus	ephippifer	16.441947	-95.029572	14	10	240	36	UCM
Micrurus	ephippifer	16.441947	-95.029572	20	8	260	25	UCM
Micrurus	ephippifer	16.444826	-94.960169	17	4			UMMZ
Micrurus	ephippifer	16.62	-94.97	17	6			UIMNH
Micrurus	ephippifer	16.922152	-96.319816	22	5	550	53	AMNH
Micrurus	ephippifer	16.947222	-95.5167	21	6			UIMNH
Micrurus	ephippifer	17.0751	-96.69473	26	5	217	21	EHT
Micrurus	ephippifer	17.11163	-96.63467	26	5	503	47	AMNH
Micrurus	ephippifer	17.129189	-96.626084	20	4	688	72	AMNH
Micrurus	ephippifer	17.129189	-96.626084	21	4	471	45	AMNH
Micrurus	ephippifer	17.129189	-96.626084	22	6	533	57	AMNH
Micrurus	ephippifer	17.129189	-96.626084	21	6	503	70	UCM
Micrurus	ephippifer	17.1291	-	20	6	603	89	UCM

rus		89	96.626084					
Micrurus	ephippifer	17.129189	-96.626084	18	6			UIMNH
Micrurus	ephippifer	17.129189	-96.626084	18	6	440	58	UIMNH
Micrurus	ephippifer	17.324947	-97.0043	22	7	675	98	AMNH
Micrurus	ephippifer	17.324947	-97.0043	27	6	557	58	AMNH
Micrurus	ephippifer	17.34178	-96.47847	27	5	840	82	AMNH
Micrurus	ephippifer	17.4616	-97.22922	26	6	493	67	AMNH
Micrurus	proximans	21.178	-105.1396	18	5	555	65	AMNH
Micrurus	proximans	21.249946	21.249946	22	7			AM?
Micrurus	proximans	21.626143	-105.15492	18	6	527	77	LBSC
Micrurus	tener maculatus	22.293664	-97.884858	17	4	722	65	BM
Micrurus	tener maculatus	22.293664	-97.884858	15	5	537	74	HM
Micrurus	bernadi	19.950139	-97.489722	19	7			MZFC
Micrurus	bernadi	19.950139	-97.489722	31	8			MZFC
Micrurus	browni	15.924833	-96.421	20	8			MZFC
Micrurus	browni	16.227639	-93.709389	20	5			MZFC
Micrurus	browni	16.475972	-94.12958	18	7			MZFC
Micrurus	browni	16.555083	-94.183826	15	6			MZFC
Micru	browni	16.5550	-	13	3			MZFC

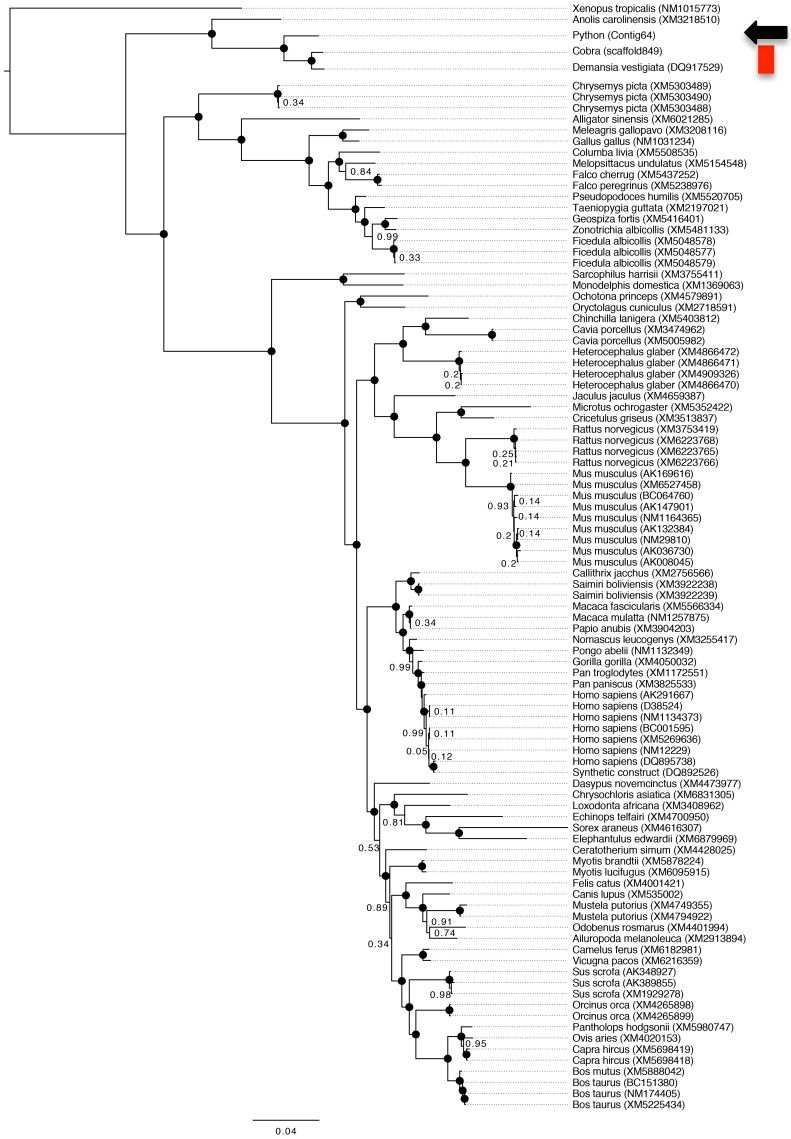
rus		83	94.183826			
Micrurus	browni	16.579639	-94.214556	21	7	MZFC
Micrurus	browni	17.539583	-99.450194	13	6	MZFC
Micrurus	browni	17.539583	-99.450194	11	4	MZFC
Micrurus	browni	17.905459	-101.342284	19	8	MZFC
Micrurus	browni	18.2302	-103.2	21	7	MZFC
Micrurus	browni	19.194463	-100.132813	14	6	MZFC
Micrurus	diastema	18.331527	-95.092779	18	7	MZFC
Micrurus	diastema	16.89966	-90.967348	32	11	MZFC
Micrurus	diastema	17.695556	-96.323139			MZFC
Micrurus	diastema	18.2625	-96.816	23	9	MZFC
Micrurus	diastema	18.54194	-96.9125	25	7	MZFC
Micrurus	diastema	18.891278	-97.015194	10	4	MZFC
Micrurus	diastema	19.768333	-96.860556	22	7	MZFC
Micrurus	diastema	19.77297	-96.867062	24	8	MZFC
Micrurus	diastema	21.138314	-98.412514	18	7	MZFC
Micrurus	distans	19.214573	-104.206209	12	5	MZFC
Micrurus	distans	19.217818	-104.21935	12	5	MZFC

Micrurus	distans	19.692896	-104.423627	18	7	MZFC
Micrurus	distans	20.30672	-103.241199	18	6	MZFC
Micrurus	distans	20.830833	-103.850883	19	7	MZFC
Micrurus	distans	20.830833	-103.850883	21	5	MZFC
Micrurus	distans	22.689281	-105.117487	16	4	MZFC
Micrurus	fulvius	19.392356	-103.672689	23	5	MZFC
Micrurus	fulvius	19.412008	-103.604103	16	5	MZFC
Micrurus	limbatus	18.388056	-94.94	25	2	MZFC
Micrurus	nigrocinctus	16.554495	16.554495	14	5	MZFC
Micrurus	proximans	19.691225	-104.415032	19	3	MZFC
Micrurus	proximans	19.4152	-104.012	20	6	MZFC

Appendix C

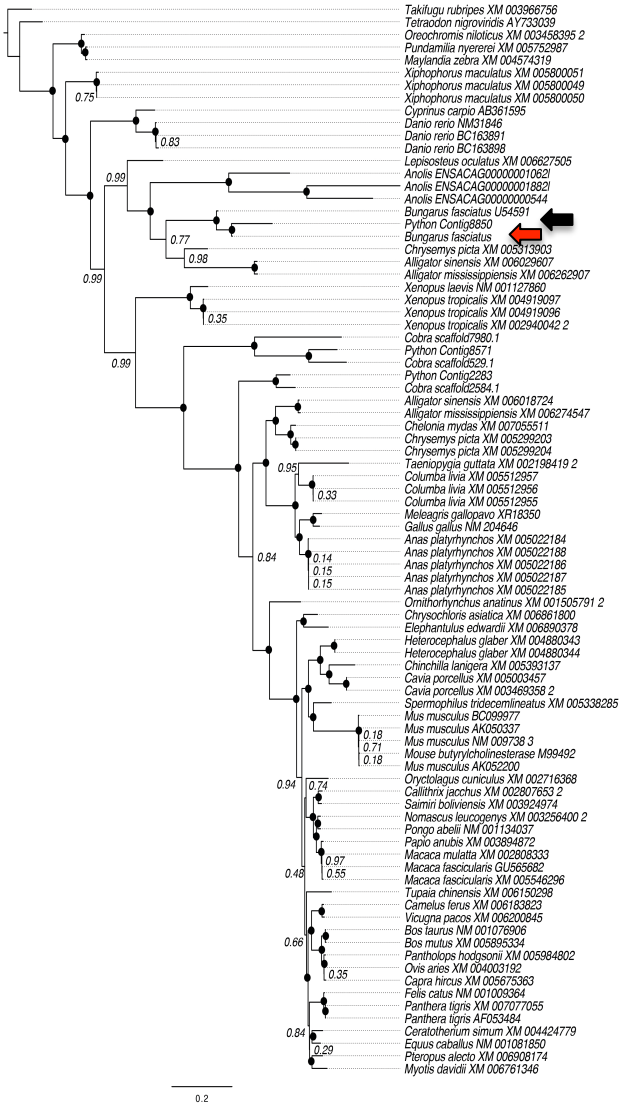
Supplementary Material for Chapter 3

5’-nucleotidase



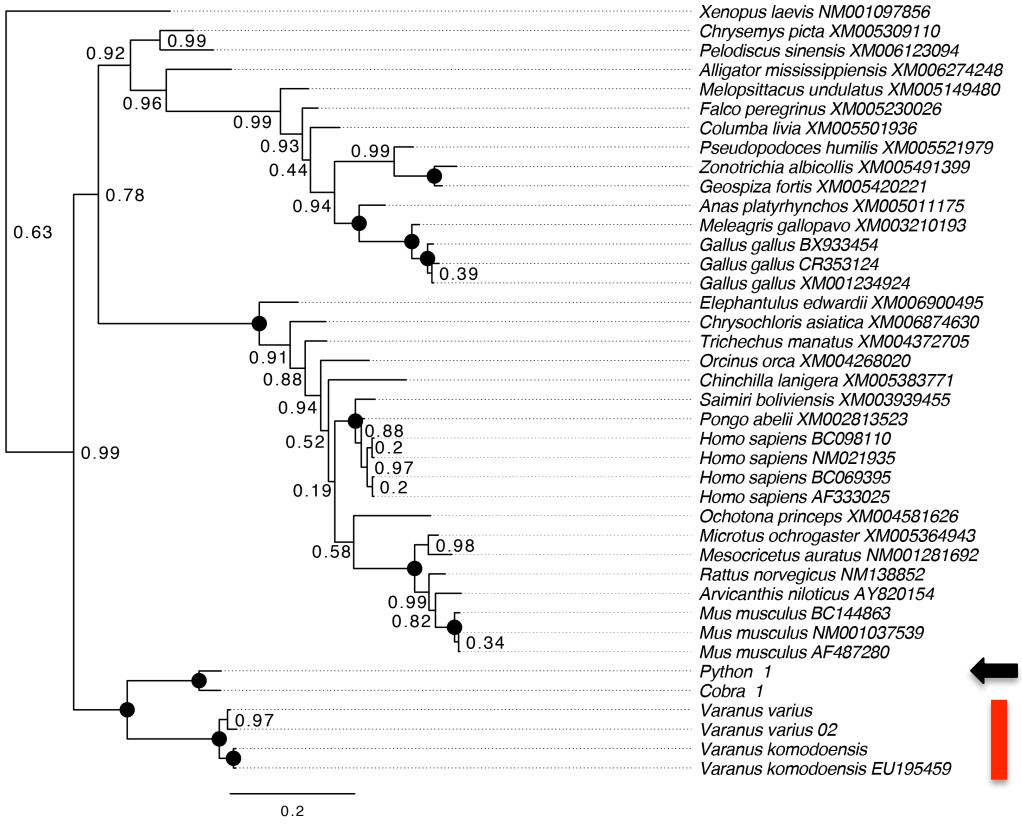
Supplementary Figure 1. Bayesian phylogenetic tree of 5’-nucleotidase. Red lines indicate toxin genes from venomous species; black arrow indicates position of python ortholog to venom toxins. “Cobra” refers to the king cobra (*Ophiophagus Hannah*). Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Acetylcholinesterase



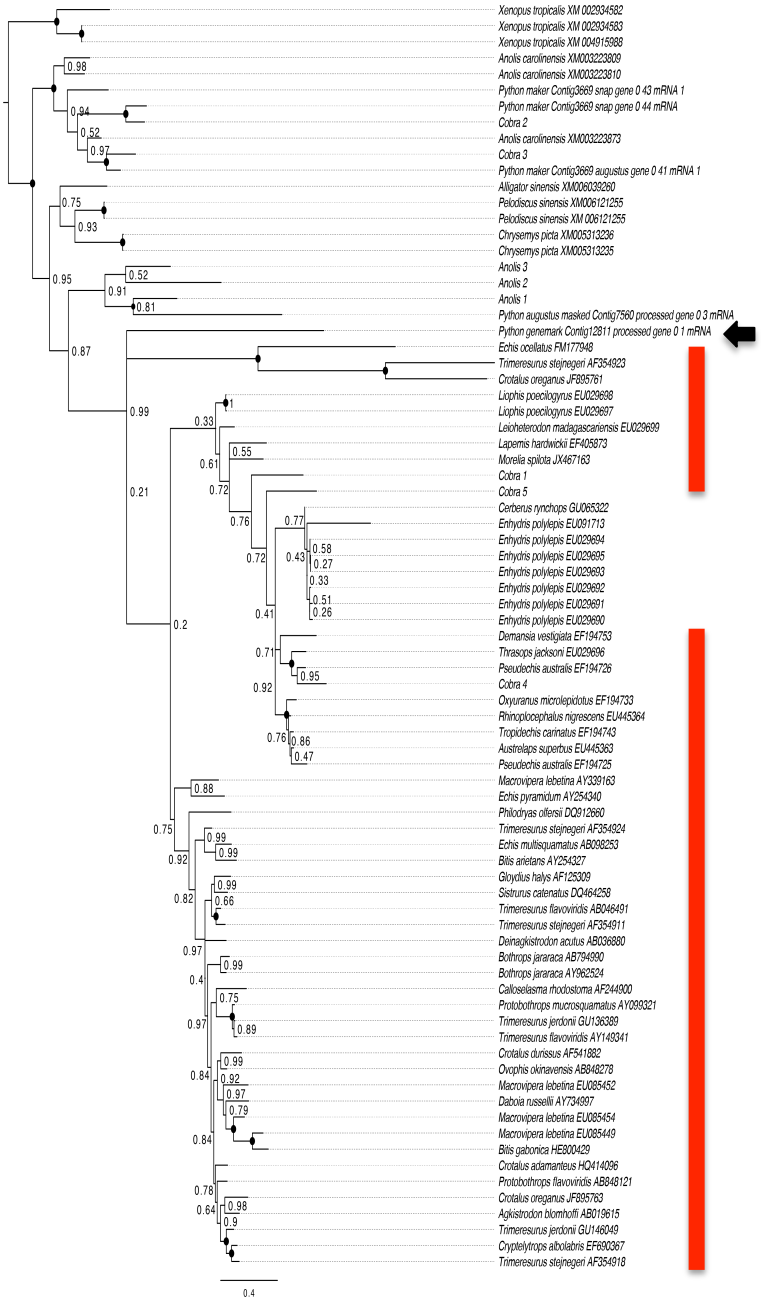
Supplementary Figure 2. Bayesian phylogenetic tree of Acetylcholinesterase. Red arrow indicate toxin genes from venomous species; black arrow indicates position of python ortholog to venom toxins. “Cobra” refers to the king cobra (*Ophiophagus hannah*). Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

AVIToxin



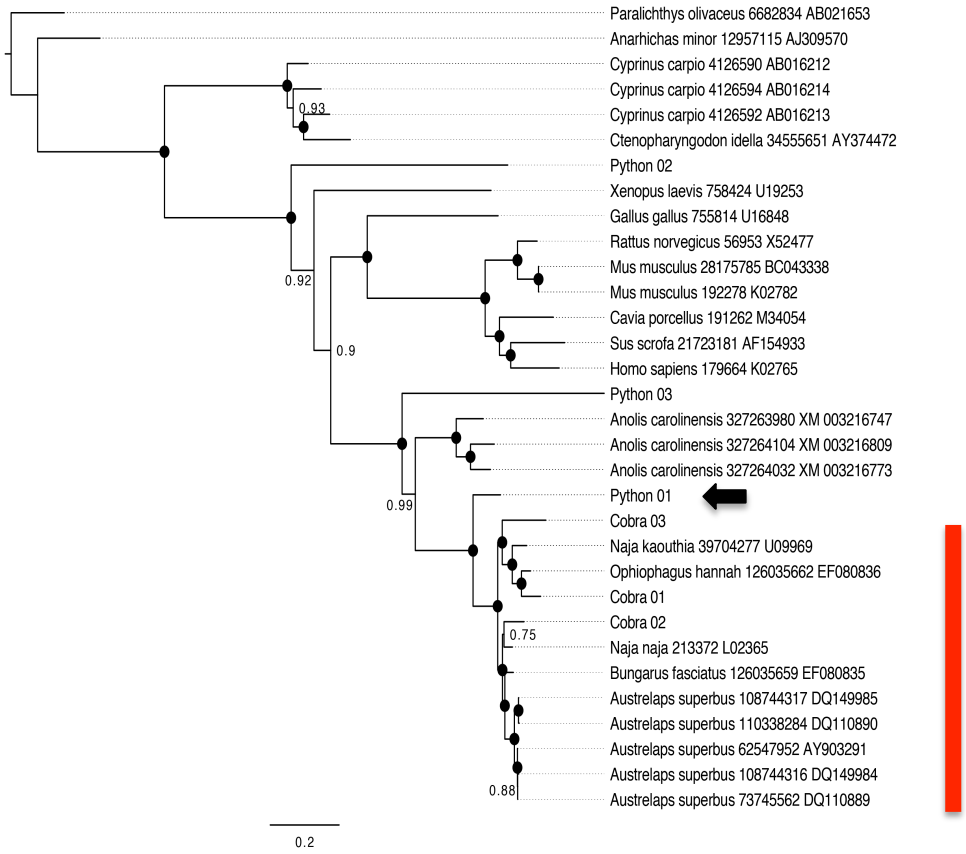
Supplementary Figure 3. Bayesian phylogenetic tree of AVIToxin. Red lines indicate toxin genes from venomous species; black arrow indicates position of python ortholog to venom toxins. “Cobra” refers to the king cobra (*Ophiophagus Hannah*). Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

C-type lectin



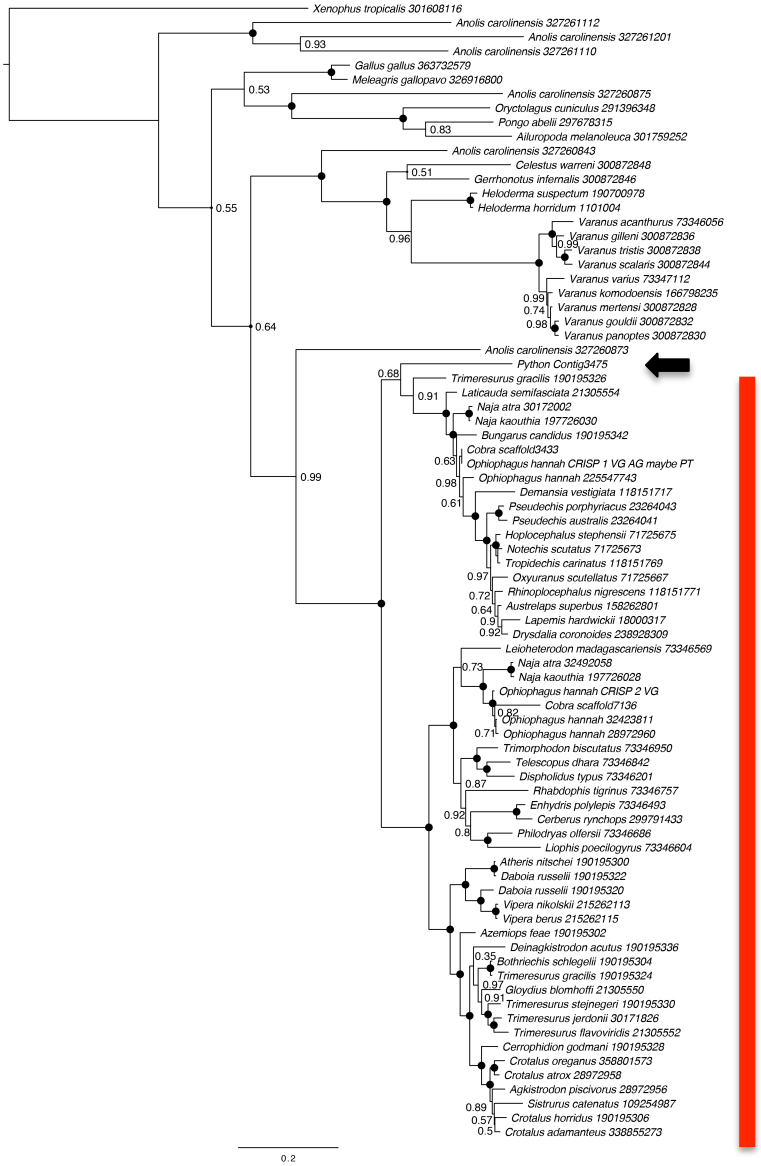
Supplementary Figure 4. Bayesian phylogenetic tree of C-type lectin. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Cobra Venom



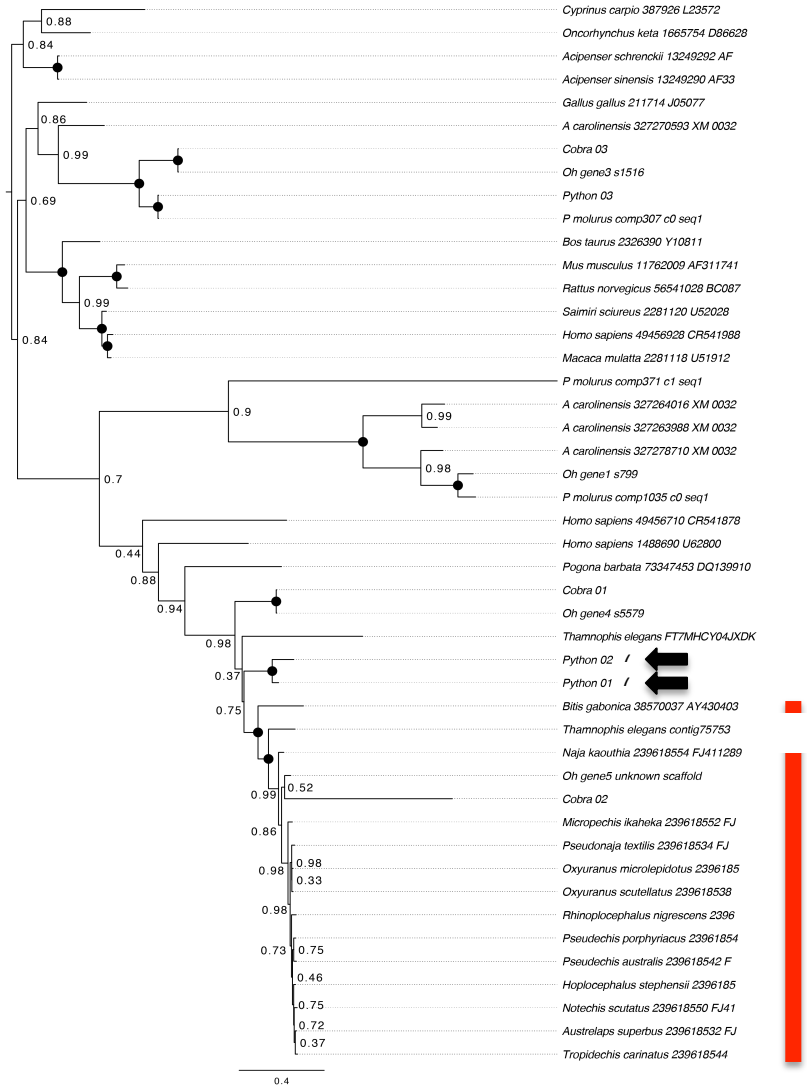
Supplementary Figure 5. Bayesian phylogenetic tree of Cobra Venom Factor. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

CRISp



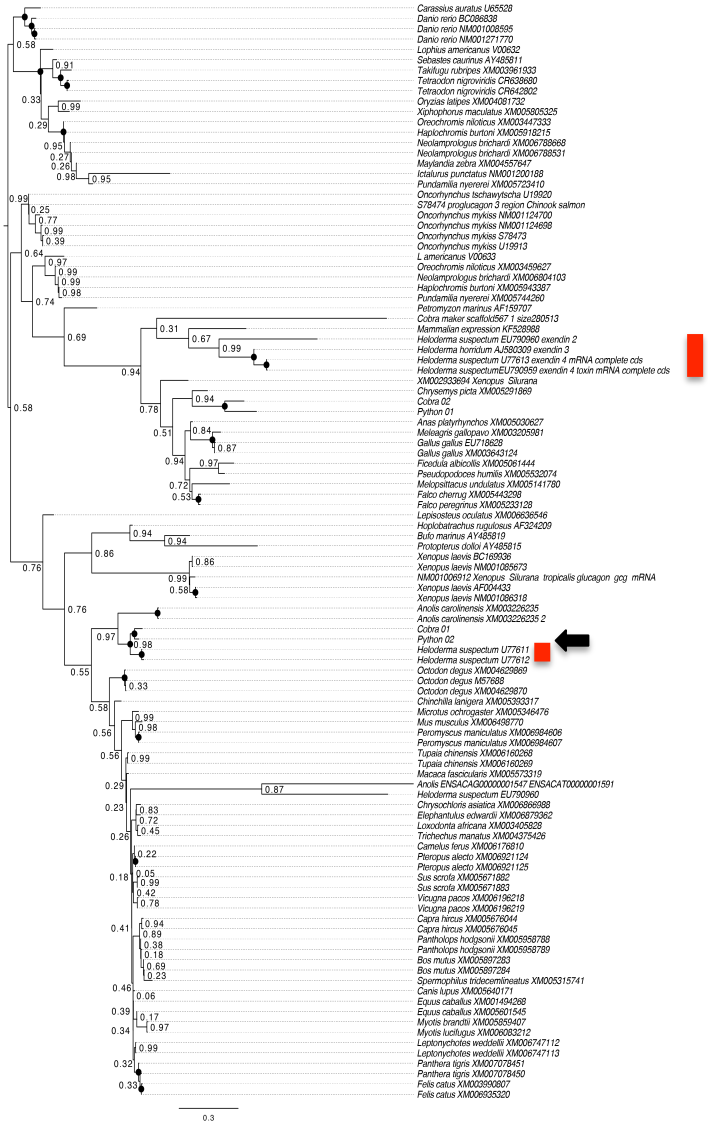
Supplementary Figure 6. Bayesian phylogenetic tree of CRISp. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Cystatin



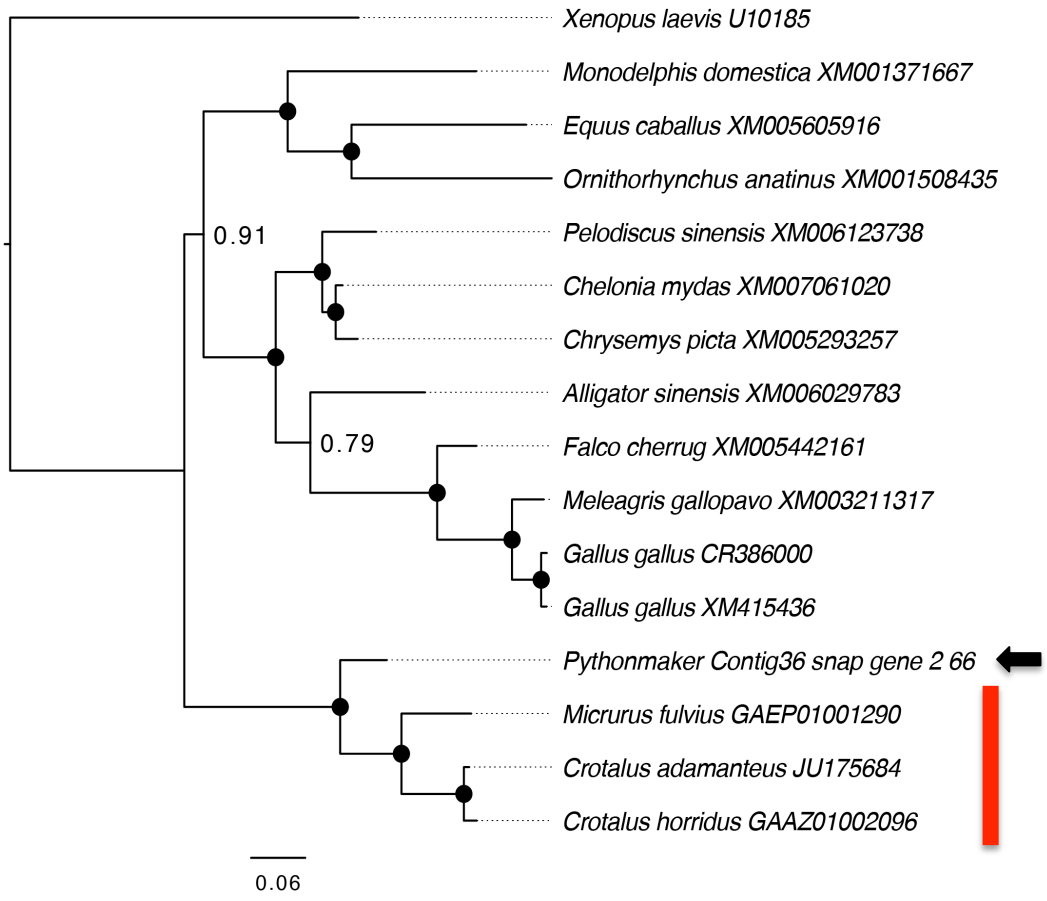
Supplementary Figure 7. Bayesian phylogenetic tree of Cystatin. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Exendin



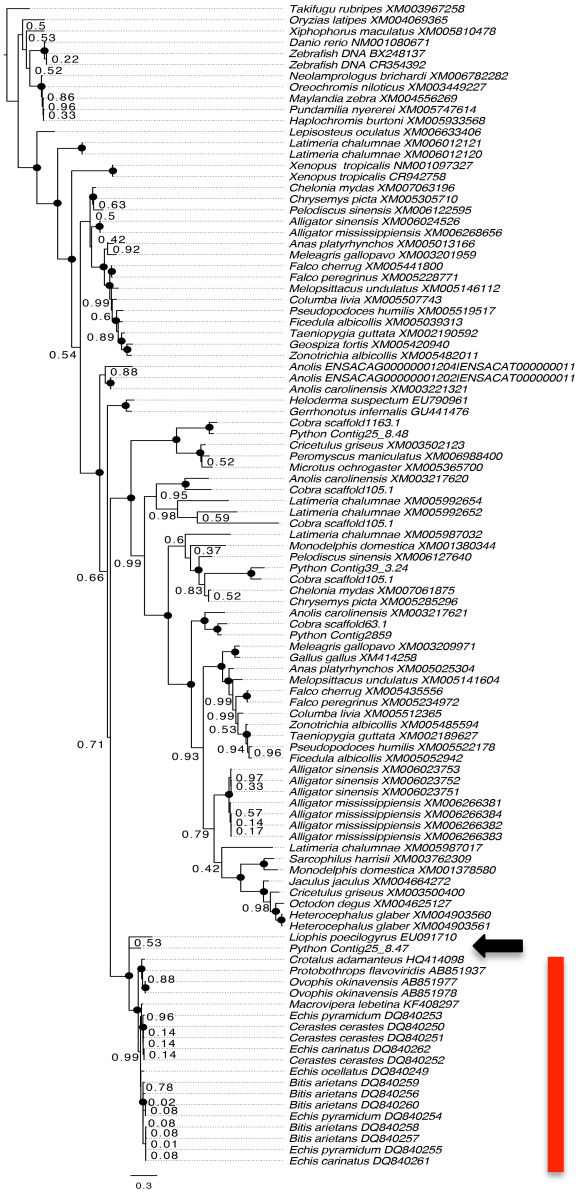
Supplementary Figure 8. Bayesian phylogenetic tree of Exendin. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Exonuclea



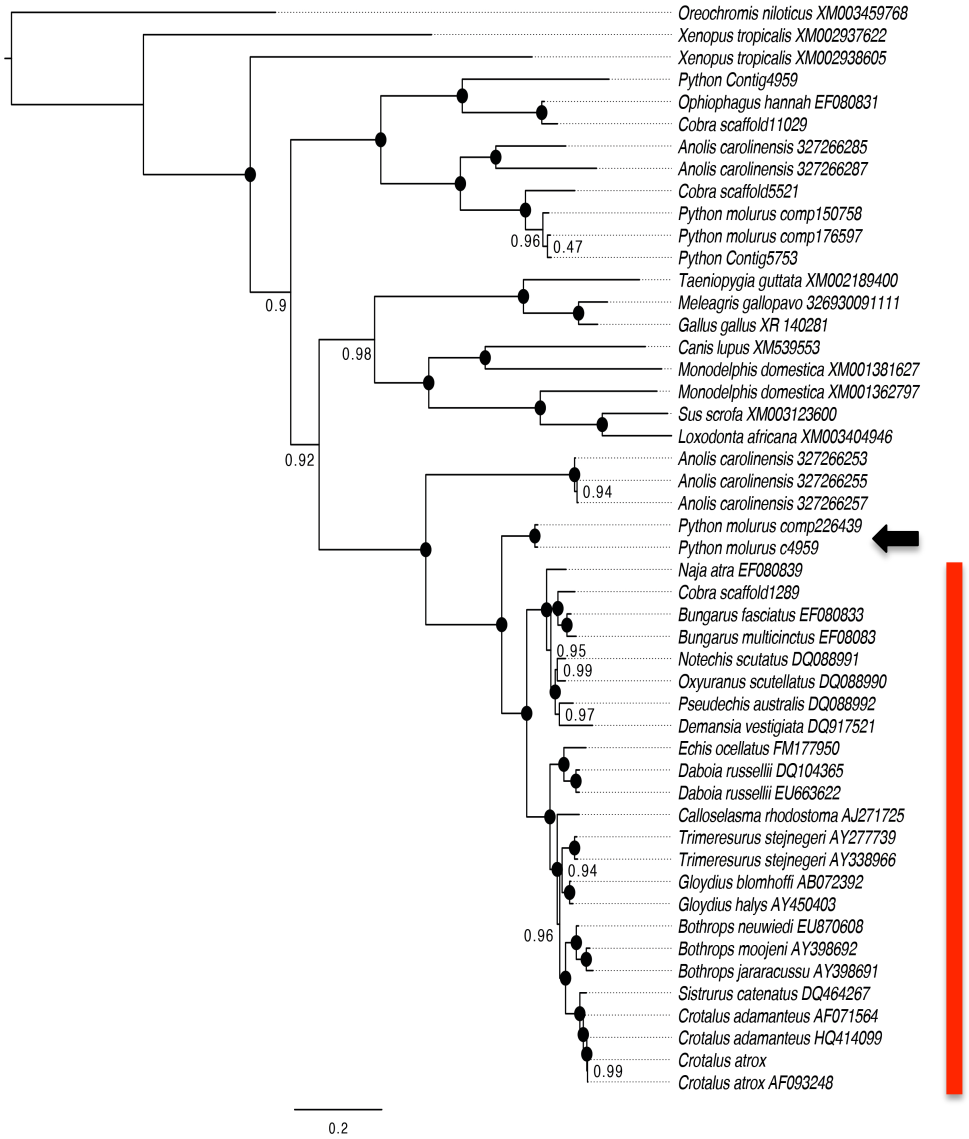
Supplementary Figure 9. Bayesian phylogenetic tree of Exonuclease. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Hyaluronidase

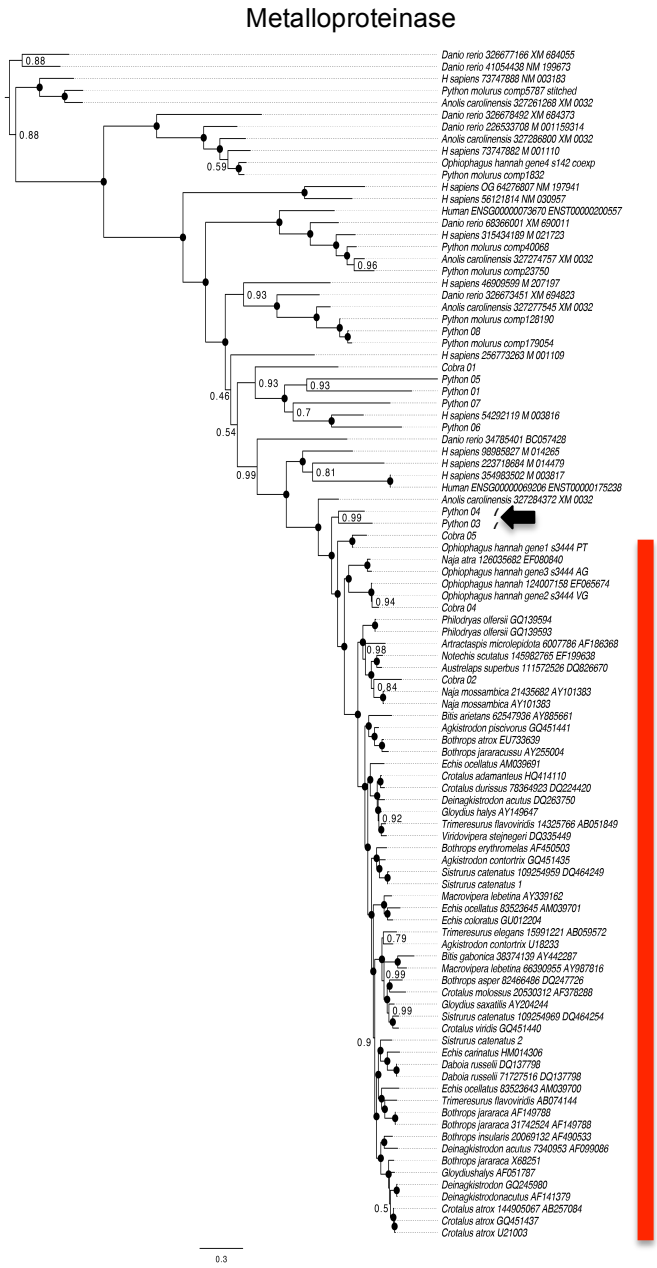


Supplementary Figure 10. Bayesian phylogenetic tree of Hyaluronidase. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

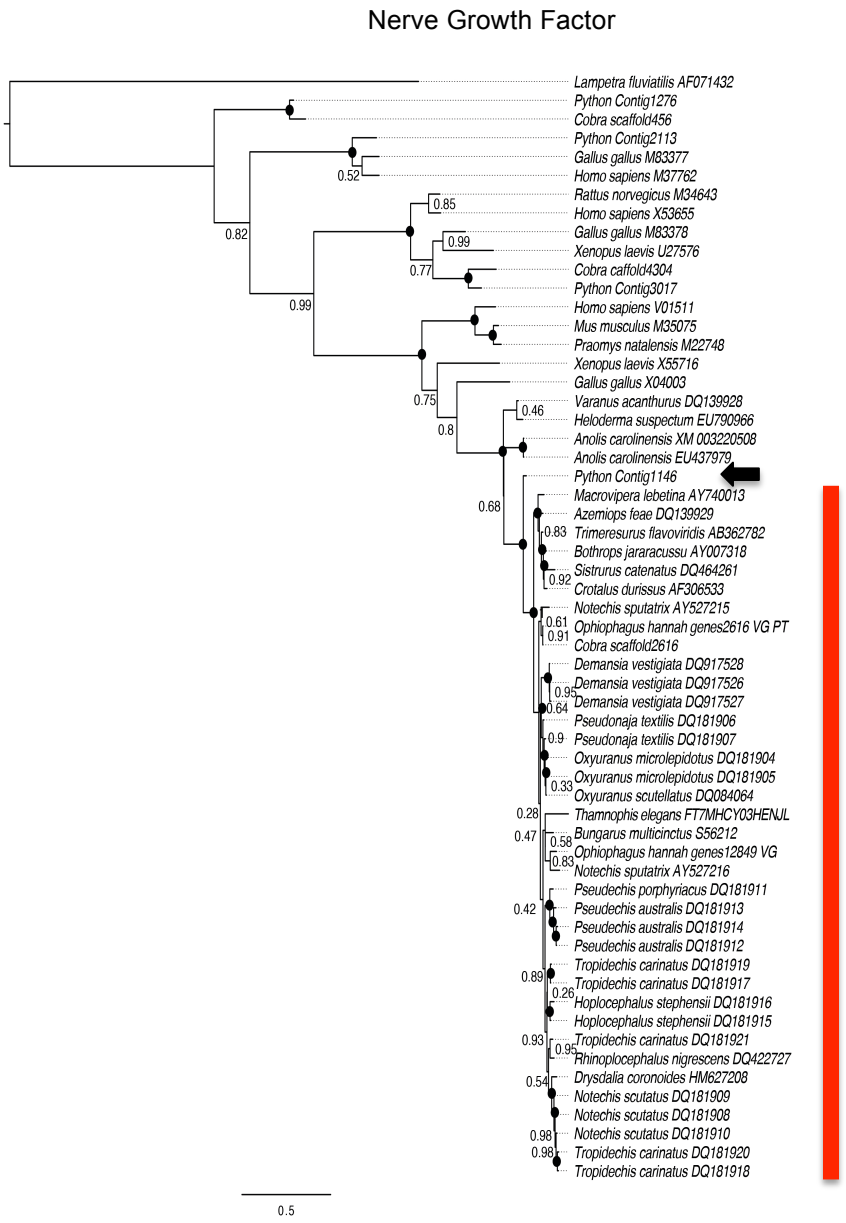
L-Amino Acid Oxidase



Supplementary Figure 11. Bayesian phylogenetic tree of L-Amino Acid Oxidase. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

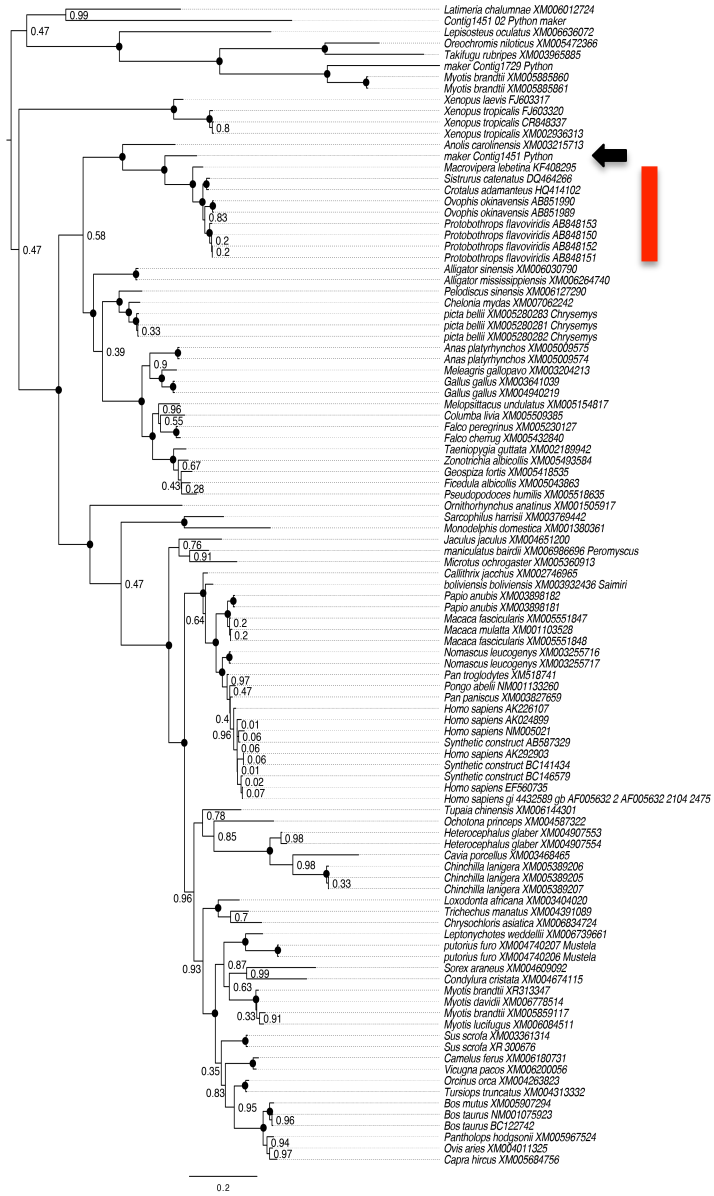


Supplementary Figure 12. Bayesian phylogenetic tree of Metalloproteinase. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

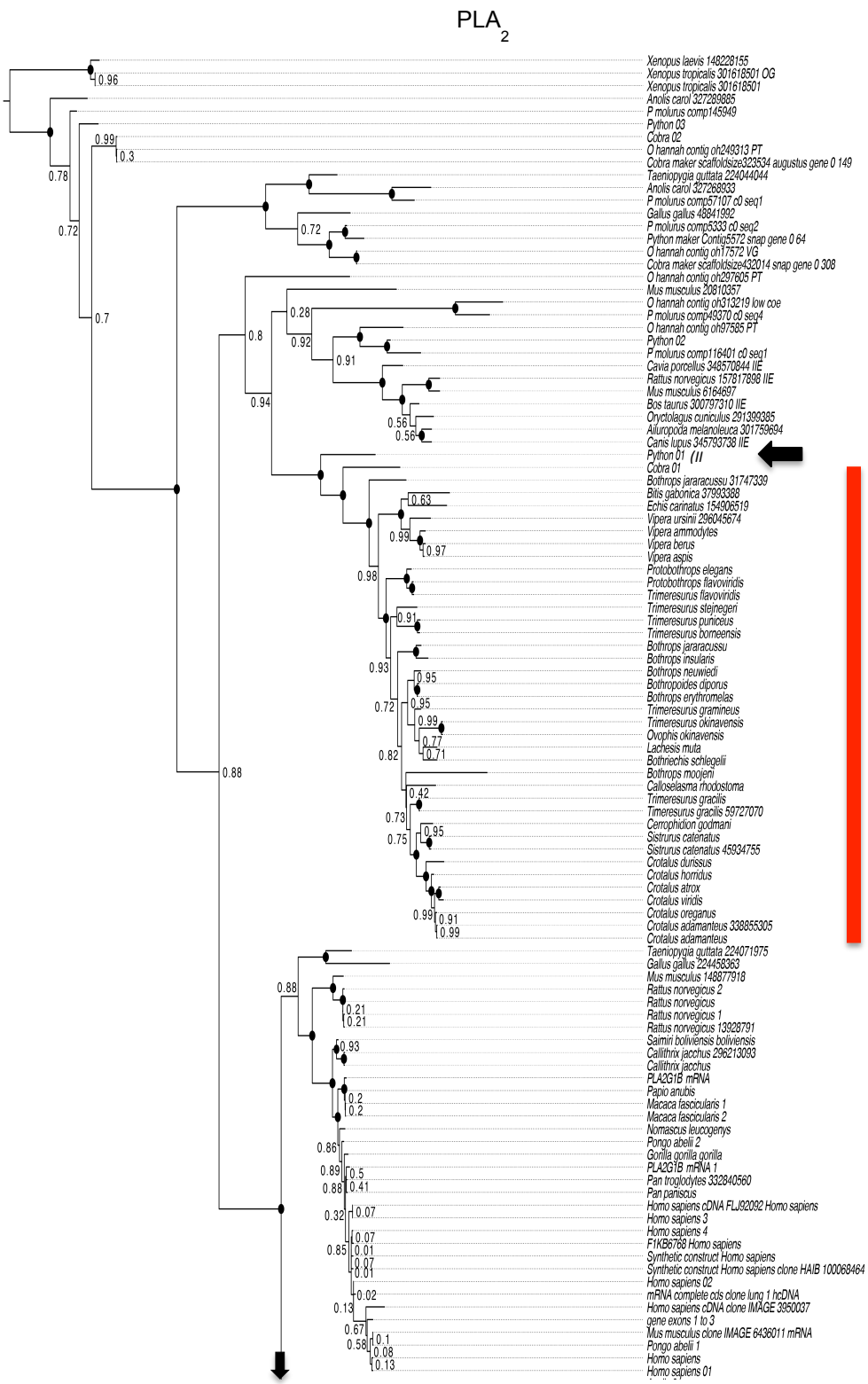


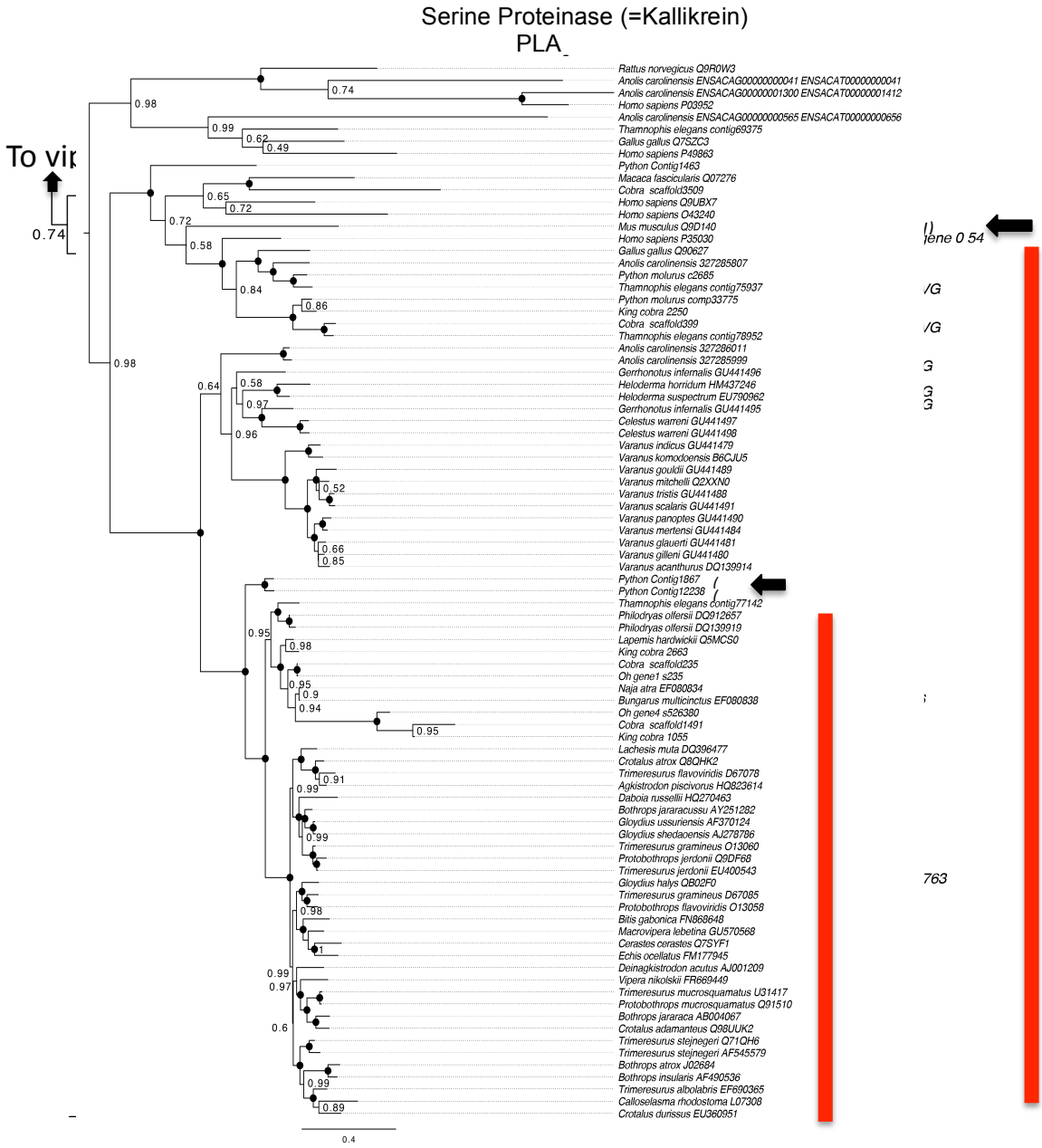
Supplementary Figure 13. Bayesian phylogenetic tree of Nerve Growth Factor. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Phosphodiesterase



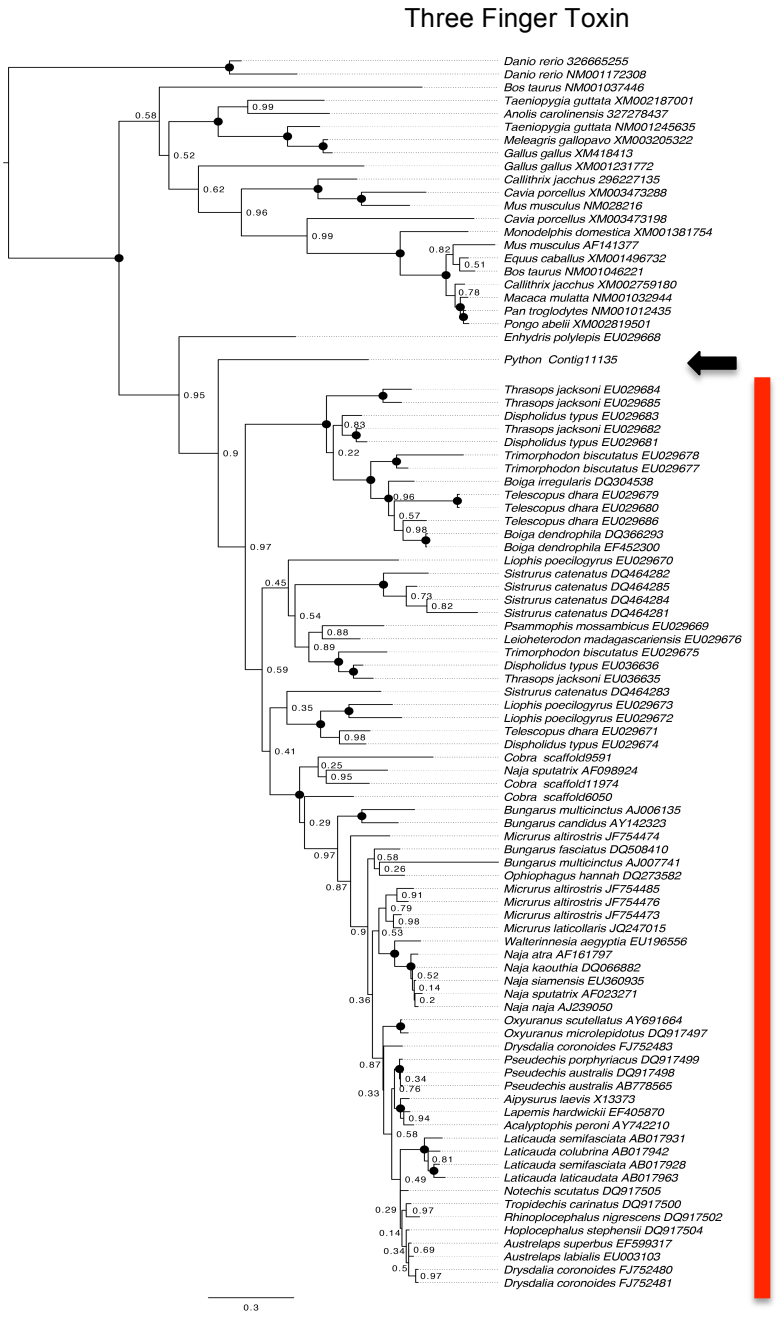
Supplementary Figure 14. Bayesian phylogenetic tree of Phosphodiesterase. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.


PLA2
Xenopus laevis 148228155
Xenopus tropicalis 301618501 OG
Xenopus tropicalis 301618501
Anolis carol 327269865
P molurus comp145949
Python 03
Cobra 02
O hannah contig oh249313 PT
Cobra maker scaffoldsize323534 augustus gene 0 149
Taeniopygia guttata 224044044
Anolis carol 327268933
P molurus comp57107 c0 seq1
Gallus gallus 48841992
P molurus comp5333 c0 seq2
Python maker Contig5572 snap gene 0 64
O hannah contig oh17572 VG
Cobra maker scaffoldsize432014 snap gene 0 308
O hannah contig oh297605 PT
Mus musculus 20810357
O hannah contig oh313219 low coe
P molurus comp49370 c0 seq4
O hannah contig oh97585 PT
Python 02
P molurus comp116401 c0 seq1
Cavia porcellus 348570844 IIE
Rattus norvegicus 157817898 IIE
Mus musculus 6164697
Bos taurus 300797310 IIE
Oryctolagus cuniculus 291399385
Ailuropoda melanoleuca 301759694
Canis lupus 345793738 IIE
Python 01 (II)
Cobra 01
Bothrops jararacussu 31747339
Bitis gabonica 37993388
Echis carinatus 154906519
Vipera ursinii 296045674
Vipera ammodytes
Vipera berus
Vipera aspis
Protobothrops elegans
Protobothrops flavoviridis
Trimeresurus flavoviridis
Trimeresurus stejnegeri
Trimeresurus puniceus
Trimeresurus borneensis
Bothrops jararacussu
Bothrops insularis
Bothrops neuwiedi
Bothropoides diporus
Bothrops erythromelas
Trimeresurus gramineus
Trimeresurus okinavensis
Ovophis okinavensis
Lachesis muta
Bothriechis schlegelii
Bothrops moojeni
Calloselasma rhodostoma
Trimeresurus gracilis
Trimeresurus gracilis 59727070
Cerrophidion godmani
Sistrurus catenatus
Sistrurus catenatus 45934755
Crotalus durissus
Crotalus horridus
Crotalus atrox
Crotalus viridis
Crotalus oreganus
Crotalus adamanteus 338855305
Crotalus adamanteus
Taeniopygia guttata 224071975
Gallus gallus 224458363
Mus musculus 148877918
Rattus norvegicus 2
Rattus norvegicus
Rattus norvegicus 1
Rattus norvegicus 13928791
Saimiri boliviensis boliviensis
Callithrix jacchus 296213093
Callithrix jacchus
PLA2G1B mRNA
Papio anubis
Macaca fascicularis 1
Macaca fascicularis 2
Nomascus leucogenys
Pongo abelii 2
Gorilla gorilla gorilla
PLA2G1B mRNA 1
Pan troglodytes 332840560
Pan paniscus
Homo sapiens cDNA FLJ92092 Homo sapiens
Homo sapiens 3
Homo sapiens 4
F1KB6768 Homo sapiens
Synthetic construct Homo sapiens
Synthetic construct Homo sapiens clone HAIB 100068464
Homo sapiens 02
mRNA complete cds clone lung 1 hcDNA
Homo sapiens cDNA clone IMAGE 3950037
gene exons 1 to 3
Mus musculus clone IMAGE 6436011 mRNA
Pongo abelii 1
Homo sapiens
Homo sapiens 01
To elapids



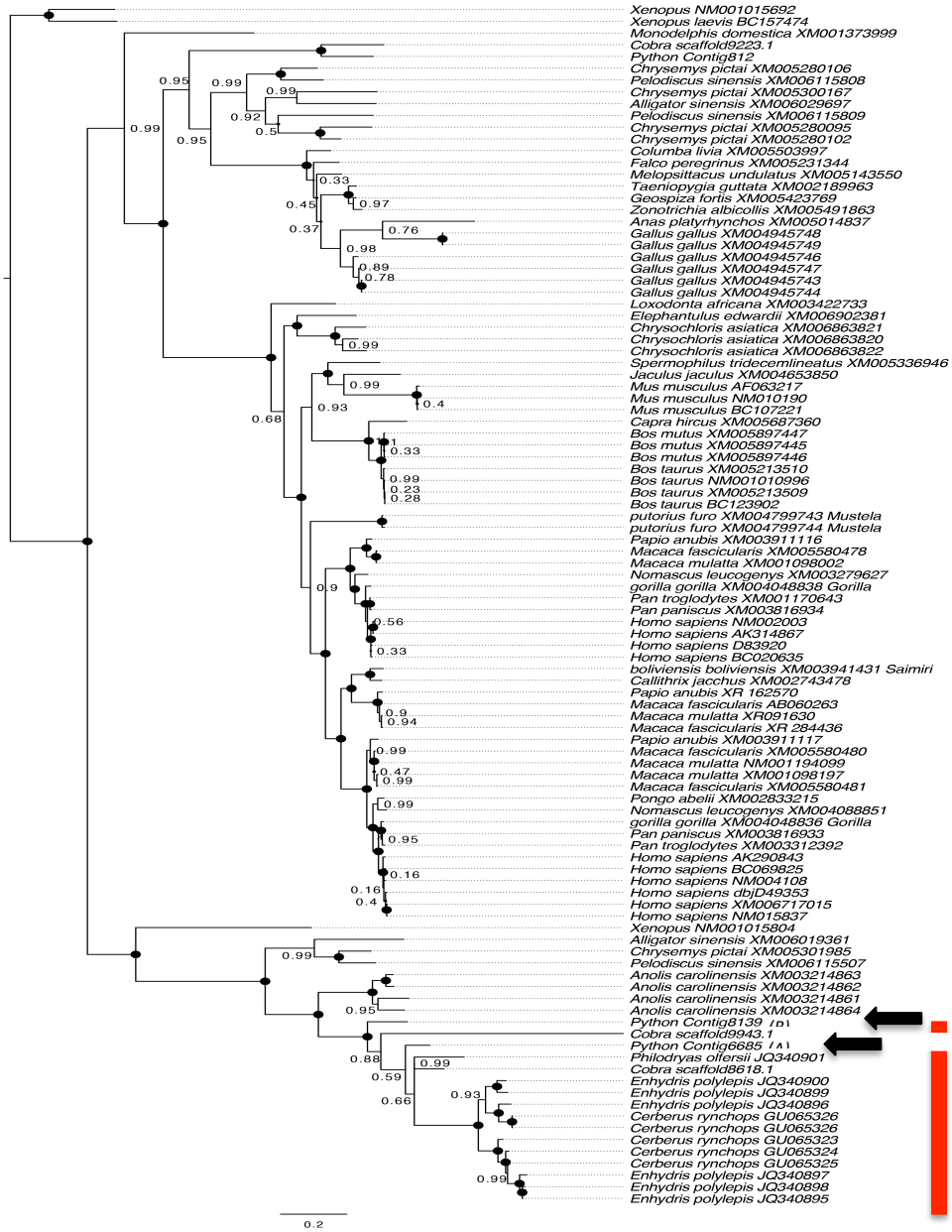
Supplementary Figure 15. Bayesian phylogenetic tree of PLA_2 . Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Supplementary Figure 16. Bayesian phylogenetic tree of Serine Proteinase (=Kallikrein). Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

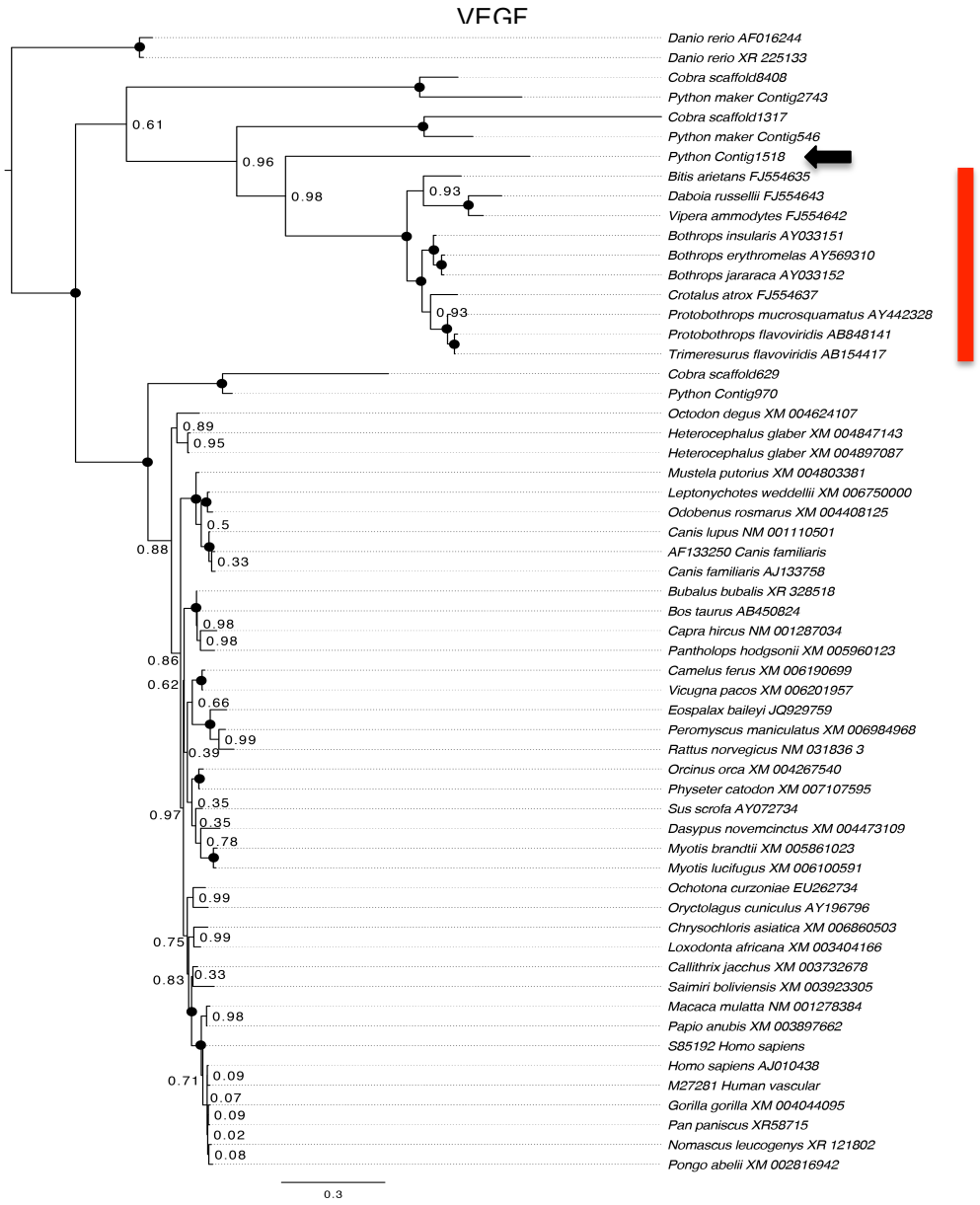


Supplementary Figure 17. Bayesian phylogenetic tree of Three Finger Toxin. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Veficolin



Supplementary Figure 18. Bayesian phylogenetic tree of Veficolin. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.



Supplementary Figure 19. Bayesian phylogenetic tree of VEGF. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Vespryn

Supplementary Figure 20. Bayesian phylogenetic tree of Vespryn. Black arrow indicates position of the python, while red dash indicates the venomous species. Numbers at nodes represent posterior support values. Black circles indicate posterior support = 1.0.

Supplemental Table 3. Partitioning strategy and models of nucleotide evolution used in this study.

Gene	Partitions	1st codon	2nd codon
Three Finger Toxin	1, 2, 3	GTR+G	HKY+G
5'-nucleotidase	1+2; 3	GTR+I+G	GTR+I+G
Acetyloniesterase	1+2; 3	GTR+G	GTR+G
AVIToxin	1+2; 3	HKY+G	HKY+G
BPP/Natriuretic Peptide	1+2; 3	GTR+G	GTR+G
C-type lectin	1+2; 3	SYM+G	SYM+G
Cobra Venom Factor	1+2; 3	HKY+G	HKY+G
CRISp	1, 2, 3	GTR+I+G	GTR+I+G
Crotamine/Crotasine	1+2; 3	HKY+G	HKY+G
Cystatin	1+2; 3	HKY+G	HKY+G
Exendin	1+2; 3	JK	JK
Exonuclease	1, 2, 3	JK	GTR
Hyaluronidase	1+2; 3	JK+G	JK+G
L-Amino Acid Oxidase	1+2; 3	GTR+I+G	GTR+I+G
Metalloproteinase	1, 2, 3	HKY+I+G	HKY+G
Nerve Growth Factor	1+2; 3	GTR+I+G	GTR+I+G

Phosphodiesterase	1+2; 3	JK+G		JK+G
				SYM+I+
PLA2	1, 2, 3	SYM+I+G	G	
Sarafotoxin	1+2; 3	K80+G		K80+G
				GTR+I+
Serine proteinase	1, 2, 3	GTR+I+G	G	
Veficolin	1+2; 3	GTR+G		GTR+G
VEGF	1+2; 3	SYM+G		SYM+G
				GTR+I+
Vespryn	1+2; 3	GTR+I+G	G	
Waprin	1+2; 3	JK+G		JK+G

Supplementary Table 4. Summary of raw RNAseq reads and mapped reads for each tissue used for gene expression analysis.

Tissue	Raw Reads*	Mapped Reads*
Blood	25,959,588	7,906,944
Ovary	20,820,092	11,642,601
Stomach	19,264,312	9,372,517
Muscle	20,062,380	13,427,461
Brain	51,687,904	20,314,823
Heart	17,692,718	4,883,094
Spleen	26,700,204	12,849,006
Kidney	42,635,186	13,241,991
Rictal Gland	32,268,332	17,060,542
Liver	11,327,681	2,564,670
Testes	11,520,424	4,901,337
Small Intestine	55,930,617	18,978,054

*paired-end reads counted as two separate reads.

Supplementary Table 5. Fisher's exact test for each expression level bin for all genes and all venom homologs inferred. Bolded P-values indicate significant test ($p<0.05$).

Bin	All genes	Venom homologs	P-value
Bin 0	111,833	137	0.3170
Bin 0 - 1	39,263	59	0.0300
Bin 1 - 10	66,843	81	0.5600
Bin 10 -100	68,738	50	0.0002
Bin 100 - 1,000	16,542	17	0.8120
Bin 1,000 - 10,000	1,288	4	0.0620
Bin >10,000	89	0	1.0000
TOTAL	304,596	348	

Supplementary Table 6. Fisher's exact test for each expression level bin for all genes and venom homologs inferred using gene phylogenies. Bold P-values indicate significant tests ($p<0.05$).

Bin	All genes	Venom homologs	P-value
Bin 0	111,833	119	0.3080
Bin 0 - 1	39,263	47	0.1670
Bin 1 - 10	66,843	67	0.8890
Bin 10 -100	68,738	46	0.0023
Bin 100 - 1,000	16,542	17	0.7790
Bin 1,000 - 10,000	1,288	4	0.0400
Bin >10,000	89	0	1.0000
TOTAL	304,596	300	

Supplementary Table 7. Fisher's exact test for each expression level bin for all genes and venom homologs with known cytotoxic activity. Bolded P-values indicate significant tests ($p<0.05$).

Bin	All genes	Venom homologs	P-value
Bin 0	111,833	39	0.0070
Bin 0 - 1	39,263	19	0.0130
Bin 1 - 10	66,843	21	0.5100
Bin 10 -100	68,738	3	2.471E-06
Bin 100 - 1,000	16,542	2	0.3310
Bin 1,000 - 10,000	1,288	0	1.0000
Bin >10,000	89	0	1.0000
TOTAL	304,596	84	

References

Ahmed M, Latif N, Khan R, Ahmad A, Rocha J, Mazzanti C, Bagatini M, Morsch V and Schetinger M. 2012. Enzymatic and biochemical characterization of *Bungarus sindanus* snake venom acetylcholinesterase. Journal of Venomous Animals and Toxins including Tropical Diseases 18: 236-243.

Alape-Girón A, Persson B, Cederlund E, Flores-Díaz M, Gutiérrez JM, Thelestam M, Bergman T and JoÈrnvall H. 1999. Elapid venom toxins: multiple recruitments of ancient scaffolds. European Journal of Biochemistry 259: 225-234.

Amaral A. 1929. IV. Phylogeny of the rattlesnakes. Bulletin of the Antivenin Institute of America 3: 6-8.

Amman BR and Bradley RD. 2004. Molecular evolution in Baiomys (Rodentia : Sigmodontinae): Evidence for a genetic subdivision in B-musculus. J Mammal 85: 162-166.

Anderson CG and Greenbaum E. 2012. Phylogeography of Northern Populations of the Black-Tailed Rattlesnake (*Crotalus molossus* Baird and Girard, 1853), with the Revalidation of *C. ornatus* Hallowell, 1854. Herpetol Monogr: 19-57.

Arevalo E, Davis SK and Sites JW. 1994. Mitochondrial-DNA Sequence Divergence and Phylogenetic-Relationships among 8 Chromosome Races of the Sceloporus-Grammicus Complex (Phrynosomatidae) in Central Mexico. Systematic Biology 43: 387-418.

Armstrong BL and Murphy JB. 1979. The natural history of Mexican rattlesnakes. Lawrence: Museum of Natural History, University of Kansas.

Bandelt HJ, Forster P and Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16: 37-48.

Bates HW. 1862. XXXII. Contributions to an Insect Fauna of the Amazon Valley. Lepidoptera: Heliconidæ. Transactions of the Linnean Society of London 23: 495-566.

Bergsten J. 2005. A review of long-branch attraction. Cladistics 21: 163-193.

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

Brattstrom BH. 1964. Evolution of the pit vipers. Transactions of the San Diego Society of Natural History 13: 185–268.

Brenes O, Munoz E, Roldan-Rodriguez R and Diaz C. 2010. Cell death induced by *Bothrops asper* snake venom metalloproteinase on endothelial and other cell lines. Experimental and molecular pathology 88: 424-432.

Brodie III E and Janzen F. 1995. Experimental studies of coral snake mimicry: generalized avoidance of ringed snake patterns by free-ranging avian predators. Functional Ecology: 186-190.

Brodie III ED. 1993. Differential avoidance of coral snake banded patterns by free-ranging avian predators in Costa Rica. Evolution: 227-235.

Brust A, Sunagar K, Undheim EA, et al. 2013. Differential evolution and neofunctionalization of snake venom metalloprotease domains. Molecular & Cellular Proteomics 12: 651-663.

Bryson RW, de Oca ANM and Velasco JR. 2008. Phylogenetic position of *Porthidium hespere* (Viperidae : Crotalinae) and phylogeography of arid-adapted hognosed pitvipers based on mitochondrial DNA. Copeia: 172-178.

Bryson RW, Murphy RW, Graham MR, Lathrop A and Lazcano D. 2011a. Ephemeral Pleistocene woodlands connect the dots for highland rattlesnakes of the *Crotalus intermedius* group. J Biogeogr 38: 2299-2310.

Bryson RW, Murphy RW, Lathrop A and Lazcano-Villareal D. 2011b. Evolutionary drivers of phylogeographical diversity in the highlands of Mexico: a case study of the *Crotalus triseriatus* species group of montane rattlesnakes. J Biogeogr 38: 697-710.

Burbrink FT and Castoe TA. 2009. Molecular phylogeography of snakes. In: Seigel R, Mullin S, editors. Snakes: Ecology and Conservation. Ithaca, New York: Cornall University Press. p. 38-77.

Cadle JE. 1992. Phylogenetic relationships among vipers: immunological evidence. In: Campbell JA, Brodie Jr. ED, editors. Biology of the Pitvipers. Tyler, Texas: Selva. p. 41-48.

Campbell JA and Flores-Villela O. 2008. A new long-tailed rattlesnake (Viperidae) from Guerrero, Mexico. Herpetologica 64: 246-257.

Campbell JA and Lamar WW. 2004a. The venomous reptiles of the Western Hemisphere. Ithaca: Comstock Pub. Associates.

Campbell JA and Lamar WW. 2004b. The Venomous Reptiles of the Western Hemisphere. Ithaca, NY: Cornell University Press.

Casewell NR, Harrison RA, Wüster W and Wagstaff SC. 2009. Comparative venom gland transcriptome surveys of the saw-scaled vipers (Viperidae: *Echis*) reveal substantial intra-family gene diversity and novel venom transcripts. BMC Genomics 10: 564.

Casewell NR, Huttley GA and Wuster W. 2012. Dynamic evolution of venom proteins in squamate reptiles. Nature communications 3: 1066.

Casewell NR, Wagstaff SC, Wüster W, et al. 2014. Medically important differences in snake venom composition are dictated by distinct postgenomic mechanisms. Proceedings of the National Academy of Sciences.

Casewell NR, Wuster W, Vonk FJ, Harrison RA and Fry BG. 2013. Complex cocktails: the evolutionary novelty of venoms. Trends in ecology & evolution 28: 219-229.

Castoe TA, de Koning AP, Hall KT, et al. 2013. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. Proceedings of the National Academy of Sciences USA 110: 20645-20650.

Castoe TA and Parkinson CL. 2006. Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). Mol Phylogenet Evol 39: 91-110.

Castoe TA, Smith EN, Brown RM and Parkinson CL. 2007a. 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). Zoological Journal of the Linnean Society 151: 809-831.

Castoe TA, Spencer CL and Parkinson CL. 2007b. Phylogeographic structure and historical demography of the western diamondback rattlesnake (Crotalus atrox): A perspective on North American desert biogeography. Mol Phylogenet Evol 42: 193-212.

Castoe TA, Streicher JW, Meik JM, et al. 2012. Thousands of microsatellite loci from the venomous coralsnake Micrurus fulvius and variability of select loci across populations and related species. Molecular Ecology Resources 12: 1105-1113.

Catchen J, Hohenlohe PA, Bassham S, Amores A and Cresko WA. 2013. Stacks: an analysis tool set for population genomics. Molecylar Ecology 22: 3124-3140.

Catchen JM, Amores A, Hohenlohe P, Cresko W and Postlethwait JH. 2011. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. G3-Genes Genom Genet 1: 171-182.

Clement M, Posada D and Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. Mol Ecol 9: 1657-1659.

Correa-Netto C, Junqueira-de-Azevedo IDM, Silva DA, et al. 2011. Snake venomics and venom gland transcriptomic analysis of Brazilian coral snakes, *Micrurus altirostris* and *M. corallinus*. Journal of Proteomics 74: 1795-1809.

Cousin X, Bon S, Massoulié J and Bon C. 1998. Identification of a Novel Type of Alternatively Spliced Exon from the Acetylcholinesterase Gene of *Bungarus fasciatus.* Molecular forms of Acetylcholinesterase in the snake liver and muscle. Journal of Biological Chemistry 273: 9812-9820.

Daza JM, Castoe TA and Parkinson CL. 2010. Using regional comparative phylogeographic data from snake lineages to infer historical processes in Middle America. Ecography 33: 343-354.

de Roodt AR, Paniagua-Solis JF, Dolab JA, Estévez-Ramiréz J, Ramos-Cerrillo B, Litwin S, Dokmetjian JC and Alagón A. 2004. Effectiveness of two common antivenoms for North, Central, and South American Micrurus envenomations. Clinical Toxicology 42: 171-178.

Deshimaru M, Ogawa T, Nakashima K-i, et al. 1996. Accelerated evolution of crotalinae snake venom gland serine proteases. FEBS Letters 397: 83-88.

Devitt TJ. 2006. Phylogeography of the Western Lyresnake (Trimorphodon biscutatus): testing aridland biogeographical hypotheses across the Nearctic-Neotropical transition. Mol Ecol 15: 4387-4407.

Doley R, Mackessy SP and Kini RM. 2009. Role of accelerated segment switch in exons to alter targeting (ASSET) in the molecular evolution of snake venom proteins. BMC Evolutionary Biology 9: 146.

Doley R, Tram NN, Reza MA and Kini RM. 2008. Unusual accelerated rate of deletions and insertions in toxin genes in the venom glands of the pygmy copperhead (*Austrelaps labialis*) from Kangaroo island. BMC Evolutionary Biology 8: 70.

Douglas ME, Douglas MR, Schuett GW and Porras LW. 2006. Evolution of rattlesnakes (Viperidae; *Crotalus*) in the warm deserts of western North America shaped by Neogene vicariance and Quaternary climate change. Molecular Ecology 15: 3353-3374.

Douglas ME, Schuett GW, Porras LW and T. HA. 2002. Phylogeography of the western rattlesnake (*Crotalus viridis*) complex, with emphasis on the Colorado Plateau. In: Schuett GW, Hoggren M, Douglas ME, Greene HW, editors. Biology of the vipers. Eagle Mountain, Utah: Eagle Mountain Publishing. p. 11-50.

Drummond AJ and Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. Bmc Evolutionary Biology 7. doi: Artn 214

Ducancel F, Matre V, Dupont C, Lajeunesse E, Wollberg Z, Bdolah A, Kochva E, Boulain J-C and Menez A. 1993. Cloning and sequence analysis of cDNAs encoding precursors of sarafotoxins. Evidence for an unusual "rosary-type" organization. Journal of Biological Chemistry 268: 3052-3055.

Dunn ER. 1919. Two new crotaline snakes from western Mexico. Proceedings of the Biological Society of Washington: 213–216.

Earl DA and Vonholdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4: 359-361.

Eaton DAR. 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics 30: 1844-1849. doi: Doi 10.1093/Bioinformatics/Btu121

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32: 1792-1797.

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

Felsenstein J. 2004. Inferring Phylogenies. Sunderland, MA: Sinauer Associates.

Flicek P, Amode MR, Barrell D, et al. 2014. Ensembl 2014. Nucleic Acids Research 42: D749-D755.

Foote R and MacMahon JA. 1977. Electrophoretic studies of rattlesnake (*Crotalus* & *Sistrurus*) venom: taxonomic implications. Comparative Biochemistry and Physiology Part B: Comparative Biochemistry 57: 235-241.

Fox JW and Serrano SM. 2008. Exploring snake venom proteomes: multifaceted analyses for complex toxin mixtures. Proteomics 8: 909-920.

Fraser DF. 1973. Variation in the coral snake Micrurus diastema. Copeia: 1-17.

Frey HM, Lange RA, Hall CM, Delgado-Granados H and Carmichael ISE. 2007. A Pliocene ignimbrite flare-up along the Tepic-Zacoalco rift: Evidence for the initial stages of rifting between the Jalisco block (Mexico) and North America. Geol Soc Am Bull 119: 49-64.

Fry BG. 2005. From genome to “venome”: molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. Genome research 15: 403-420.

Fry BG, Casewell NR, Wüster W, Vidal N, Young B and Jackson TN. 2012. The structural and functional diversification of the Toxicofera reptile venom system. Toxicon 60: 434-448.

Fry BG, Roelants K, Champagne DE, et al. 2009. The Toxicogenomic Multiverse: Convergent Recruitment of Proteins Into Animal Venoms. Annual Review of Genomics and Human Genetics 10: 483-511.

Fry BG, Undheim EA, Ali SA, et al. 2013. Squeezers and leaf-cutters: differential diversification and degeneration of the venom system in toxicoferan reptiles. Molecular & cellular proteomics : MCP 12: 1881-1899.

Fry BG, Vidal N, Norman JA, et al. 2006. Early evolution of the venom system in lizards and snakes. Nature 439: 584-588.

Fry BG, Winter K, Norman JA, et al. 2010. Functional and structural diversification of the Anguimorpha lizard venom system. Molecular & Cellular Proteomics 9: 2369-2390.

Fry BG and Wüster W. 2004. Assembling an arsenal: origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences. Molecular Biology and Evolution 21: 870-883.

Fujimi T, Nakajyo T, Nishimura E, Ogura E, Tsuchiya T and Tamiya T. 2003. Molecular evolution and diversification of snake toxin genes, revealed by analysis of intron sequences. Gene 313: 111-118.

Garza-Garcia A, Harris R, Esposito D, Gates PB and Driscoll PC. 2009. Solution structure and phylogenetics of Prod1, a member of the three-finger protein superfamily implicated in salamander limb regeneration. PLoS One 4: e7123.

Githens TS and George ID. 1931. Comparative studies on the venoms of certain rattlesnakes. Bulletin of the Antivenin Institute of America 5: 31-34.

Gloyd HK. 1940. The rattlesnakes, genera *Sistrurus* and *Crotalus*. A study in zoogeography and evolution. Chicago,.

Greene H and McDiarmid R. 2005. Wallace and Savage: Heroes, theories, and venomous snake mimicry. Ecology and evolution in the tropics: a herpetological perspective University of Chicago Press, Chicago: 190-208.

Greene HW. 2000. Snakes: the evolution of mystery in nature: University of California Press.

Greene HW and Cundall D. 2000. Evolutionary biology - Limbless tetrapods and snakes with legs. Science 287: 1939-1941.

Greene HW and McDiarmid RW. 1981. Coral snake mimicry: does it occur. Science 213: 1207-1212.

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

FASTX-Toolkit [Internet]. 2015. Available from: http://hannonlab.cshl.edu/fastx_toolkit/index.html

Hargreaves AD, Swain MT, Hegarty MJ, Logan DW and Mulley JF. 2014. Restriction and recruitment-gene duplication and the origin and evolution of snake venom toxins. Genome Biology and Evolution 6: 2088-2095.

Harper GR and Pfennig DW. 2008. Selection overrides gene flow to break down maladaptive mimicry. Nature 451: 1103-U1106.

Heled J and Drummond AJ. 2010. Bayesian Inference of Species Trees from Multilocus Data. Mol Biol Evol 27: 570-580.

Heyborne WH and Mackessy SP. 2013. Identification and characterization of a taxon-specific three-finger toxin from the venom of the Green Vinesnake (Oxybelis fulgidus; family Colubridae). Biochimie 95: 1923-1932.

Hijmans RJ, Cameron SE, Parra JL, Jones PG and Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25: 1965-1978.

Hillis DM and Bull JJ. 1993. An Empirical-Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis. Syst Biol 42: 182-192.

Hinman KE, Throop HL, Adams KL, Dake AJ, McLauchlan KK and McKone MJ. 1997. Predation by free-ranging birds on partial coral snake mimics: The importance of ring width and color. Evolution 51: 1011-1014.

Holman JA. 2000. Fossil snakes of North America: Origin, Evolution, Distribution, Paleoecology. USA: Indiana University Press.

Huelsenbeck JP and Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17: 754-755.

Ivanov CP and Ivanov OC. 1979. The evolution and ancestors of toxic proteins. Toxicon 17: 205-220.

Ivanov OC. 1981. The evolutionary origin of toxic proteins. Toxicon 19: 171-178.

Jadin RC, Reyes-Velasco J and Smith EN. 2010. Hemipenes of the long-tailed rattlesnakes (serpentes: Viperidae) from Mexico. Phyllomedusa 9: 69-73.

Jiang Y, Li Y, Lee W, Xu X, Zhang Y, Zhao R, Zhang Y and Wang W. 2011. Venom gland transcriptomes of two elapid snakes (Bungarus multicinctus and Naja atra) and evolution of toxin genes. BMC genomics 12: 1.

Junqueira-de-Azevedo IdL and Ho PL. 2002. A survey of gene expression and diversity in the venom glands of the pitviper snake *Bothrops insularis* through the generation of expressed sequence tags (ESTs). Gene 299: 279-291.

Kass RE and Raftery AE. 1995. Bayes Factors. J Am Stat Assoc 90: 773-795.

Kelly CMR, Barker NP, Villet MH and Broadley DG. 2009. Phylogeny, biogeography and classification of the snake superfamily Elapoidea: a rapid radiation in the late Eocene. Cladistics 25: 38-63.

Keogh JS. 1998. Molecular phylogeny of elapid snakes and a consideration of their biogeographic history. Biological journal of the Linnean Society 63: 177-203.

Kini RM and Chinnasamy A. 2010. Nucleotide sequence determines the accelerated rate of point mutations. Toxicon 56: 295-304.

Klauber LM. 1952. Taxonomic studies of the rattlesnakes of mainland Mexico. . Bulletin of the Zoological Society of San Diego 26: 1-143.

Klauber LM. 1956. Rattlesnakes, their habits, life histories, and influence on mankind. Berkeley,: Published for the Zoölogical Society of San Diego by the University of California Press.

Klauber LM. 1972. Rattlesnakes: their habits, life histories, and influence on mankind. Berkeley,: Published for the Zoological Society of San Diego by the University of California Press.

Kochva E. 1978. Oral Glands of the Reptilia. In: Gans C, editor. Biology of the Reptilia, Vol 8. London and New York: Academic Press. p. 43-161.

Kordiš D and Gubenšek F. 2000. Adaptive evolution of animal toxin multigene families. Gene 261: 43-52.

Kreil G. 1995. Hyaluronidases—a group of neglected enzymes. Protein Science 4: 1666-1669.

LaDuc TJ. 2003. Allometry and size evolution in the rattlesnake, with emphasis on predatory strike performance. The University of Texas.

Lanfear R, Calcott B, Ho SY and Guindon S. 2012a. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol 29: 1695-1701. doi: 10.1093/molbev/mss020

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

Lavin MF ES, Birrell G, St. Pierre L, Guddat L, de Jersey J, Masci P. 2010. Snake venom nerve growth factors. In: Handbook of venoms and toxins of reptiles. In: SP M, editor. Handbook of venoms and toxins of reptiles. Boca Raton, Florida: CRC Press. p. 377-392.

Lavin-Murcio PA and Dixon JR. 2004. A new species of coral snake (Serpentes, Elapidae) from the Sierra de Tamaulipas, Mexico. Phyllomedusa: Journal of Herpetology 3: 3-7.

Lawing AM and Polly PD. 2011. Pleistocene climate, phylogeny, and climate envelope models: an integrative approach to better understand species' response to climate change. Plos One 6: e28554.

Leaché AD, Chavez AS, Jones LN, Grummer JA, Gottscho AD and Linkem CW. 2015. Phylogenomics of Phrynosomatid lizards: conflicting signals from sequence capture versus restriction site associated DNA sequencing. Genome biology and evolution 7: 706-719.

Lee C. 1972. Chemistry and pharmacology of polypeptide toxins in snake venoms. Annual review of pharmacology 12: 265-286.

Li H and Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754-1760. doi: Doi 10.1093/Bioinformatics/Btp324

Linnaeus C. 1758. Systema naturae. Regnum Animale.

Lynch VJ. 2007. Inventing an arsenal: adaptive evolution and neofunctionalization of snake venom phospholipase A2 genes. BMC Evolutionary Biology 7: 2.

Ma Y, Zhao Y, Zhao R, Zhang W, He Y, Wu Y, Cao Z, Guo L and Li W. 2010. Molecular diversity of toxic components from the scorpion Heterometrus petersii venom revealed by proteomic and transcriptome analysis. Proteomics 10: 2471-2485.

Mackessy SP. 2002. Biochemistry and pharmacology of colubrid snake venoms. Toxin Reviews 21: 43-83.

Mackessy SP. 2010a. Evolutionary trends in venom composition in the Western Rattlesnakes (*Crotalus viridis* sensu lato): Toxicity vs. tenderizers. Toxicon 55: 1463-1474.

Mackessy SP. 2010b. The field of reptile toxinology: snakes, lizards and their venoms. In: Mackessy S, editor. Handbook of Venoms and Toxins of Reptiles. Boca Raton, FL. : CRC Press/Taylor & Francis Group, . p. 3-23.

Mackessy SP and Baxter LM. 2006. Bioweapons synthesis and storage: the venom gland of front-fanged snakes. Zoologischer Anzeiger 245: 147-159.

Mackessy SP, Sixberry NM, Heyborne WH and Fritts T. 2006. Venom of the Brown Treesnake, *Boiga irregularis:* Ontogenetic shifts and taxa-specific toxicity. Toxicon 47: 537-548.

Margres MJ, Aronow K, Loyacano J and Rokyta DR. 2013. The venom-gland transcriptome of the eastern coral snake (*Micrurus fulvius*) reveals high venom complexity in the intragenomic evolution of venoms. BMC Genomics 14: 1-18.

McDiarmid RW, Campbell JA and Touré TSA. 1999. Snake species of the world: a taxonomic and geographic reference. Vol. 1: Washington, DC: Herpetologists' League.

Miller KB and Bergsten J. 2012. Phylogeny and classification of whirligig beetles (Coleoptera: Gyrinidae): relaxed-clock model outperforms parsimony and time-free Bayesian analyses. Syst Entomol 37: 706-746.

Milne I, Lindner D, Bayer M, Husmeier D, McGuire G, Marshall DF and Wright F. 2009. TOPALi v2: a rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. Bioinformatics 25: 126-127.

Minton SA. 1956. Some properties of North American pit viper venoms and their correlation with phylogeny. In: Buckley EE, Porges N, editors. Venoms. Washington, D.C.: American Asociation for the Advancement of Science. p. 145-151.

Minton SA. 1992. Serologic relationships among pit vipers: evidence from plasma albumins and immunodiffusion. In: Campbell JA, Brodie Jr. ED, editors. Biology of the pitvipers. Tayler, Texas: Selva. p. 155-161.

Müller F. 1879. Ituna and Thyridia: a remarkable case of mimicry in butterflies. Transactions of the Entomolgical Society of London 1879: 20-29.

Murphy RW, Fu J, Lathrop A, Feltham JV and Kovac V. 2002. Phylogeny of the rattlesnakes (*Crotalus* and *Sistrurus*) inferred from sequences of five mitochondrial DNA genes. In: Schuett GW, Hoggren M, Douglas ME, Greene HW, editors. Biology of the vipers. Eagle Mountain, Utah: Eagle Mountain Publishing. p. 69-92.

Nakashima K-i, Nobuhisa I, Deshimaru M, et al. 1995. Accelerated evolution in the protein-coding regions is universal in crotalinae snake venom gland phospholipase A2 isozyme genes. Proceedings of the National Academy of Sciences USA 92: 5605-5609.

Navarro-Siguenza AG, Peterson AT, Nyari A, Garcia-Deras GM and Garcia-Moreno J. 2008. Phylogeography of the Buarremon brush-finch complex (Aves, Emberizidae) in Mesoamerica. Mol Phylogenet Evol 47: 21-35.

Nelsen DR, Nisani Z, Cooper AM, Fox GA, Gren EC, Corbit AG and Hayes WK. 2014. Poisons, toxungens, and venoms: redefining and classifying toxic biological secretions and the organisms that employ them. Biological Reviews 89: 450-465.

Pahari S, Mackessy SP and Kini RM. 2007. The venom gland transcriptome of the Desert Massasauga Rattlesnake (*Sistrurus catenatus edwardsii*): towards an understanding of venom composition among advanced snakes (Superfamily Colubroidea). BMC Molecular Biology 8: 115.

Palmer CA, Hollis DM, Watts RA, Houck LD, McCall MA, Gregg RG, Feldhoff PW, Feldhoff RC and Arnold SJ. 2007. Plethodontid modulating factor, a hypervariable salamander courtship pheromone in the three-finger protein superfamily. FEBS Journal 274: 2300-2310.

Parkinson CL. 1999. Molecular Systematics and Biogeographical History of Pitvipers as Determined by Mitochondrial Ribosomal DNA Sequences. Copeia 1999: 576-586. doi: 10.2307/1447591

Parkinson CL. 2002. Multigene phylogenetic analysis of pitvipers with comments on their biogeography. In: Schuett GW, Hoggren M, Douglas ME, Greene HW, editors. Biology of the Vipers. Eagle Mountain, Utah: Eagle Mountain Publishing. p. 93-110.

Parmley D and Holman JA. 2007. Earliest fossil record of a Pigmy Rattlesnake (Viperidae : *Sistrurus* Garman). Journal of Herpetology 41: 141-144.

Peterson BK, Weber JN, Kay EH, Fisher HS and Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS One 7: e37135.

Pfennig DW, Harcombe WR and Pfennig KS. 2001. Frequency-dependent batesian mimicry - Predators avoid look-alikes of venomous snakes only when the real thing is around. Nature 410: 323-323.

Posada D. 2008. jModelTest: phylogenetic model averaging. Mol Biol Evol 25: 1253-1256. doi: 10.1093/molbev/msn083

Pritchard JK, Stephens M and Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Pyron RA, Burbrink FT, Colli GR, de Oca ANM, Vitt LJ, Kuczynski CA and Wiens JJ. 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.

Racine JS. 2012. RStudio: A Platform-Independent IDE for R and Sweave. Journal of Applied Econometrics 27: 167-172.

Rádis-Baptista G, Kubo T, Oguiura N, Svartman M, Almeida T, Batistic RF, Oliveira EB, Vianna-Morgante ÂM and Yamane T. 2003. Structure and chromosomal localization of the gene for crotamine, a toxin from the South American rattlesnake, *Crotalus durissus terrificus*. Toxicon 42: 747-752.

Rambaut A and Drummond A. 2007. Tracer v. 1.5. Computer program and documentation distributed by the authors. In.

Rannala B, Huelsenbeck JP, Yang ZH and Nielsen R. 1998. Taxon sampling and the accuracy of large phylogenies. Systematic Biology 47: 702-710.

Reyes-Velasco J, Gruenwald CI, Jones JM and Weatherman GN. 2010. Rediscovery of the rare autlan long-tailed rattlesnake, *Crotalus lannomi*. Herpetological Review 41: 19-25. .

Robinson MD, McCarthy DJ and Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139-140.

Rogones T. VSEARCH GitHub repository.

Ronquist F, Teslenko M, van der Mark P, et al. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic biology 61: 539-542.

Roze J. 1967. A checklist of the new world venomous coral snakes (Elapidae), with descriptions of new forms. American Museum Novitates: 1-60.

Ruder T, Sunagar K, Undheim EA, et al. 2013. Molecular phylogeny and evolution of the proteins encoded by coleoid (cuttlefish, octopus, and squid) posterior venom glands. Journal of Molecular Evolution 76: 192-204.

Sherbrooke WC, Westphal MF and Brush T. 2006. Responses of greater roadrunners during attacks on sympatric venomous and nonvenomous snakes. The Southwestern Naturalist 51: 41-47.

Siigur E, Aaspõllu A and Siigur J. 2001. Sequence diversity of *Vipera lebetina* snake venom gland serine proteinase homologs – result of alternative-splicing or genome alteration. Gene 263: 199-203.

Silvestro D and Michalak I. 2012. raxmlGUI: a graphical front-end for RAxML. Org Divers Evol 12: 335-337.

Slowinski JB. 1995. A phylogenetic analysis of the New World coral snakes (Elapidae: Leptomicrurus, Micruroides, and Micrurus) based on allozymic and morphological characters. J Herpetol: 325-338.

Slowinski JB and Keogh JS. 2000. Phylogenetic Relationships of Elapid Snakes Based on Cytochrome b mtDNA Sequences. Molecular Phylogenetics and Evolution 15: 157-164.

Smith SM. 1975. Innate recognition of coral snake pattern by a possible avian predator. Science 187: 759-760.

Smith SM. 1977. Coral-snake pattern recognition and stimulus generalisation by naive great kiskadees (Aves: Tyrannidae). Nature 1977: 535-536

Stamatakis A, Hoover P and Rougemont J. 2008. A Rapid Bootstrap Algorithm for the RAxML Web Servers. Systematic biology 57: 758-771.

Stille B. 1987. Dorsal Scale Microdermatoglyphics and Rattlesnake (*Crotalus* and *Sistrurus*) Phylogeny (Reptilia, Viperidae, Crotalinae). Herpetologica 43: 98-104.

Sunagar K, Jackson TN, Undheim EA, Ali SA, Antunes A and Fry BG. 2013. Three-Fingered RAVERs: Rapid Accumulation of Variations in Exposed Residues of Snake Venom Toxins. Toxins 5: 2172-2208.

Takasaki C, Itoh Y, Onda H and Fujino M. 1992. Cloning and sequence analysis of a snake, Atractaspis engaddensis gene encoding sarafotoxin S6C. Biochem Bioph Res Co 189: 1527-1533.

Tamiya T and Fujimi TJ. 2006. Molecular evolution of toxin genes in Elapidae snakes. Molecular Diversity 10: 529-543.

Tanner WW. 1966. A new rattlesnake from western Mexico. Herpetologica 22: 298–302.

Terrat Y and Ducancel F. 2013. Are there unequivocal criteria to label a given protein as a toxin? Permissive versus conservative annotation processes. Genome Biol 14: 406.

Thompson JD, Higgins DG and Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research 22: 4673-4680.

Uetz P and Etzold T. 1996. The EMBL/EBI Reptile Database. Herpetological Review 27: 174-175. The Reptile Database [Internet]. 2015 [cited 2015 April 4]. Available from: http://www.reptile-database.org

Undheim EA, Jones A, Clauser KR, Holland JW, Pineda SS, King GF and Fry BG. 2014. Clawing through evolution: toxin diversification and convergence in the ancient lineage Chilopoda (Centipedes). Molecular Biology and Evolution: msu162.

Vaidya G, Lohman DJ and Meier R. 2011. SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. Cladistics 27: 171-180.

Vaiyapuri S, Wagstaff SC, Harrison RA, Gibbins JM and Hutchinson EG. 2011. Evolutionary analysis of novel serine proteases in the venom gland transcriptome of *Bitis gabonica rhinoceros*. PLoS One 6: e21532.

Vidal N. 2002. Colubroid systematics: evidence for an early appearance of the venom apparatus followed by extensive evolutionary tinkering. Toxin Reviews 21: 21-41.

Villa RA and Uriarte-Garzon P. 2011. *Crotalus stejnegeri* (long-tailed rattlesnake). Herpetological Review 42: 572.

Vonk FJ, Casewell NR, Henkel CV, et al. 2013a. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proceedings of the National Academy of Sciences USA 110: 20651-20656.

Vonk FJ, Casewell NR, Henkel CV, et al. 2013b. The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proceedings of the National Academy of Sciences of the United States of America 110: 20651-20656.

Wallace AR. 1871. Contributions to the theory of natural selection: Macmillan.

Whittington CM, Papenfuss AT, Kuchel PW and Belov K. 2008. Expression patterns of platypus defensin and related venom genes across a range of tissue types reveal the possibility of broader functions for OvDLPs than previously suspected. Toxicon 52: 559-565.

Whittington CM, Papenfuss AT, Locke DP, et al. 2010. Novel venom gene discovery in the platypus. Genome Biol 11: R95.

Wuster W, Ferguson JE, Quijada-Mascarenas JA, Pook CE, Salomao Mda G and Thorpe RS. 2005. Tracing an invasion: landbridges, refugia, and the phylogeography of the Neotropical rattlesnake (Serpentes: Viperidae: Crotalus durissus). Mol Ecol 14: 1095-1108.

Xie W, Lewis PO, Fan Y, Kuo L and Chen M-H. 2011. Improving Marginal Likelihood Estimation for Bayesian Phylogenetic Model Selection. Systematic biology 60: 150-160.

Zweifel RG. 1959. Additions to the herpetofauna of Nayarit, Mexico. American Museum Novitates: 1-13.

Zwickl DJ and Hillis DM. 2002. Increased taxon sampling greatly reduces phylogenetic error. Systematic Biology 51: 588-598.

Biographical Information

Jacobo Reyes-Velasco was born in Guadalajara, Mexico on September 25th, 1986 to Oscar Reyes-Ruvalcaba and Aranzazu Velasco-Lafarga. During his early years in the mountains of the state of Jalisco and in Chihuahua, in norther Mexico, but later his family established in the state of Colima, in the west coast of Mexico, were he grew up. He showed a fascination for nature from an early age, which sometimes got him into trouble at school. He went to college to study biology at the Universidad de Colima, but later transferred to the Universidad de Guadalajara, were he obtain his BS in biology in 2009. Jacobo latter received a Ph.D. in Quantitative Biology from the University of Texas at Arlington (2015). He has conducted extensive fieldwork in the tropics and has authored and coauthored several articles in peer-reviewed journals such as *Herpetological Review, Molecular Phylogenetics and Evolution, Chelonian Conservation and Biology, Zootaxa and Molecular Biology and Evolution.* His early work with the conservation of freshwater turtles earned him several national and international awards, including the Presidential Youth Medal, awarded by the president of Mexico (2003). His research has been recognized by several institutions including the University of Texas at Arlington and the National Geographic Society. Jacobo's research focuses on evolutionary biology and includes biogeography, systematics, taxonomy and conservation of reptiles and amphibians, as well as on the evolution of complex phenotypes. Jacobo will be conducting postdoctoral research in East Africa while hosted by the New York University campus Abu Dhabi, in the United Arab Emiriates. Apart from research in biology he enjoys exploring the outdors, a weird mix of reggae and electronic music, playing the didgeridoo and wildlife photography. He is also a funding member of *Entorno Biotico*, a non-for-profit organization dedicated to sustainable development, conservation and research in western Mexico. Jacobo is also an avid vexillologist (he really likes flags).