Toxic Forms Most Beautiful: The Evolutionary Dynamics of Rear-Fanged Snake Venoms

by

Peter A. Cerda

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Ecology and Evolutionary Biology)
in the University of Michigan
2023

Doctoral Committee:

Assistant Professor Alison Davis Rabosky, Co-Chair
Professor Thomas Duda, Jr., Co-Chair
Professor James Bardwell
Professor Daniel Rabosky

Peter A. Cerda

pacerda@umich.edu

ORCID iD: 0000-0001-5275-8541

Dedication

I would like to dedicate this dissertation to my parents Maria and Mago Cerda, my siblings David, Michael, Melissa, Cyndi, and Linda, my many nieces and nephews for always supporting me, and to Patsy DeLacey for joining me on this wild ride.

Acknowledgements

I would like to thank my advisors Thomas Duda and Alison Davis Rabosky, for their never-ending support and guidance during my time at Michigan. I would like to thank my committee members, Dan Rabosky and James Bardwell, for their help and support on these projects. I would also like to thank my co-authors on the projects included in my dissertation, Jenna Crowe-Riddell, Haley Martens, and Kristy Srodawa, for being amazing collaborators.

I am lucky to be surrounded by fantastic lab members and would like to thank all past and present members of the herpetology and mollusk divisions at the University of Michigan Museum of Zoology. I would also like to thank the EEB community for being supportive of me as a scientist and as a person. I am grateful for the friends I have made here and thank them endlessly for helping me make this place my home.

Support for the research included in this dissertation was granted by the University of Michigan, including the Department of Ecology and Evolutionary Biology, the Museum of Zoology, Rackham Graduate School, and faculty start-up funds granted to my Co-advisor Alison Davis Rabosky. It was also supported by the Theodore Roosevelt Memorial Fund and the David and Lucile Packard Foundation granted to Dan Rabosky.

Table of Contents

List of Tables

List of Figures

Abstract

A chief goal in evolutionary biology is to understand how ecological factors and genetic processes influence the evolution of traits. However, a complicating factor is that traits are often encoded by multiple genes, most of which are unknown in non-model organisms. This challenge results in studies of trait variation that lack insight into the mechanisms or evolutionary forces acting on them. My dissertation takes advantage of a trait comprised of direct gene products — where the link between genetics and traits is clearly understood — therefore providing a tractable system for studying the evolution of genes and phenotypes. Specifically, my dissertation tests hypotheses about phylogenetic patterns of trait diversity and the molecular evolution of genes and gene families responsible for complex trait. My dissertation uses snake venom, a "cocktail" of toxins that functions for prey capture and/or predator defense, as a system to study the evolution of complex traits. Venom components interact directly with other organisms by targeting specific physiological functions, are typically encoded by large gene families, and are usually under high selective pressures. Several genetic mechanisms are associated with the generation of variation within a lineage, including rapid gene evolution, gene duplication, and changes in gene expression. Venoms are integral to prey acquisition in snake species, and it is presumed that diet is the strongest selective pressure shaping venom evolution. Most of our understanding of snake venom evolution comes from viperid and elapid venoms. However, we know very little about the venoms of the highly diverse group of colubrid snakes. Colubrid snakes comprise over 50% of snake species, are found worldwide, inhabit different habitat types,

and exhibit a broad range of dietary preferences. Thus, by excluding colubrid snakes, we have an incomplete picture of the evolution of snake venoms. I leverage extensive collections of snake venom glands from Neotropical species to characterize toxin expression profiles and determine patterns of toxin gene family evolution. I specifically use a transcriptomic approach to determine the composition of several genera of snakes to test if front fanged species have a more complex venom than rear-fanged species. I found that species of both rear-fanged snakes and front-fanged elapids have a less complex venom than viperid snakes. I determined the evolutionary history of a dominate venom component, C-type lectins, in rear-fanged *Helicops* species, finding that the gene family was comprised of two clades undergoing different rates of evolution. A highly diversified and rapidly evolving clade also possessed a novel insertion which I predict to be highly influential in prey capture. Finally, I characterized the venom gland transcriptome of several colubrid snakes which were dominated by the same gene family, three finger toxins, and determined that the gene family is under high rates of evolution and uncovered possible heterodimeric interactions among expressed transcript products. My work on rear-fanged snake venom evolution shows the importance of examining undescribed species as novel structural and regulatory patterns can be uncovered, though improved knowledge of toxin function and interaction with prey items is needed to fully understand the drivers of venom evolution.

Chapter 1 Introduction

One of the major goals of evolutionary biology is to understand the molecular and genetic mechanisms underlying the evolution of phenotypes. How can a simple change in amino acid sequence result in such spectacular changes in an organism's morphology or physiology? Much focus has been placed on developmental genes, such as the *HOX* gene family, which are highly conserved genes responsible for animal body segmentation [1,2]. Of particular interest are genes and gene families which are responsible for traits associated with species interactions, such as immune system genes like those encoding MHC [3]. These genes often play a role in local adaptation [4,5], but understanding the evolution of these genes and the resulting trait in non-model organisms has been a challenge as multiple genes often play a role in the development of the phenotypic trait, many of which are unknown [2,5].

Many studies that link genotype to phenotype have become iconic examples of natural selection in action. For example, some species of Darwin's finches exhibit a diversity of beak morphologies, allowing species to occupy different niche space on the same island [5,6]. Signatures of strong directional selection are found in two loci, *ALX1* and *HMGA2*, associated with changes in beak shape and size respectively [5,6]. Of particular interest are traits associated with species interactions. For example, variation in coat color of *Peromyscus polionotus*, a species of deer mice, is presumed to be the result of local adaptation, as populations exhibit a phenotype that potentially decreases predation risk in their immediate habitat [7]. This is due to a mutation in two genes responsible for pigment color, *Agouti* and *Mc1r* [7]. However, current

investigations of trait evolution in non-model systems are limited to using candidate loci due to the fact that many of these traits are polygenic and loci that contribute to theses phenotypes are mostly unknown [4,5]. Here I propose to investigate the evolution of an adaptive trait that is vital in species interactions in a tractable system.

Venomous taxa represent an exceptional system to explore questions related to the mechanisms responsible for the evolution of traits associated with species interactions. Venom is a toxic substance typically used by organisms (*e.g*., jellyfish, snails, wasps, spiders, scorpions, snakes, and some mammals) for prey capture and/or defense [8–13]. For example, some species of snakes and cone snails primarily use their venoms for prey capture, while slow lorises utilize their venom to deter predation [14–16]. Venom is typically composed of gene products that interact directly with other organisms by targeting specific physiological functions of prey or predators [17–21]. For example, type 1 α-neurotoxin, a member of the three finger toxin gene family (3FTx), actively blocks nicotinic acetylcholine receptors, resulting in paralysis [22]. Thus, venomous systems represent an opportunity to trace regulatory and structural changes in venom genes to changes in the venom phenotype and effects it has on prey or predators [9,23–25]. However, challenges remain when making interpretations about trait evolution. For example, studies of venom composition or transcriptomes represent a snapshot of the venom phenotype as venom expression can change throughout an individual's life or depend on if the animal has used its venom recently [13,26–28]. Nether the less, venomous taxa remain a highly attractive system to study the evolution of complex traits in non-model organisms [29,30].

Due to the medical significance of snake bites, we understand the evolution of viperid (vipers and pit vipers) and elapid (cobras, kraits, and coral snakes) venoms [22]. However much less work has been done on in the traditionally 'non-venomous' taxa such as members of the

Colubrinae and Dipsadinae groups, even though many of these species have enlarged teeth and connected glands that produce medically significant bites [31]. The venom delivery system in these snakes differs from their relatives. Viperid and elapid snakes both have a high-pressure delivery system in which the venom gland is constricted by the adductor muscle to excrete the venom, typically stored in a basal lumen, through hollow fangs located at the anterior end of the jaw [22]. Alternatively, rear-fanged snakes have a low-pressure delivery system, in which the Duvernoy's gland (homologous to the venom gland and referred to as a venom gland throughout this proposal) is not constricted by a muscle [22,32]. Thus, the venom, which is stored intracellularly, is excreted slowly into prey via grooved fangs located at the posterior end of the snake's jaw [22,33] These differences result in a relatively longer handling period during prey accusation [22].

At least among "traditionally venomous" taxa like elapids and viperids, venom phenotypes are known to be highly variable within a lineage. There are several potential reasons for this. Geographic variation is commonly observed in many viperid and elapid linages [15,34,35]. Perhaps the most common explanation for this variation is adaptation to local ecologies due to differences in selective pressures, such as diet [15,24]. While defensive venoms have evolved in snakes [36], diet appears to be the biggest factor influencing venom evolution in snakes [15,37]. Other evolutionary processes act independently in geographically isolated populations, such as gene family restructuring or diversifying selection, resulting in differences in venom components and composition [38,39]. Furthermore, there are sources of variation within an individual. Ontogenic changes provide perhaps the most drastic source of variation, as several linages expresses a different venom than adults [13]. Changes in expression patterns are correlated with prey acquisition, as expression of venom components is upregulated after use

[28]. This is presumably the due to the need to replenish venom after utilizing it for prey capture [28]. Due to these sources of variation within a species, it might be similarly difficult for colubrids to make generalized statements about the venom composition of a single species and make comparisons across species as it is for elapid and viperid snakes. The aim of my first chapter is to characterize the venom gland transcriptomes of several genera of rear-fanged snakes and compare them to front fanged species to determine if colubrid species exhibit the same level of variation as front fanged species [40].

Several genetic and gene regulatory mechanisms contribute to the diversity of snake venoms. Strong positive selection has been found to act on venom genes, resulting in high rates of evolution [41,42]. Furthermore, high rates of gene duplication have been found to play a major role in generating the diversity of venom genes [42]. Many gene families appear to follow the gene birth and death model of gene family evolution, meaning that genes family members are gained via gene duplication events or lost due to deletion or pseudogenization in the genome, resulting in restructured gene families [38,43,44]. Changes in gene expression also influence the evolution of snake venoms, as selection for increased gene expression may be the underlying mechanism responsible for myotoxin gene family expansion among populations of eastern diamondback rattlesnake [45]. Much or our knowledge about venom gene family evolution in snakes comes from the medically relevant front fanged families Viperidae and Elapidae. Much less is known about the evolution of these gene families in rear-fanged colubrine and dipsadine snake. While their venom is comprised of multiple components commonly found among venomous snakes, it is typically dominated by relatively few toxins, resulting in a less complex venom than their front-fanged relatives [22,40]. The aim of my second and third chapters are to

determine the evolutionary patterns of highly expressed toxin gene families recovered from a genus of dipsadine snakes (chapter 3) and from several species of colubrine snakes (chapter 4).

Together, my dissertation addresses questions of venom complexity and evolution in a largely neglected group of snakes. By determining the patterns of rear-fanged snake venom evolution, we are better able to make inferences about the evolution of venoms across all taxa. Further, future comparisons can be made about how genes, gene families, and ultimately traits associated with species interactions, such as MHC, evolve.

Chapter 2 Divergent Specialization of Simple Venom Gene Profiles among Rear-Fanged Snake Genera (*Helicops* and *Leptodeira*, Dipsadinae, Colubridae)[1]

2.1 Abstract

Many venomous animals express toxins that show extraordinary levels of variation both within and among species. In snakes, most studies of venom variation focus on front-fanged species in the families Viperidae and Elapidae, even though rear-fanged snakes in other families vary along the same ecological axes important to venom evolution. Here we characterized venom gland transcriptomes from 19 snakes across two dipsadine rear-fanged genera (*Leptodeira* and *Helicops*, Colubridae) and two front-fanged genera (*Bothrops*, Viperidae; *Micrurus*, Elapidae). We compared patterns of composition, variation, and diversity in venom transcripts within and among all four genera. Venom gland transcriptomes of rear-fanged *Helicops* and *Leptodeira* and front-fanged *Micrurus* are each dominated by expression of single toxin families (C-type lectins, snake venom metalloproteinase, and phospholipase A2, respectively), unlike highly diverse front-fanged *Bothrops* venoms. In addition, expression patterns of congeners are much more similar to each other than they are to species from other genera. These results illustrate the repeatability of simple venom profiles in rear-fanged snakes and the potential for relatively constrained venom composition within genera.

[1] Peter A. Cerda; Jenna M. Crowe-Riddell; Deise J. P. Gonçalves; Drew A. Larson; Thomas F. Duda; Alison R. Davis Rabosky. 2022. "Divergent Specialization of Simple Venom Gene Profiles among Rear-Fanged Snake Genera (*Helicops* and *Leptodeira*, Dipsadinae, Colubridae)." *Toxins* 14, no. 7: 489.

2.2 Introduction

Many animal groups use venoms that are comprised of toxic compounds to subdue prey and defend against predators [30]. Venom composition tends to differ considerably among closely related taxa within these groups, possibly due to differences in the type and diversity of prey species [15,37,46–50]. While venoms of a variety of taxa have been characterized [30,47,51,52], the venom composition of members of related but understudied groups is important for reconstructing the evolution of these complex phenotypes across taxa.

Much of our knowledge of snake venoms is based on information from front-fanged species of the families Viperidae (e.g., rattlesnakes) and Elapidae (e.g., coral snakes) [4,37,43,53]. These snakes inject venom through hypodermic needle-like fangs that are either hinged (viperids) or fixed (elapids) at the front of the mouth (Figure 2-1B) and utilize a high-pressure venom delivery system (Figure 2-1A) [22]. While most venom components can be found throughout venomous snake species, the venom composition of these two families tends to be drastically different [22]. Viperid venoms are typically hemorrhagic or cytotoxic and largely contain snake venom metalloproteinases (SVMPs), snake venom serine proteases (SVSPs), and phospholipase A2s (PLA2s) [54]. On the other hand, venoms of elapids are usually neurotoxic and primarily contain either PLA2s, three-finger toxins (3FTxs), or a combination of these two types [54]. While these toxin families dominate the venom profiles of front-fanged snakes, other toxins, such as C-type lectins (CTLs), are also found at varying levels [54]. Front-fanged species can differ in venom complexity [23], ranging from “simple” venoms comprised mostly of one toxin family to “complex” venoms comprised of many toxin families [37] that may be associated with inter- and intraspecific differences in predator–prey interactions [15,37,48,55].

Less of our knowledge on snake venoms comes from rear-fanged snakes of the family Colubridae (sensu Pyron et al. [56]), which makes up approximately half of all snake species [31]. Venomous colubrids produce toxins in the oral Duvernoy's gland (hereafter referred to as the "venom gland"; Figure 2-1A), which are delivered through grooved or ungrooved fangs located at the back of the mouth (Figure 2-1B) via a low-pressure venom delivery system [31,57]. Although the family contains approximately 700 venom-producing paraphyletic rear-fanged species [22], venoms have only been investigated in a few of them (Table 2-1). Emerging trends seem to suggest that colubrid venoms tend to have relatively simple compositions and that the subfamilies appear to follow a compositional dichotomy [58]. Venoms of the dipsadine subfamily tend to be dominated by SVMPs, such as some viperids, while venoms of the Colubrinae subfamily largely contain 3FTxs, such as some elapids (Table 2-1 and references within). Variation in the venoms of colubrid snakes has not been well studied, but there is evidence of ontogenetic shifts in venom composition of *Boigia irregularis* [26], as well as some geographic variation in the venom composition of *Tantilla nigriceps* [59]. Nonetheless, the venom compositions of rear-fanged snakes are still largely unknown, limiting our ability to both describe general patterns in the diversity of colubrid venoms as well as accurately model the evolutionary dynamics of snake venoms more broadly.

To increase our understanding of the venom composition of rear-fanged snakes, we characterized venom gland transcriptomes of members of the dipsadine genera *Helicops* and *Leptodeira*. While venoms of members of these genera have not previously been examined through RNA sequencing approaches, they have been subject to functional studies, and so some inferences about their composition are available [60–65]. For example, *Leptodeira annulata* and its subspecies *Leptodeira annulata pulchriceps* have venoms that differ in terms of serine

protease activity, but both appear to be rich in SVMPs and PLA2s as they exhibit proteolytic, hemorrhagic, and neurotoxic activities [61–63]. The venom of *Helicops angulatus* exhibits neurotoxic but not hemorrhagic activity and contains a cysteine-rich secretory protein, termed helicopsin, that causes respiratory paralysis in mice [64,65]. The absence of hemorrhagic activity implies that venoms of *H. angulatus* do not contain SVMPs, as these metalloproteases tend to induce hemorrhaging [66]. Hence, not all dipsadine snakes apparently possess a viperid-like venom, and some members of this subfamily may instead contain members of other toxin types.

We provide a first characterization of the venom gland transcriptomes of members of the dipsadine genera, *Helicops* and *Leptodeira*, and compare them with venom gland transcriptomes of front-fanged snakes of the genera *Bothrops* (Viperidae) and *Micrurus* (Elapidae). To do this, we sequenced venom gland transcriptomes from 14 species of these four genera, including multiple individuals of four species. We determined (i) which major toxin families are expressed, (ii) whether these venoms are simple or complex, and (iii) inter- and intraspecific levels of variation in venom composition among members of these genera. Although the venom profiles of *Bothrops* and *Micrurus* species have been characterized previously, we generated new sequence data for these genera to enable effective comparisons of transcriptomes generated via the same library preparation methods, sequencing approaches, and bioinformatic procedures.

2.3 Methods

2.3.1 Sampling

We collected individuals of *Bothrops*, *Helicops*, *Leptodeira*, and *Micrurus* species from multiple localities in Nicaragua and Peru between 2016 and 2018 (Table 2-2). Within ten minutes of euthanasia, we excised venom gland tissues and stored them in RNALater (Invitrogen, Carlsbad, CA, USA). While at remote locations, we stored samples at room

temperature for up to three weeks during active collection expeditions. Then we stored the samples at −20 °C prior to export to the University of Michigan and subsequently at −80 °C until RNA extraction. We deposited whole voucher specimens in the Museo de Historia Natural, Universidad Nacional Mayor de San Marcos (MUSM) in Lima, Peru, and the University of Michigan Museum of Zoology (UMMZ; Table 2-2).

2.3.2 Extraction, Library Preparation, and Sequencing

We extracted the total RNA from venom glands using the PureLink RNA mini kit (Life Technologies, Carlsbad, CA, USA), following the recommended protocol for animal tissue. The total RNA was stored at −80 °C until submission to the University of Michigan's Advanced Genomics Core, where the RNA was quantified with a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA), and the size was visualized with an Aligent TapeStation (Santa Clara, CA, USA). Poly-A tail selected libraries were constructed with Illumina TruSeq RNASeq (San Diego, CA, USA) and NEBNext Ultra II RNA (Ipswich, MA, USA) library kits, and 150 bp paired-end sequencing was conducted on Illumina HiSeq4000 or Illumina NovaSeq6000 machines (Table 2-3).

2.3.3 Bioinformatics

We assessed the raw read quality for each individual using FastQC v0.11.6 [67] and used Trimmomatic v0.38 to trim adapters and remove low-quality reads [68] for downstream phylogenetic and transcriptomic analysis. We assembled the transcriptome of each individual using two methods, Trinity v2.6.6 [69] and Extender [70], a seed-and-extend assembler that has been shown to recover a high level of isoform diversity in snake venom families [71]. The

extender assemblies were generated with combined trimmed forward and reverse paired reads that were merged with PEAR v0.9.6 [72].

We merged the assembled transcriptomes produced by Trinity and Extender to construct a single transcriptome for each individual. For each transcriptome, we followed a previously published protocol to filter out low-quality transcripts and chimeras [73,74] using transRate v1.0.3 [75] and a BLAST-based method [76], respectively. We used Corset v1.07 [77] to remove duplicate transcripts and select a single representative transcript (the longest) for each putative gene, using SALMON v0.11.2 [78] as our aligner option. To find open reading frames, we used transDecoder [79] and BLAST [80]; CD-HIT was used to reduce redundancy at 95% similarity [81]. We estimated RNA abundance (i.e., transcripts per million [TPM]) with the align_and_estimate_abundance.pl script in Trinity [69], which calls on RSEM v1.2.28 [82] and Bowtie2 v2.3.4.1 [83].

We created a custom database of non-toxin and toxin nucleotide sequences by downloading venom gland transcriptomes of *Crotalus adamanteus* [70], *Crotalus horridus* [84], *Micrurus fulvius* [85], *Boiga irregularis* [58], *Hypsiglena* sp. [58], and *Spilotes sulphureus* [86], the annotated genome of *Ophiophagus hannah* [87], and sequences of snake venom matrix metalloproteinase [88] from NCBI GenBank. A combination of these species has been used previously to identify putative rear-fanged snake venom components and represent a diversity of venomous snakes [86]. We used BLASTn [80] to identify known toxins in our transcriptomes using this custom database. We wrote a custom script in R [89] to annotate our nucleotide sequences, sort non-toxin and toxin transcripts by identifying annotations that matched known toxins from the custom database using the ‘grep’ function, and associate sequence identity with transcript abundance estimates. To determine the composition of toxin transcripts in the venom

gland transcriptomes, we wrote a custom R script that counted the number of venom gene contigs and total TPM for each snake toxin gene family across individuals. Toxin sequences with a TPM of 0 were removed. We divided the total TPM of each toxin family over the total TPM of all toxin genes to give a proportion of each toxin family present in the venom gland transcriptome.

2.3.4 Assembly of Mitochondrial Sequences and Phylogenetic Tree Estimation

We assembled mitochondrial sequences for our samples using HybPiper v1.3.1 [90] and a target file consisting of sequences from 16 fully annotated, publicly available mitochondrial genomes downloaded from GenBank [91–96]. To prepare the target file, all sequences annotated as rRNA, tRNA, and protein-coding genes were extracted from the mitogenome using Geneious Prime 2020.2.3 (https://www.geneious.com (accessed on 5 January 2022)). We then manually curated sequence label formatting and combined all sequences into a single HybPiper target file. We processed the raw sequence reads for all of the transcrioptomes with trimmomatic [68] and used the options "ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10:2:TRUE SLIDINGWINDOW:5:20 LEADING:20 TRAILING:20 MINLEN:36" to trim adapters and reads with a PHRED score of less than 20. We spot-checked the trimmed reads with FASTQC [67]. We combined reads that became unpaired during trimming into a single file per sample and assembled target mitochondrial genes with HybPiper using forward, reverse, and unpaired reads and default settings. The HybPiper pipeline calls on Exonerate [97], BLAST+ [80], Biopython [98], BWA [99], SAMtools [100], GNU Parallel [101], and SPAdes [102].

The assemblies for protein coding, rRNA, and informative tRNA sequences were aligned with MAFFT v7.271 [103] and the options "--maxiterate 1000", "--ep 0.123", and "--genafpair". Columns with >70% gaps were removed, and alignments were concatenated into a supermatrix

with the pxclsq and pxcat commands in phyx, respectively [104]. We estimated a phylogenetic tree using IQ-TREE v 2.1.3 [105] and the options "-m TEST" and "-mset raxml" to determine the best-fitting model, which had 19 partitions, variously assigned GTR+F, GTR+F+I, GTR+F+G4, and GTR+F+I+G4 models. The maximum likelihood tree was visualized with Figtree v.1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/ (accessed on 10 January 2022)).

2.3.5 Complexity and Variation

To estimate the levels of venom complexity, we calculated Shannon indices (H') [106] based on each unique toxin and their respective TPM expressed in the transcriptomes of the individuals examined. For cases in which multiple individuals of a species were examined, we averaged the *H'* values from these individuals to estimate the venom complexity of the species. We also averaged the values across species to estimate the relative levels of venom complexity of the genera. We quantified the extent that samples differ in levels of intra- and interspecific variation of venom composition with calculations of pairwise proportional dissimilarity (PPD) values (i.e., Brays–Curtis distances, [107]) based on proportions of toxin families recovered. We used the PPD values to estimate and compare levels of intraspecific variation in venom composition for species from which multiple individuals were examined (i.e., *Helicops angulatus*, *Leptodeira annulata*, *Micrurus lemniscatus*, and *Micrurus obscurus*). We further averaged PPD values that were calculated among species of genera to evaluate levels of interspecific variation in venom composition.

2.3.6 diceCT and Segmentation

We used diffusible iodine contrast-enhanced computed tomography (diceCT) to scan a representative specimen from each genus using 1.25% Lugol's iodine solution and a Nikon

Metrology XTH 225ST microCTscanner (Xtect, Tring, UK) at the UMMZ, following protocols outlined in Callahan et al. [108]. We segmented the maxillary bone and venom gland using the 'draw' and 'thresholding' tools in Volume Graphics Studio Max version 3.2 (Volume Graphics, Heidelberg, Germany).

2.4 Results

2.4.1 Venom Gland Gene Family Recovery

We produced and analyzed the venom gland transcriptomes of 19 individuals of six rear-fanged snake species (*Helicops* n = 5 and *Leptodeira* n = 4; Table 2-2) and eight front-fanged snake species (*Bothrops* n = 3 and *Micrurus* n = 8; Table 2-2). The number of paired reads recovered per sample ranged from 12,936,858 to 24,938,732 (Table 2-3). We found that the total toxin sequence count was associated with sequencing platforms used: transcriptomes that were sequenced on a HiSeq 4000 produced a higher toxin transcript count than those on a NovaSeq 6000, including transcriptomes of conspecifics or congeners. The number of unique toxin transcripts recovered from the venom gland of *Helicops* species ranged from 47 to 91 (Figure 2-1C). The most frequently identified toxin families in Helicops were snake venom metalloproteinase (SVMPs) and C-type lectins (CTLs). Several copies of cysteine-rich secretory proteins (CRiSPs) were recovered from two individuals of *H. angulatus* and one individual of *H. leopardinus* (Figure 2-1C). The number of unique toxins recovered from *Leptodeira* venom gland transcriptomes ranged from 29 to 99 (Figure 2-1D). Similar to the rear-fanged *Helicops*, the most common toxin families in *Leptodeira* were SVMPs, CTLs, and PLA2s (Figure 2-1D). The number of unique toxins recovered from *Micrurus* species ranged from 37 to 141. These largely included SVMPs, PLA2s, and CTLs transcripts (Figure 2-1E). The number of unique toxin transcripts recovered from *Bothrops* species ranged from 89 to 131, and the most common

toxin families recovered were SVMPs, snake venom serine proteinases (SVSPs), CTLs, and PLA2s (Figure 2-1F).

2.4.2 Venom Gland Transcriptome Expression

In total, we found that total toxin expression encompassed between 17% and 91% of the total expression in the venom gland transcriptomes studied (Table 2-3). We found that CTLs were the most highly expressed toxin family of Helicops and comprised between 63 and 99% of all toxins expressed (Figure 2-2). SVMPs were the second most abundant toxin family expressed (0.6–36%) in all but one individual of *H. angulatus*, for which CRiSPs toxins made up 24% of the venom expression profile (Figure 2-2). We found that SVMPs also comprised a considerable portion of the venom gland expression profile of rear-fanged *Leptodeira* species and encompassed 83–98% of the expressed toxins (Figure 2-2). The second most highly expressed toxins in *Leptodeira* were CTLs which ranged from 0.7 to 7% of expressed toxins (Figure 2-2). In front-fanged *Micrurus* species, we found that venom gland expression was dominated by one or two toxin families. All *Micrurus* species expressed PLA2s at a high level (49–99%; Figure 2-2). Kunitz-type serine protease inhibitors were the second most highly expressed toxin in both individuals of *M. lemniscatus* that we examined (22–26%; Figure 2-2), whereas 3FTxs were the second most highly expressed toxins of all other *Micrurus* species (13–21%; Figure 2-2). In front-fanged *Bothrops atrox* and *Bothrops brazili*, SVMPs had the highest expression, comprising 48 and 50% of the venom profile, respectively. (Figure 2-2). Bradykinin-potentiating peptides were the second most abundant toxin family found in *B. atrox* (28%), while CTLs were the second most abundant toxin found in *B. brazili* (20%; Figure 2-2). SVMPs and CTLs were expressed at similar levels (~30%) in *Bothrops bilineatus* (Figure 2-2).

2.4.3 Complexity and Variation

We calculated Shannon diversity indices to compare levels of venom complexity across individuals. These values ranged from a low of 0.662 for *Helicops leopardinus* to a high of 3.251 for *Bothrops brazili* (Figure 2-2). Overall, the four genera ranked from lowest to highest levels of venom complexity as follows: *Helicops* (1.289), *Micrurus* (1.602), *Leptodeira* (1.954), and *Bothrops* (3.014). We further estimated beta diversity statistics to determine levels of intra- and interspecific variation in venom composition. For the four species for which multiple individuals were examined, *L. annulata* and *M. lemniscatus* exhibited the lowest values (0.097 in both cases), while *H. angulatus* (0.239) and *M. obscurus* (0.497) had higher values. Within-genera comparisons showed that *Leptodeira* (0.097) and *Bothrops* (0.355) exhibited the lowest and highest values for genera, respectively, whereas *Helicops* (0.190) and *Micrurus* (0.180) had intermediate values. While intraspecific variation was found among members of these genera, we saw a greater similarity in venom gland transcriptome composition among individuals within genera than among genera (Figure 2-2).

2.5 Discussion

We used a transcriptomic approach to characterize and compare patterns of toxin variation in the venom gland transcriptomes of members of two rear-fanged (*Helicops* and *Leptodeira*) and two front-fanged (*Bothrops* and *Micrurus*) snake genera. While snake venom metalloproteinases (SVMPs) transcripts dominate the venom profile of the *Leptodeira* species, C-type lectins (CTLs) are the most highly expressed toxin family of *Helicops* species (Figure 2-2). Venom profiles of the front-fanged species are similar to those reported from these taxa previously [109–111]. While we were able to recover toxin transcripts from numerous toxin families across all individuals, we found that toxin expression within genera is typically dominated by only a single toxin class: CTLs in *Helicops*; phospholipase A2s (PLA2s) in

Micrurus; and SVMPs in both *Leptodeira* and *Bothrops*. Shannon diversity indices suggest that while complexity may vary among individuals, differences may be due to the relative contribution of each underlying transcript rather than differences in toxin family abundance. We also observed low levels of variation among individuals of the same genera but high levels of variation across genera.

The dipsadine rear-fanged species that we examined mostly exhibited low levels of venom complexity, given that their venom gland expression profiles were largely composed of transcripts of only single toxin classes (Figure 2-2). In line with previous results, front-fanged snakes tended to exhibit higher levels of venom complexity. Increased venom complexity has been correlated with large dietary breadth in both venomous cone snails and North American pit vipers [37,47]. The venoms of several rear-fanged snakes are known to have a "simpler" venom than their front-fanged relatives [58,86,112–114]. The lack of complexity observed for rear-fanged snakes may be due to the highly specialized diets that many colubrid snakes exhibit [22]. Broadly, *Leptodeira* appear to specialize in frogs while Helicops specialize in fish, though a formal analysis of dietary specialization in these two genera has not been performed [115,116]. However, the toxicological diversity of specific toxin families from the species described here has not been determined. The investigation of toxin function may reveal physiological targets and functions specific to prey types, which have been found in neurotoxins described from several rear-fanged snakes [26,86].

Several members of the colubrid subfamily Dipsadinae have been shown to use a hemorrhagic venom that is largely comprised of SVMPs [58,113]. The *Leptodeira* species we examined exhibited this type of venom profile. However, *Helicops* species had venom gland expression profiles that are dominated by transcripts of a single toxin family, CTLs. No previous

studies have shown snakes that produce venom dominated by CTLs. While functional attributes of CTLs of rear-fanged snakes are not known, CTLs of front-fanged snakes are multifunctional heterodimers that affect hemostasis by acting as anticoagulants, which can cause hemorrhaging, or as procoagulants, which can cause blood clotting [117]. In addition, CTLs of front-fanged snakes have been shown to evolve rapidly [118]. Recently, Xie et al. [119] found extensive duplication of novel dimeric CTLs genes unique to *H. leopardinus*. The novel CTLs found in *H. leopardinus* were shown to be under positive selection, but the distribution of these CTLs across the genus *Helicops* is currently unknown [119]. A previous study of *H. angulatus* found its venom lacked hemorrhagic activity, which is typically associated with SVMPs and CTLs [64]. Instead, its venom was shown to exhibit neurotoxic activity, likely due to the presence of a previously uncharacterized cysteine-rich secretory protein (CRiSP) named helicopsin [64]. Of the *H. angulatus* individuals examined here, only two possess CRiSP transcripts, and only one expressed CRiSP at a considerable level (Figure 2-2). It is unknown if the CRiSP transcripts found in our Peruvian individuals are similar to helicopsin that was previously isolated from an individual of this species from Venezuela [64].

Our sampling included multiple individuals of some species and multiple species of four genera to evaluate patterns of intra- and interspecific variation of venom profiles as inferred from transcriptome data. The levels of variation within species differed considerably. Intraspecific variation of venom profiles of *L. annulata* and *M. lemniscatus* were low, while those of *H. angulatus* and *M. obscurus* were quite distinct. For example, individuals of *H. angulatus* differed in terms of the relative abundance of CRiSPs, SVMPs, and CTLs, despite being from the same locality. Further, the described absence of hemorrhagic activity of venoms of *H. angulatus* from Venezuela [65] suggests that this species exhibits geographic variation in venom. While venom

gland profiles generated here generally match venom profiles that were described previously in other studies, geographic variation in venom composition may occur among some of the species surveyed here as well [53,110,111,120]. For example, while pooled venom of *B. brazili* from Pará, Brazil, contained 33% PLA2s, 27% SVMPs, and 14% SVSPs [120], SVMPs and CTLs were the predominant toxin components that are represented in the venom gland transcriptome of an individual from the Madre de Dios region of Peru. Geographic variation is rather common in predatory venomous species as populations presumably adapt to local prey assemblages [39,121]. However, caution should be taken when comparing venom gland transcriptomes and venom proteomes as there are cases in which there is a lack of correspondence between the abundance of expressed transcripts and translated proteins [110]. Further, the proportion of toxin expression in the whole transcriptome varied widely among the specimens examined (17.15–91.81%; Table 2-3). The variation in venom expression can possibly be attributed to differences in toxin expression over time [28]. For example, toxin expression is higher after feeding events to replenish the used venom proteins [28].

The two individuals of *M. obscurus* that we examined differed considerably in venom complexity, with the venom profile of one individual being nearly completely dominated by PLA2s, while that of the other individual was more complex and contained wapirins, 3FTxs, and Kunitz-type serine protease inhibitors (Figure 2-2; note that wapirins are included in the "Other" category in the figure). These two individuals differed in age, as one was a juvenile (*M. obscurus* 0665), while the other was an adult (*M. obscurus* 1054). Thus, the difference observed in *M. obscurus* may be due to an ontogenic shift from a more "complex" venom that is expressed by juveniles to a "simple" venom that is expressed by adults. Ontogenetic shifts in venom composition have been observed in many snake species, and this change is potentially due to

differences in the diet as individuals age [13,26,55,122]. However, more intensive sampling is needed to determine if this is indeed the case in *M. obscurus*.

The species studied here displayed varying levels of interspecific variation in venom gland transcriptomes, which is a common pattern observed across venomous taxa [11,30,37,47]. Barua and Mikheyev [123] found that while many combinations of venom components were possible, different species tend to show similar venom profiles despite phylogenetic relatedness, albeit with different proportions of respective venom compositions. While interspecific variation was observed among species examined here, we do note that most of our taxa had venom gland profiles similar to those expected given their respective families [58,123], except for species of *Helicops*. It is not clear how or why the *Helicops* species examined here arrived at a potentially novel venom phenotype. The use of CTLs by *Helicops* species may be a more efficient strategy to capture their aquatic prey [115,116]. Proteomic and functional studies should be performed on Helicops venoms to determine the abundance of CTL proteins in these venoms and how they might be used in prey capture. Future exploration of colubrid venoms will help us further our understanding of how convergent and novel venom phenotypes have evolved across all venomous snakes.

2.6 Funding

This research was supported by the University of Michigan to A.D.R. through startup funds and to P.A.C. through the C.F. Walker Scholarship and Graduate Research Awards. This research was also supported by a Theodore Roosevelt Memorial Fund award from the American Museum of Natural History to P.A.C. Field research was partially supported through funding from the David and Lucile Packard Foundation to Dan Rabosky and the University of Michigan to Iris Holmes.

2.7 Author Contributions

Conceptualization, P.A.C., T.D. and A.D.R.; Methodology, P.A.C., J.C.R., D.J.P.G. and D.A.L.; Formal Analysis, P.A.C., D.J.P.G. and D.A.L.; Investigation, P.A.C., D.J.P.G. and D.A.L.; Resources, P.A.C. and A.D.R.; Data Curation, P.A.C., D.J.P.G and D.A.L.; Writing – Original Draft Preparation, P.A.C.; Writing – Review & Editing, P.A.C., J.C.R., D.J.P.G., D.A.L., T.D. and A.D.R.; Visualization, P.A.C., J.C.R. and A.D.R; Supervision, T.D. and A.D.R.; Project Administration, A.D.R.; Funding Acquisition, P.A.C. and A.D.R. All authors have read and agreed to the published version of the manuscript.

2.8 Data Availability Statement:

The raw sequence data for each individual are available on NCBI SRA under BioProject PRJNA843733 and BioSample accession SAMN28768753-SAMN28768771.

2.9 Acknowledgments

We thank Dan Rabosky, Rudi von May, Joanna Larson, Jose Martínez-Fonseca, and Iris Holmes for assistance in organizing and running field collection in Peru and Nicaragua, along with the many people who helped with collection in the field. We thank Conservación Amazónica (ACCA) and Project Amazonas for support at field stations. We also thank SERFOR (Servicio Nacional Forestal y de Fauna Silvestre) in Peru, MARENA (Ministerio del Ambiente y los Recursos Naturales) in Nicaragua, and the U.S. Fish and Wildlife Service for collection, export, and import permits. We thank Matt Holding and Darin Rokyta for access to the Extender program, and Ramon Nagesan, Sean Callahan, and Greg Schneider for assistance with CT

scanning. We also thank the Davis Rabosky and Duda lab members and anonymous reviewers for helpful comments that strengthened this manuscript.

Table 2-1 Summary of venom gland transcriptome toxin profiles of several species of rear-fanged colubrid snakes.

Summary of venom gland transcriptome toxin profiles of several species of rear-fanged colubrid snakes. SVMP = snake venom metalloproteinase, 3FTx = three-finger toxin, CRiSP = cystine rich secretory protein, svMMP = snake venom matrix metalloproteinase, CTL = C-type lectins, Kunitz = Kunitz-type serine protease.

Subfamily	**Species**	**Major Venom Component(s)**	**Reference**
Colubrinae	*Ahaetulla prasina*	SVMPs, 3FTxs	[124]
	Boiga irregularis	3FTxs, SVMPs	[58]
	Dispholidus typus	SVMPs	[125]
	Spilotes sulphureus	3FTxs	[86]
	Tantilla nigriceps	3FTxs, CRiSPs, SVMPs	[59]
	Trimorphodon quadruplex	3FTxs, SVMPs	[112]
Dipsadinae	*Borikenophis portoricensis*	SVMPs	[124]
	Conophis lineatus	svMMPs	[126]
	Hypsiglena sp.	SVMPs, CRiSPs	[58]
	Phalotris mertensi	Kunitzs, SVMPs, CTLs	[127]
	Philodryas olfersii	SVMPs, CNPs	[128]
	Thamnodynastes strigatus	svMMPs	[129]

Table 2-2 Specimen information for samples sequenced in this study.

Specimen information for samples sequenced in this study. Numbers at the end of species names are field codes used to identify individuals. MUSM = Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, UMMZ = University of Michigan Museum of Zoology, EBLA = Estación Biológica Los Amigos, EBMS = Estación Biológica Madre Selva, LBM = Las Brisas del Mogotón, EBVC = Estación Biológica Villa Carmen, RB = Refugio Bartola, mm = millimeters, g = grams, F = female, M = male, J = juvenile, A = adult.

Family	Taxon	Museum Accession No.	Date Captured	Station, Country	SVL (mm)	Mass (g)	Sex	Age
Viperidae	*Bothrops atrox* 0365	MUSM 35721	21 March 2016	EBLA, Peru	589	81	F	J
	Bothrops bilineatus 0065	UMMZ 245084	11 March 2016	EBLA, Peru	744	85	F	A
	Bothrops brazili 1278	MUSM 36922	1 December 2016	EBLA, Peru	606	76	M	A
Elapidae	*Micrurus annellatus* 3275	UMMZ 248450	26 November 2018	EBLA, Peru	497	18.11	F	A
	Micrurus hemprichii 1810	UMMZ 246857	18 January 2017	EBMS, Peru	740	86	M	A
	Micrurus lemniscatus 0249	UMMZ 245082	16 March 2016	EBLA, Peru	715	65	M	A
	Micrurus lemniscatus 0336	MUSM 35905	21 March 2016	EBLA, Peru	725	50	F	A
	Micrurus nigrocinctus 3053	UMMZ 247142	22 May 2018	LBM, Nicaragua	717	64.8	F	A
	Micrurus obscurus 0665	UMMZ 246859	7 November 2016	EBVC, Peru	261	5.19	M	J
	Micrurus obscurus 1054	UMMZ 246860	22 November 2016	EBLA, Peru	775	81	M	A
Colubridae	*Helicops angulatus* 0143	UMMZ 245053	13 March 2016	EBLA, Peru	373	48.36	F	A
	Helicops angulatus 3440	UMMZ 248879	2 December 2018	EBLA, Peru	411	60	F	A

Family	Taxon	Museum Accession No.	Date Captured	Station, Country	SVL (mm)	Mass (g)	Sex	Age
	Helicops angulatus 3559	MUSM 39826	9 December 2018	EBLA, Peru	307	24.19	F	A
	Helicops leopardinus 1812	UMMZ 246808	18 January 2017	EBMS, Peru	685	220	F	A
	Helicops polylepis 1932	UMMZ 246809	18 January 2017	EBMS, Peru	823	600	F	A
	Leptodeira annulata 0468	UMMZ 245059	24 March 2016	EBLA, Peru	463	18.56	M	A
	Leptodeira annulata 0497	UMMZ 245060	27 March 2016	EBLA, Peru	590	38.02	F	A
	Leptodeira rhombifera 3241	UMMZ 247098	12 June 2018	Tecomapa, Nicaragua	665	169.5	F	A
	Leptodeira septentrionalis 3176	UMMZ 247099	3 June 2018	RB, Nicaragua	654	113.2	F	A

Table 2-3 Library preparation, sequencing platform, sequencing output, and percent of the transcriptome that was comprised of toxin transcripts for individuals used in the study.

Library preparation, sequencing platform, sequencing output, and percent of the transcriptome that was comprised of toxin transcripts for individuals used in the study.

Family	Taxon	Library Preparation	Illumina Platform	Reads Pairs	Percent Toxin Expression
Viperidae	*Bothrops atrox* 0365	TruSeq RNASeq	HiSeq 4000	22,392,182	47.97%
	Bothrops bilineatus 0065	NEBNext Ultra II	NovaSeq 6000	17,985,945	74.53%
	Bothrops brazili 1278	NEBNext Ultra II	NovaSeq 6000	14,816,998	47.25%
Elapidae	*Micrurus annellatus* 3275	NEBNext Ultra II	NovaSeq 6000	16,398,046	39.12%
	Micrurus hemprichii 1810	NEBNext Ultra II	NovaSeq 6000	19,149,986	33.68%
	Micrurus lemniscatus 0249	NEBNext Ultra II	NovaSeq 6000	16,413,190	19.58%
	Micrurus lemniscatus 0336	TruSeq RNASeq	HiSeq 4000	24,938,732	53.04%
	Micrurus nigrocinctus 3053	NEBNext Ultra II	NovaSeq 6000	18,981,692	63.38%
	Micrurus obscurus 0665	NEBNext Ultra II	NovaSeq 6000	15,955,904	40.44%
	Micrurus obscurus 1054	NEBNext Ultra II	NovaSeq 6000	16,791,601	48.82%
Colubridae	*Helicops angulatus* 0143	TruSeq RNASeq	HiSeq 4000	23,374,958	17.15%
	Helicops angulatus 3440	NEBNext Ultra II	NovaSeq 6000	17,274,735	31.03%
	Helicops angulatus 3559	NEBNext Ultra II	NovaSeq 6000	15,797,921	19.57%
	Helicops leopardinus 1812	NEBNext Ultra II	NovaSeq 6000	17,894,322	52.98%
	Helicops polylepis 1932	NEBNext Ultra II	NovaSeq 6000	12,936,858	91.81%
	Leptodeira annulata 0468	NEBNext Ultra II	NovaSeq 6000	19,191,193	27.03%
	Leptodeira annulata 0497	TruSeq RNASeq	HiSeq 4000	20,844,579	36.83%
	Leptodeira rhombifera 3241	NEBNext Ultra II	NovaSeq 6000	17,838,000	41.59%
	Leptodeira septentrionalis 3176	NEBNext Ultra II	NovaSeq 6000	17,147,031	46.39%

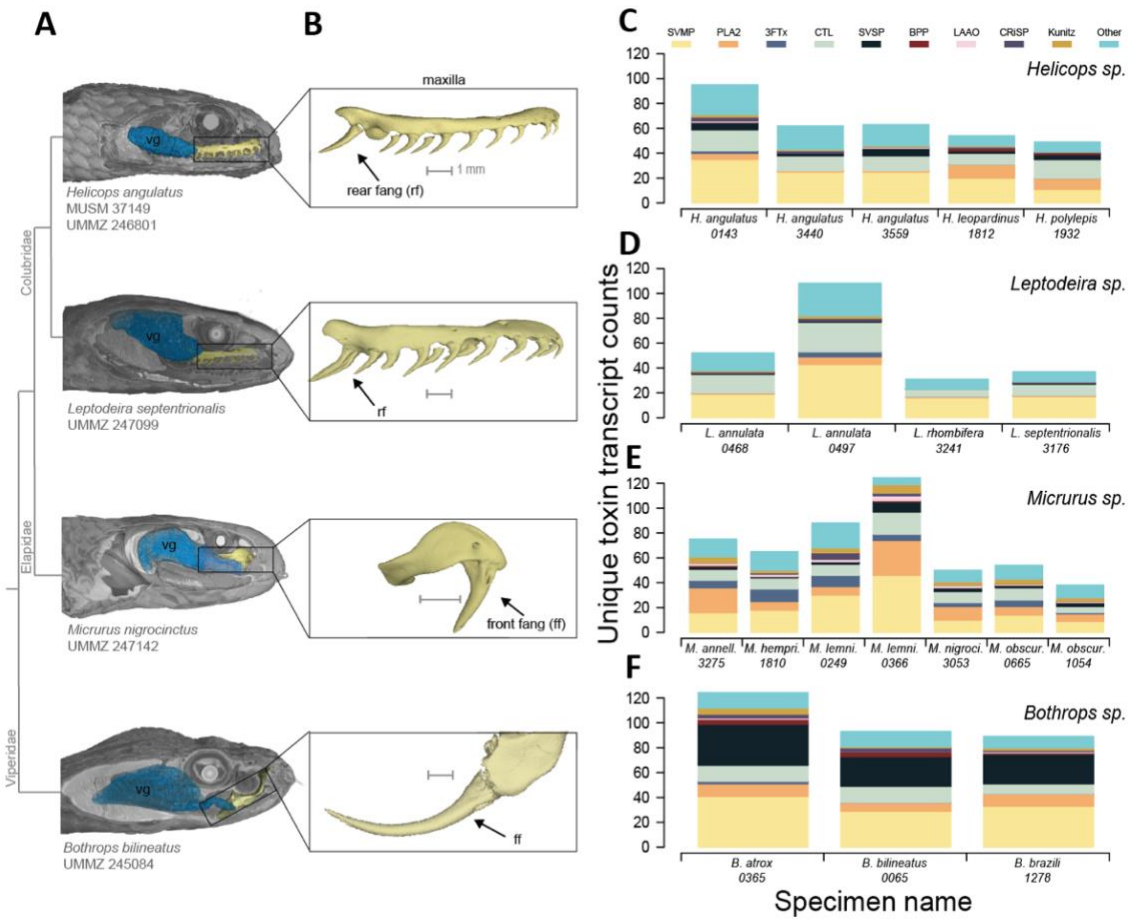

Figure 2-1 Representative venom delivery system for each genus and venom transcript counts of individuals.

(A) Soft-tissue scan of the venom delivery system for each genus mapped onto as simplified phylogeny developed in this study. (B) Isolated maxilla bone from each representative genus showing position of fangs (arrow). (C) Unique toxin transcripts recovered from each *Helicops* individual. (D) Unique toxin transcripts recovered from each *Leptodeira* individual. (E) Unique toxin transcripts recovered from each *Micrurus* individual. (F) Unique toxin transcripts recovered from each *Bothrops* individual. MUSM = Museo de Historia Natural, Universidad Nacional Mayor de San Marco, UMMZ = University of Michigan Museum of Zoology, SVMP =

snake venom metalloproteinase, PLA2 = phospholipase A2, 3FTx = three-finger toxin, CTL = C-type lectin, SVSP = snake venom serine proteinase, BPP = bradykinin-potentiating peptides, LAAO = L-amino acid oxidase, CRiSP = cystine rich secretory protein, Kunitz = Kunitz-type serine protease. Micro-CT scans of specimens vouchered in the University of Michigan Museum of Zoology (UMMZ) and Museo de Historia Natural de la Universidad Nacional Major de San Marcos (MUSM). We deposited these micro-CT scans used for venom and fang morphology for public access in the Morphosource 'Scan All Snakes' Project ID00000C374 (https://www.morphosource.org/Detail/ProjectDetail/Show/project_id/374).

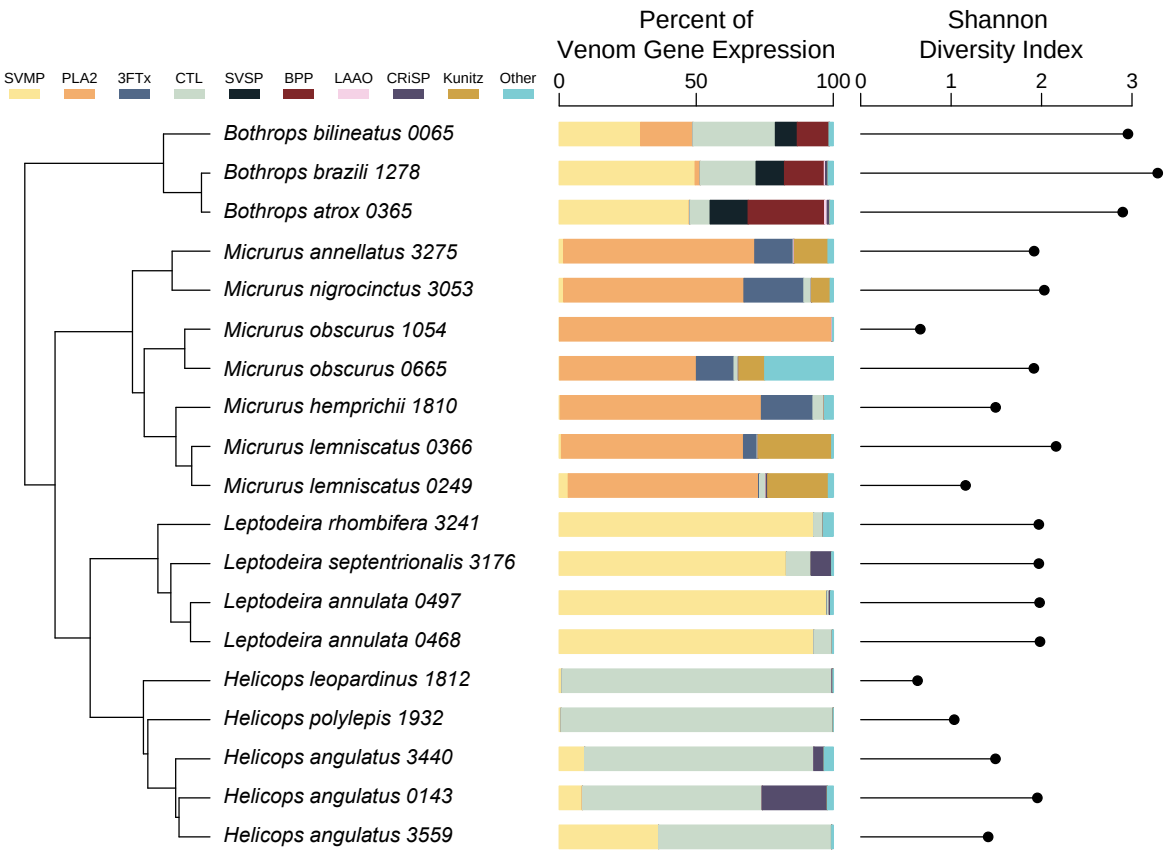

Figure 2-2 Expression of toxin gene families from venom gland transcriptomes and overall venom transcriptome diversity.

Expression of toxin gene families from venom gland transcriptomes (center) and overall venom transcriptome diversity (right) mapped to phylogeny inferred from mitochondrial gene sequences (left). Note that these data suggest there is generally higher similarity in venom profiles among individuals within genera than among genera for both metrics. *Bothrops* and *Micrurus* species are front-fanged, while *Leptodeira* and *Helicops* species are rear-fanged. SVMP = snake venom metalloproteinase, PLA2 = phospholipase A2, 3FTx = three-finger toxin, CTL = C-type lectin, SVSP = snake venom serine proteinase, BPP = bradykinin-potentiating peptides, LAAO = L-

amino acid oxidase, CRiSP = cystine-rich secretory protein, Kunitz = Kunitz-type serine *protease.*

Chapter 3 High Expression and Diversification Linked to a Novel Insertion in *Helicops* (Dipsadinae, Colubridae) C-Type Lectin Venom Genes.

3.1 Abstract

Highly diverse genes families can drive phenotypic evolution through diversification of both gene expression patterns and neofunctionalization. However, understanding how trait evolution is impacted by the interplay between gene family evolution and expression is difficult because most traits are polygenic and are comprised of unknown genes. Here, we leverage a tractable system, snake venom, to determine how changes in gene expression and sequence affect the evolutionary rates of the snake venom C-type lectin gene family in several species of *Helicops* (Dipsadinae, Colubridae). We found that the gene family was comprised of two sets of sequences that differed greatly in rates of diversification and expression. The highly expressed toxins were mostly restricted to a single, rapidly diverging cluster of sequences. Further, we found that an insertion was associated with this cluster of sequences, indicating a possible neofunctionalization event. Together, our results suggest that a combination of regulatory and structural changes has resulted in the diversification of C-type lectins in *Helicops* species. We propose that the novel insertion is being maintained by selection and may be vital to prey capture, as sequences possessing the insertion are the dominant toxin expressed in the venom transcriptomes of these species. Our findings show how structural and regulatory changes and rates of evolution interact to produce a putative novel trait.

3.2 Introduction

Gene families are groups of duplicated genes that share similar sequences and functions [130] and play a prominent role in organismal biology [131–135]. For example, the major histocompatibility complex gene family consists of numerous gene copies which encode proteins that can differ in how effective they are at recognizing and binding to different pathogens, directly impacting an individual's fitness [136,137]. Gene families often evolve under a birth-and-death model of gene family evolution, where genes are duplicated and then subject to several fates such as neofunctionalization, subfunctionalization, or loss from the genome [138,139]. This model of evolution results in restructured gene families that are comprised of numerous members that vary among species, populations, and individuals [132,140,141]. While the functional divergence of gene family members can be responsible for the evolution of many traits, changes in gene copy number or gene expression can alter phenotypes as well [142–144]. For example, increased dosage effects may occur as there are more copies of the gene to produce its product [139]. However, the ability to track interactions between gene family evolution and gene expression has been difficult as many traits rely on numerous genes, which may be uncharacterized [145,146]. Here we leverage a tractable system to investigate the interaction of gene family evolution and gene expression in snake venoms.

Snake venoms are secretions comprised of numerous toxins and toxin families that are direct gene products, resulting in a system in which structural and regulatory changes of the genes can be traced to changes in venom function and composition [147]. Venoms are used to incapacitate prey or defend against predators [30,148]. As the toxins are directly linked to an individual's fitness [37], they are typically under strong positive selection [41,149,150]. Further, toxin gene families usually evolve under a gene birth and death model, resulting in differences in

venom composition and function among and within species [43,44,150]. Challenges remain when studying the evolution of snake venom expression as approaches to quantify venom expression or composition rely on a 'snapshot' of the current venom profile. Venom has been shown to change in expression as an organism ages or last fed [13,26,28]. However, the ability to link changes in gene sequence and expression makes venom an ideal system in which to study this interaction.

Changes in gene expression have been correlated with changes in abundance of the translated toxin proteins, showing a direct link between gene expression and venom composition [151]. The expression of toxins may also differ among individuals, yielding venoms that vary in composition and complexity among and within populations [45,58,147,152]. Margres et al. [45] found the that the myotoxin gene family of *Crotalus adamanteus* is under selection for increased expression, which selects for the retention of duplicated genes as more genes results in an increased abundance of myotoxins. However, increased reliance on a subset of the toxin members relaxes selection on other gene copies, allowing for mutations, and possible neofunctionalization events, to be maintained [45].

C-type lectins (CTLs) are a common toxin expressed by members of several snake families [117]. These peptides are hemotoxic (i.e., damage blood cells, organs, tissues, or disrupt blood clotting), and diverse in function, as they may affect thrombosis (blood clotting) and hemostasis (blood flow; [153–156]). CTLs also vary in form, as there is a common ancestral monomeric form consisting of identical subunits and a derived dimeric form consisting of alpha and beta subunits [119,153]. A sugar binding site in the mature protein is known to interact with target sites in prey and differs in amino acid composition (EPN and QPD, among others), suggesting selection is acting to diversify the type of carbohydrates it may bind to [31]. Our

current understanding of the evolutionary history of CTLs suggests that numerous copies of this family were present in the last common ancestor of venomous snakes [119].

While CTLs are common across snake species, they rarely are the dominant toxin expressed [123]. However, *Helicops* species (Dipsadinae, Colubridae), are an exception as they have been found to have venom gland transcriptomes dominated by CTL transcripts [40]. Recently, *Helicops leopardinus* was found to contain numerous CTL gene duplicates as well as a novel insertion within some toxin transcripts [119]. It is unclear if the duplicated transcripts of *H. leopardinus* are responsible for the high CTL expression found in *Helicops* species and if these toxins are under high selective pressures.

We used a transcriptomic approach to determine how the C-type lectins toxin gene family is evolving across members of the *Helicops* genus. We specifically tested 1) if high CTL toxin expression is associated with a specific cluster of sequences or dispersed throughout the gene tree and 2) if rates of evolution are correlated with expression level, and 3) where selection is acting within the sequence. We predicted that highly expressed toxins will be clustered in the gene tree and would be under higher rates of evolution, as in other more traditionally studied venom systems with higher toxicity levels.

3.3 Materials and Methods

3.3.1 Bioinformatics

We used RNAseq data from two *Helicops angulatus* individuals, one individual of *Helicops leopardinus*, and one individual of *Helicops polylepis* that were collected in Perú between 2016 and 2018 and previously sequenced (Table 3-1; Chapter 2; [40]). We used Trimmomatic v0.33 [68] to trim adapters and remove low-quality sequences. We then used the transcriptomic assembly and annotation guide described by Nachtigall et al. [157]. Briefly, we

used two transcriptome assemblers as per the recommendation of Holding et al. [71]. We used Trinity v2.14 [69] to assemble paired reads into a transcriptome and Pear v0.9.6 [72] to merge paired reads which were then assembled into transcriptomes with Extender [70]. We merged resulting transcriptomes for each individual and used CD-HIT v4.8.1 [81] to remove identical transcripts.

We used ToxCodAn, an annotation program that uses machine learning to identify and annotate toxins present in a transcriptome [157]. We further cleaned the output by running the ToxCodAn cleaner pipeline, which separates toxin families, checks for erroneous start codons, and removes sequences with low coverage or a transcript per million transcripts (TPM) value of 0 [157]. We used CodAn [158] to identify and annotate non-toxin transcripts. We combined annotated toxins and non-toxins into a single nucleotide file. We used ChimeraKiller (https://github.com/masonaj157/ChimeraKiller; accessed July 12th, 2022) to identify and remove chimeras from our transcripts. To cluster alleles of loci into a representative sequence, we used CD-HIT v4.8.1 [81] to reduce redundancy by 99%. We then estimated TPM values (transcripts per million reads, a proxy for expression levels) of all remaining transcripts using RSEM v1.2.28 [82]. We wrote a custom R script to separate toxin families and append the TPM values to the sequence name to easily map and track expression values to sequences [89].

3.3.2 Gene Tree Construction

To construct gene trees, we combined CTL sequences from all individuals into a single nucleotide file. We used MAFFT v7.505 [103] to align sequences with 1000 replicates and included the options "--ep = 0.123", which sets a gap penalty, and "--genafpair" options, which constructs the pairwise alignment with a local algorithm using the set gap penalty. We then used IQ-Tree v2.0.3 [159] to construct a gene tree with the MAFFT alignment. We used options "--m

MFP" [160] to find the best substitution model, "--maxiterate 100000" which uses UFBoot [161] to assess branch support across the topology after 100,000 iterations, and the "-bnni" option to reduce the impact of severe model violation on the estimation of branch support values. We used FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/ accessed March 23, 2023) to root the tree based on the presence of an amino acid motif commonly found in ancestral monomeric CTLs (similar to how Xie et al. [119] rooted their CTL gene tree).

3.3.3 Expression simulations

To test if transcripts with high expression are clustered in the gene tree, we used a simulation-based approach to generate a null distribution of expression levels across the tree and compare them to observed values. For the observed data, we first log transformed TPM values to normalize the data. Then, we calculated the distance to the root for each tip of the gene tree using the distRoot function in the package 'adephylo' [162] in R [89]. We calculated the observed correlation coefficient by regressing TPM against the distance to root, and then ran 10,000 simulations in which we randomized (without replacement) log-transformed TPM values while keeping the distance to the root for each tip standard before calculating correlation coefficients. We then calculated the probability of our observed correlation given the distribution of the randomized trials.

3.3.4 Selection tests

To test for selection acting on toxin sequences, we used PAML v 4.9 [163]. To determine if sequences associated with high expression were evolving at a faster rate than sequences with lower expression, we ran the CodeML branch-site model test [163]. We partitioned sequences in the gene tree into two main clades based on branch length and expression levels. We conducted

likelihood ratio test to determine if branches associated with sequences with high expression had significantly different evolutionary rates than a null model where all branches were evolving under a single rate [163,164]. We also used several methods implemented in HyPhy v2.5.41 [165]. To test for evidence of past episodes of selection acting on individual codon sites, we used the Mixed Effects Model of Evolution (MEME) method [166]. Finally, to test for consistent or recent selection acting on individual sites, we used the Fast Unconstrained Bayesian Approximation (FUBAR) method [167]. MEME and FUBAR both take the alignment and gene tree into account when identifying selection acting on codons.

3.4 Results:

3.4.1 Toxin recovery and expression

We found a total of 66 toxin transcripts from two individuals of *Helicops angulatus* and one individual each of *Helicops leopardinus* and *Helicops polylepis*. We were able to identify 14 unique toxin gene families from the transcriptomes. We removed one CTL transcript (Hean3440 Toxin21600) after constructing the tree as it appears to have been an erroneous sequence upon further inspection of the sequence alignment (i.e., it had numerous base pair mismatches and was much longer than other CTL sequences). The percentage of total toxin transcription within the entire transcriptome ranged from 42% to 84% (Table 3-2). Of those, CTLs were the most diverse family, with 22 members recovered across all individuals. CTLs also dominated the fraction of the transcriptome that was comprised of toxin transcripts, responsible for 82% to 92% of the total toxin transcription (Table 3-2). Other families found include Kunitz, cystine rich secretory proteins (CRiSPs), and snake venom metalloproteinase (SVMPs). However, most of these families are comprised of lowly-expressed transcripts (less than 16% of total toxin expression; Table 3-2).

3.4.2 Gene tree construction

We constructed a gene tree of the CTL toxin sequences using IQ-Tree with an alignment determined with Mafft. We found that the K2P+R2 substitution model best fit our data and so used this model to construct the gene tree. The tree two main clusters (Figure 3-2). One cluster contains sex lowly expressed members with long branches (Figure 3-2) and an EPN and QPD carbohydrate binding motif consistent with ancestral monomeric CTLs previously described (Figure 3-3;[119]). Interestingly, there is a pair of toxins from *H. angulatus* that appears to have lost this motif. The other cluster contains sequences with very short branches and exhibit high levels of expression (Figure 3-2). A key characteristic of this group of sequences is the possession of an insertion containing two to four cysteine codons, which was previously described in an individual of *H. leopardinus* (Figure 3-3; [119]). Species-specific subclades exist within Clade B which may represent evidence of recent duplications within each species. However, while sequence from *H. angulatus* and *H. leopardinus* occur uniquely together, sequences from *H. polylepis* occur with sequences of *H. angulatus*, suggesting that CTLs of *H. angulatus* share a recent common ancestor with those of *H. polylepis*.

3.4.3 Cluster specific expression bias

We detected a negative relationship between level of expression and branch length (R = -0.729). From 10,000 permutations, we found that the observed correlation coefficient was significantly different from the randomized data (Figure 3-4; P=0.0005). This negative relationship was driven by branch length (our proxy for cluster membership), as the highly diversified cluster with very short branch lengths had significantly higher expression levels than the CTL clade with much longer branch lengths.

3.4.4 Selection Test

We used PAML [163], to run a branch-site model test to determine if branches associated with sequences with large TPM values show differences in rates of evolution than those with low TPM values. We found that there was a significant difference in rates of evolution between these two sets of sequences (Table 3-3; P =0.004). Our MEME analysis found 27 sites under episodic selection, while our FUBAR analysis found 22 sites under pervasive selection (Table 3-4). 13 sites were found to be under selection by all three tests (Table 3-4). A high concentration of sites under selection were found in the novel insertion within the highly expressed cluster of sequences (Figure 3-3). We found 18 sites under selection with our MEME analysis and 19 sites with our FUBAR analysis, 13 of which were congruent within the novel insertion (Table 3-4). Our FUBAR analysis also identified 17 sites under purifying selections, many of which were cysteines.

3.5 Discussion:

We used a transcriptomic approach to determine if gene expression was correlated with rapidly diversifying sequences in a tractable trait. Consistent with a previous study, we identified numerous toxins present in the transcriptomes of *Helicops*, but total toxin expression was dominated by the CTL gene family [40]. We investigated the evolutionary history of the CTL gene family within *Helicops.* We mapped expression onto this tree and found a high correlation between group membership and gene expression, with a highly diversified group of sequences being comprised of highly expressed transcripts. Additionally, we found a difference in rates of evolution between branches associated with sequences with high levels of expression and those associated with low levels of expression. Interestingly, the rapidly highly expressed group of

sequences contains an insertion which is under high selective pressures, suggesting that it may be a novel binding site for this toxin or an important structural change in the final protein.

Positive dosage effects have been found to select for the retention of gene duplicates of highly expressed toxins [45], as an increase in toxins produced generally increase toxicity [151]. Eventually mutation will occur on the duplicates, generating genetic diversity commonly found in toxin gene families [43,45]. The toxins transcripts described here seem to support this idea, as they are highly expressed with several duplication events observed within a highly expressed cluster of sequences. Further, the highly expressed sequences were found to be under positive selection. Similar to Xie et al. [119], we found a region within expressed *Helicops* CTLs that is putative insertion. Our selection analysis found evidence of strong positive selection acting on many sites within this region. This region is thought to be key for changes to protein functionality [119], and our evidence of strong positive selection supports this idea. It is likely this region is of high importance in prey capture, as this cluster of CTLs transcripts comprise most of the toxins these snakes express. The effects of venom from *Helicops angulatus* had been examined previously on mice, but the neurotoxic effects identified were not associated with CTLs and instead attributed to a cysteine-rich secretory protein [65]. However, mice are not a dietary item for *H. angulatus,* which specializes on fish and frogs [115,116]. It is possible that the insertion in *Helicops* CTLs may produce proteins with neurotoxic effects, but full characterization of the functions of these toxins on appropriate prey models has yet to be done.

Despite being lowly-expressed, it appears that the ancestral monomeric CTLs are being maintained, as they are under purifying selection. Several of the ancestral CTLs recovered maintain a classic codon motif EPN, which is known to be a sugar binding site [21]. Additionally, one transcript with a QPD motif was recovered from *H. polylepis*, which is another

motif commonly found in ancestral CTLs [119]. It is unclear how or why these toxins are being maintained. Ancestral CTLs could be maintained due to a possible ontogenic shift in venom composition, which has been observed in several species of snake [13,26,27,168]. It is possible that *Helicops* species relies on a different venom composition to capture different prey items when it is younger. To date, the venom of juvenile *Helicops* species has not been characterized. Alternatively, the emergence of the novel insertion may have been recent, and not enough time has passed for the ancestral CTLs to be removed from the genome. However, the possibility remains that the ancestral CTLs are in fact housekeeping genes. The function of viperid ancestral CTLs has been investigated [169], but this work has not been extended to colubrid snakes [31]. There have been issues in the past where observed gene expression has been misattributed to toxins despite the transcript being expressed in similar levels throughout various tissues [170,171]. For example, phospholipase b and ficolin, among others, are considered venom components despite being expressed in multiple tissues in several species of snakes [170]

Our results show how gene family and gene expression interact to generate changes in venom composition. The change in gene expression and possible increase in the retention of duplicated genes appears to be linked with the novel insertion found within the highly diversified sequences. We propose that the insertion represents a neofunctionalization event, which has changed the toxicological effects of CTLs on their prey in such a way that they have become the dominate venom component in *Helicops*. However, it is unclear what the function of this novel CTL form is. Further, it is unknown if this insertion is restricted to *Helicops* or is more widely distributed within colubrid snakes. Studies of colubrid snake venoms have become more common recently [31,58,59,128,172], yielding discoveries of novel venom components and phenotypes [40,129]. However, this body of work pales in comparison to medically-significant

viperid and elapid snakes, despite colubrid snakes comprising over half of all snakes [31]. More studies of colubrid venom will yield insights into how the interaction of gene family evolution and gene expression drive the evolution of ecologically-driven traits.

3.6 Author Contributions

Conceptualization, P.A.C., T.D. and A.D.R.; Methodology, P.A.C. and A.D.R.; Formal Analysis, P.A.C.; Investigation, P.A.C.; Resources, P.A.C. and A.D.R.; Data Curation, P.A.C. and H.M.M.; Writing – Original Draft Preparation, P.A.C.; Writing – Review & Editing, P.A.C., H.M.M., T.D. and A.D.R.; Visualization, P.A.C., H.M.M. and A.D.R; Supervision, T.D. and A.D.R.; Project Administration, A.D.R.; Funding Acquisition, P.A.C. and A.D.R. All authors have read and agreed to the submitted version of the dissertation chapter. P.A.C = Peter A. Cerda, H.M.M = Haley M. Martens, T.D. = Thomas F. Duda Jr., A.D.R. = Alison Davis Rabosky

Table 3-1 Sample collection and reference information

Sample collection and reference information. UMMZ = University of Michigan Museum of Zoology, MUSM = Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, EBLA = Estación Biológica Los Amigos, EBMS = Estación Biológica Madre Selva,

Taxon	Museum Accession No.	Date Capture	Station, Country	GenBank SRA Numbers
Helicops angulatus 3440	UMMZ 24887	2 December 2018	EBLA, Peru	SRS13333645
Helicops angulatus 3559	MUSM 39826	9 December 2018	EBLA, Peru	SRS13333646
Helicops leopardinus 1812	UMMZ 246808	18 January 2017	EBMS, Peru	SRS13333647
Helicops polylepis 1932	UMMZ 246809	18 January 2017	EBMS, Peru	SRS13333648

Table 3-2 Expression of toxin components

Expression of toxin components. The total toxin expression is the percent of the venom gland transcriptome belonging to venom gene families as calculated from TPM values. The percent of each venom gene family is the fraction of total toxin expression each respective venom component is responsible for. CTL = C-type lectins, CRiSP = cystine-rich secretory proteins, SVMP = snake venom metalloproteinase, - = > 1% or not recovered

Taxon	% Toxin expression	% CTL expression	% Kunitz expression	% CRiSP expression	% SVMP expression
Helicops angulatus 3440	82.02%	82.27%	16.33%	1.33%	-
Helicops angulatus 3559	83.45%	91.65%	4.60%	1.67%	1.88%
Helicops leopardinus 1812	81.37%	90.2%	9.32%	-	-
Helicops polylepis 1932	42.00%	91.51%	6.14%	-	1.73%

Table 3-3 Omega and log likelihood results of CodeML branch site model test.

Model	Omega	Ln(L)	P-value
Null: One Rate	W = 0.106	-4558.891	-
Alternative: Two Rates	w1 = 0.102 w2 = 10.788	-4554.801	0.004

Table 3-4 Number of sites found to be under positive selection for MEME and FUBAR (+), number of sites under purifying selection (FUBAR -), and number of sites under selection within the insertion region.

Number of sites found to be under positive selection for MEME and FUBAR (+), number of sites under purifying selection (FUBAR -), and number of sites under selection within the insertion region.

	MEME	FUBAR +	FUBAR -	BOTH
Whole alignment	27	22	17	13
Insertion	18	19	0	13

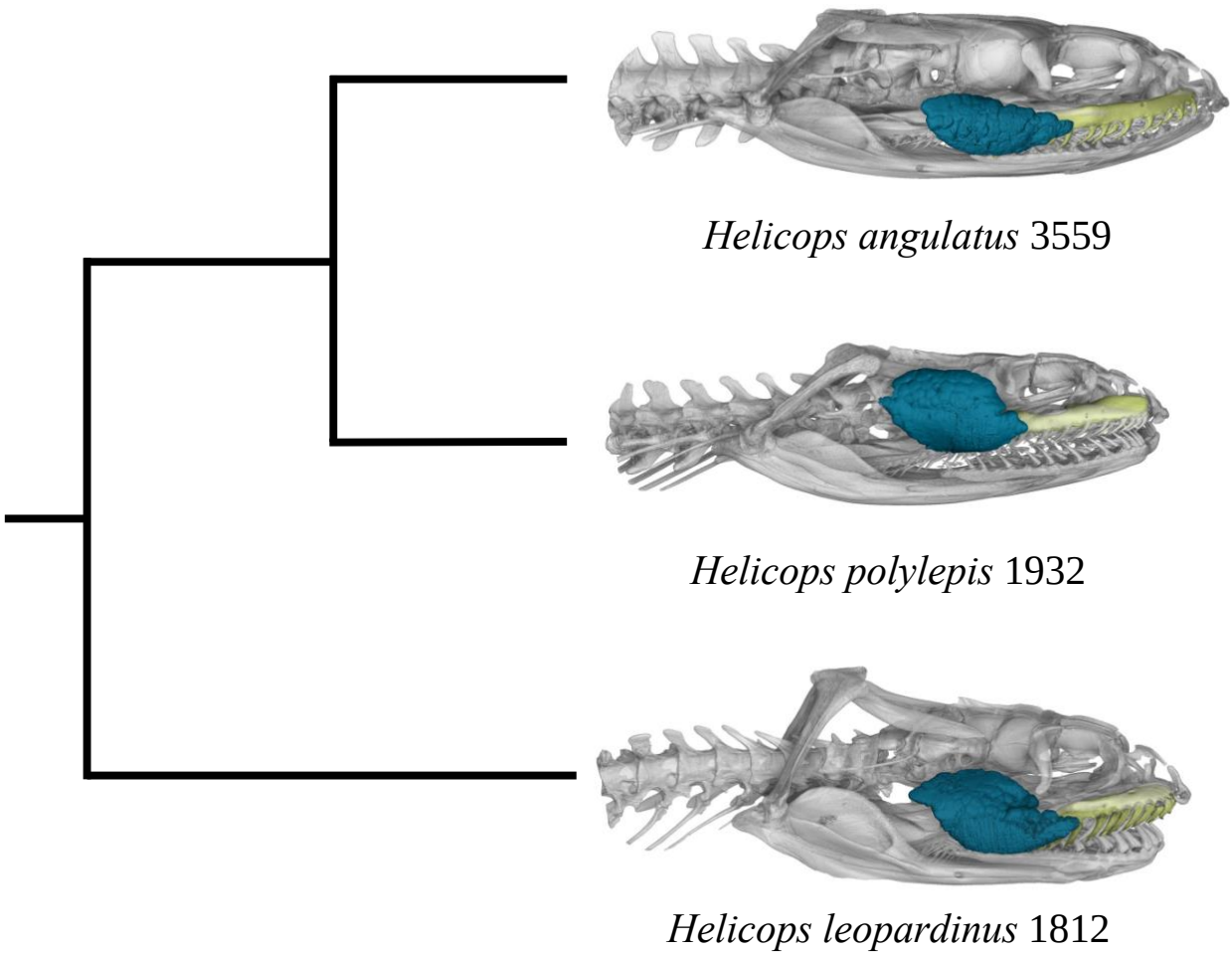

Figure 3-1 diceCT scans showing the venom delivery systems.

diceCT scans showing the venom delivery systems (venom glands, blue, and maxillary bones, yellow) for A) *Helicops angulatus* 3559, B) *Helicops polylepis* 1932 and C) *Helicops leopardinus* 1812. Phylogeny modified from Cerda et al.[40].

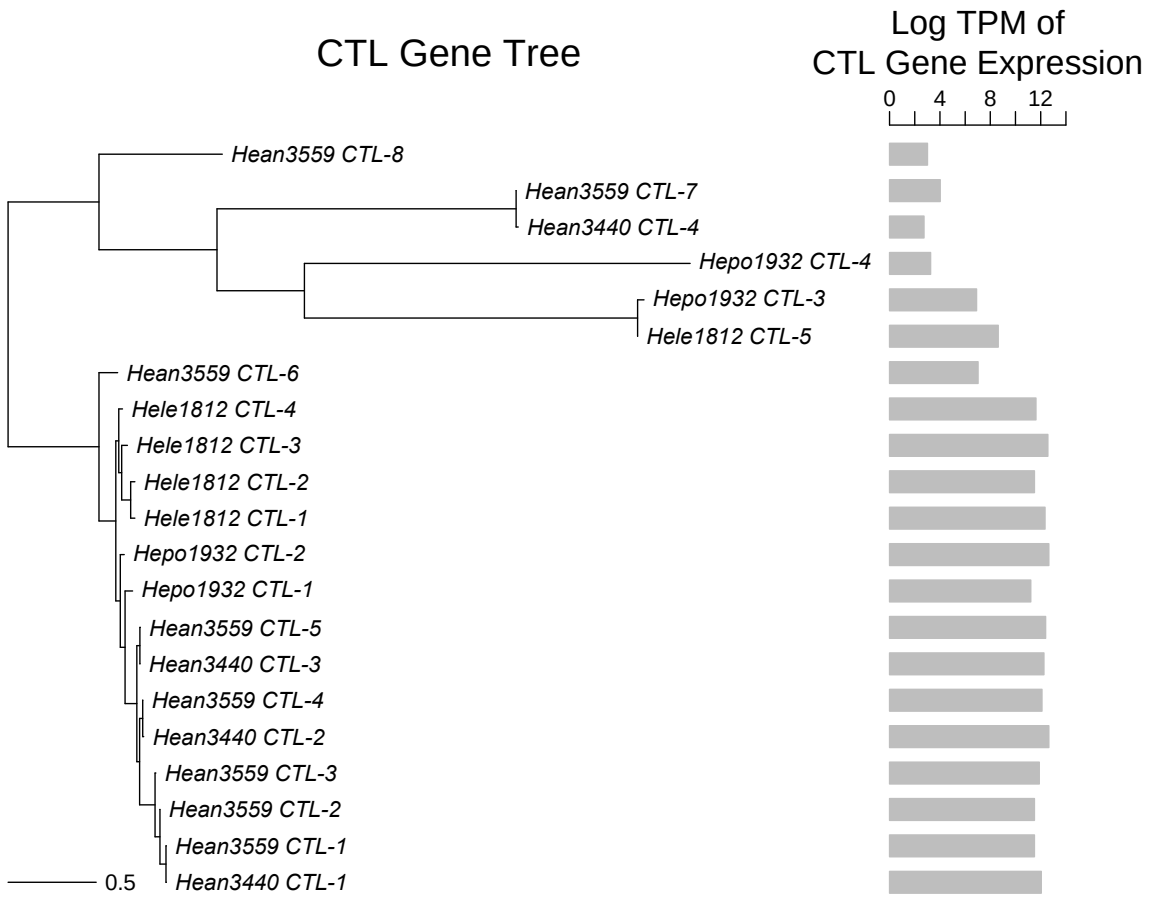

Figure 3-2 Gene tree of recovered C-type lectin transcripts.

Gene tree of recovered C-type lectin transcripts. Log TPM values are plotted with each respective transcript. Hean3440 = *Helicops angulatus* 3440, Hean3559 = *Helicops angulatus* 3559, Hele1812 = *Helicops leopardinus* 1812, Hepo1932 = *Helicops polylepis* 1932

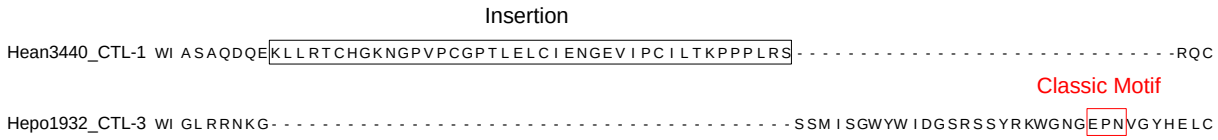

Figure 3-3 Representative sequences of the highly divergent and ancestral clades.

Representative sequences of the highly divergent and basal clades. Hean3440_CTL-1 is a member of the highly divergent clade which features an insertion (within black box) only described from *Helicops* species. Hepo1932_CTL-3 is a member of the basal clade and features a classic codon motif (within red box) found in CTLs of other species.

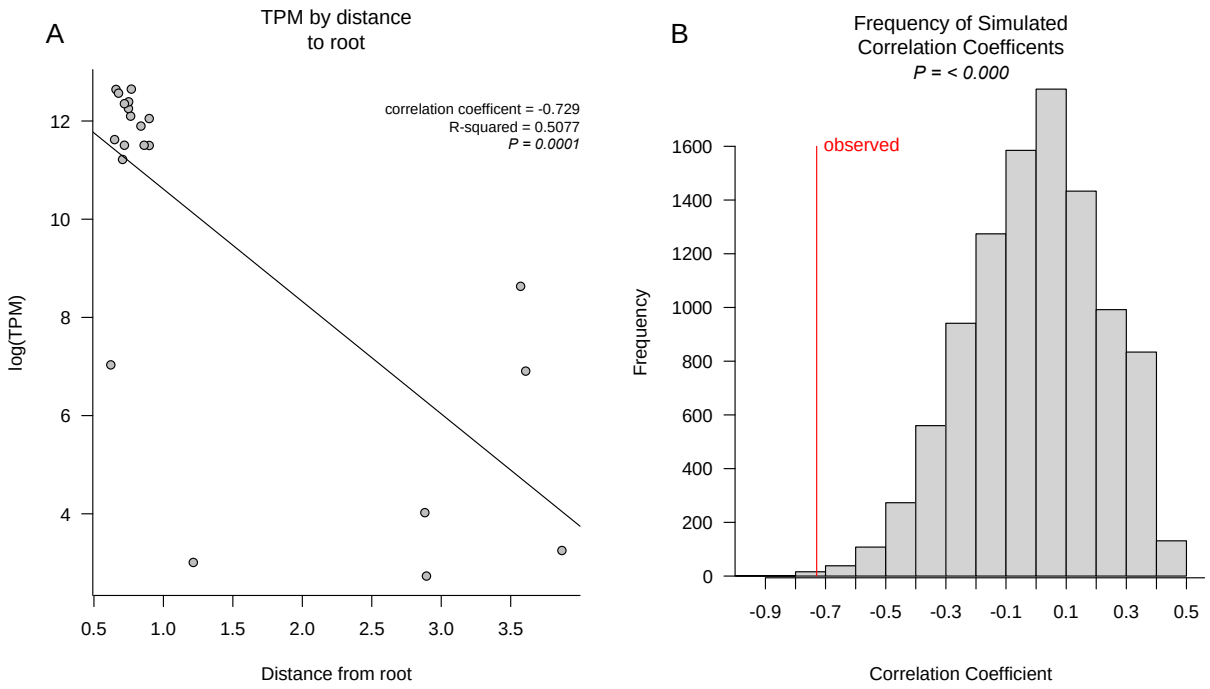

Figure 3-4 Comparison of observed correlations between branch length and expression levels to randomized permutations.

Comparison of observed correlations between branch length and expression levels to randomized permutations. A) Log TPM values for each tip taxa correlate negatively with their distance from the root. B) Distribution of randomized correlation coefficients comparing log TPM and distance from root for 1000 permutations, with the observed correlation coefficient point in red (observed correlation coefficient = -0.729; $r^2 = 0.507$ $P = 0.0101$).

Chapter 4 Evolution of Three Finger Toxin Genes of Neotropical Colubrine Snakes

4.1 Abstract

Snake venom research has historically focused on front-fanged species (Viperidae and Elapidae), limiting our knowledge of venom evolution in venomous rear-fanged snakes across the ecologically-diverse phylogeny. Three finger toxins (3FTxs) are a common neurotoxic component in the venoms of rear-fanged snakes (Colubrinae, Colubridae), but it is unclear how prevalent 3FTxs are both in expression within venom glands and more broadly among colubrine species. Here, we used a transcriptomic approach to characterize the venom expression profile of colubrine snakes from the Neotropics. All species had high levels of overall toxin expression that was dominated by neurotoxic 'non-conventional' 3FTxs, encompassing 93-99% of all toxins expressed by all individuals. To determine the evolutionary history of 3FTxs, we constructed a gene tree which revealed rapid diversification and suggests that gene duplication events are occurring with lineages. Further, we found evidence of potential novel heterodimer interactions in 3FTx protein structure found in *Chironius multiventris.* This potentially represents a case of parallel evolution of heterodimeric 3FTxs found in previously studied colubrines. Selection tests on 3FTx sequences show patterns of strong positive selection, particularly within regions that are suspected to be important in binding to target sites in prey which ultimately causes the neurotoxic effects. This work shows the importance of determining patterns of snake venom evolution across all snakes.

4.2 Introduction

Colubridae is a large family of snakes that includes approximately half of all extant snake species (i.e., >2000 species; [22]). There are roughly 700 species of venomous rear-fanged colubrine snakes, which is comparable in species richness of front-fanged and medically important Viperidae and Elapidae snake families (~700 species total; [22,33,56]). Until recently, venoms of colubrid snakes have not been studied extensively. This has been attributed to the low incidence of lethal envenomation to humans, as well as the difficulties in extracting venom from the low-pressure venom system of rear-fanged snakes [31]. Given these challenges, whole transcriptome approaches can greatly assist in venom studies of colubrids by determining venom composition and variation among and between species [40,58], identifying novel putative toxins [129] and generating new sequence data for evolutionary studies [59,119,173]. Recent analysis or rear-fanged snakes has revealed novel venom components [129], simplified venom compositions [40,58,59], and prey specific toxins [86,174,175]. However, only ~23% of published snake venom transcriptomes are of venom glands from rear-fanged snakes [176], which limits our capacity to fully evaluate the evolutionary history and potential for novel structural changes in venom genes.

Three-finger toxins (3FTxs) are small proteins that are generally non-enzymatic and bind various ligands to induce a variety of biological effects driven by structural differences including neurotoxicity, cardiotoxicity, anticoagulation, and cytotoxicity [22,177]. This toxin superfamily was first discovered in front-fanged elapid snakes [178]. The ‘conventional’ structure consists of eight conserved cysteines that create three β-stranded loops, or “fingers”, stabilized by four disulfide bonds [19,179]. Much of our knowledge of the evolution of 3FTxs comes from elapid snakes, which show high rates of gene duplication [38,180,181], strong positive selection

[38,181], and high functional diversity [19,179]. 3FTxs dominate the venom profiles of several venomous colubrid snakes of the subfamily Colubrinae [31,58,59,174,175], although the extent of this pattern is currently unknown [114]. Colubrine snakes are so far only known to possess non-conventional 3FTxs that feature an additional fifth disulfide bond located in loop I or II [179]. Like their elapid counterparts, colubrine 3FTxs undergo extensive gene duplication within taxa, though patterns of gene family evolution across colubrine species have not been thoroughly investigated [86,119]. Further, while patterns of sequence variation have been observed in 3FTx sequences of several colubrine species, sites that are under selection and may be associated with the functional diversification of these toxins have not been evaluated.

Several colubrine species have 3FTxs with taxon-specific prey toxicity enabled by heterodimer interactions among their proteins. For example, Mangrove cat-snakes, *Boiga irregularis* and *Boiga dendrophilia*, have bird and lizard specific non-conventional 3FTxs, irditoxin and denmotoxin, respectively [174,175]. Despite being phylogenetically distant Colubrinae snakes, prey-specific 3FTxs of *Boiga* species (i.e., irditoxin and denmotoxin) and *Spilotes* species (i.e., sulditoxin) form linked heterodimers which is enabled through formation of disulfide bonds between additional cysteine residues in the first and second loops of the alpha and beta subunits [86,174,175]. These same additional cysteines have been discovered through RNA-sequencing studies of other venomous colubrines such as *Boiga nigriceps, Telescopus dhara,* and *Trimorphodon biscutatus*, which suggest that these species have heterodimeric 3FTxs [182]. Our current understanding is that the heterodimeric interaction has evolved independently in *Boiga* species and other Colubrinae snakes. This suggests that parallel evolution of heterodimeric interactions between 3FTxs has occurred across different colubrine snakes, thought the extent of this pattern is not fully known.

We used a transcriptomic approach to determine the following: How widespread is the 3FTx venom profile among colubrine snakes? What patterns of selection are driving 3FTx evolution? And are there other 3FTx capable of forming novel heterodimeric interactions? We sequenced the venom gland transcriptome of individuals of four rear-fanged species of Neotropical colubrine snakes: *Chironius multiventris*, *Oxybelis aeneus*, *Rhinobothryum bovallii*, and *Spilotes sulphureus*. Except for *S. sulphureus* [86], nothing is known of the venom gland transcriptomes of these species. We included *S. sulphureus* to ensure that our sequencing and transcriptome profiling were effective (i.e., toxins recovered from *S. sulphureus* would group together). In all species examined, we found that venom gland expression was dominated by 3FTxs transcripts. We evaluated the gene trees of 3FTxs and found evidence for potential novel heterodimers in sequences of *C. multiventris*, suggesting that parallel evolution of this structural change may be common in rear-fanged colubrine snakes. We found positive selection at sites within structural loops, indicating these areas may be key binding sites that interact with prey target molecules. Overall, our results highlight the importance of exploring the venoms of understudied species.

4.3 Material and Methods

4.3.1 Specimen selection

Venom glands were excised and sequenced for *Chironius multiventris* (2 individuals), *Oxybelis aeneus* (1 individual), *Rhinobothryum bovallii* (1 individual), and *Spilotes sulphureus* (1 individual). The specimens were collected during field trips to Peru and Nicaragua (Table 4-1). We euthanized captured snakes following approved protocols from the University of Michigan Institutional Animal Care and Use Committee (Protocols #PRO00006234 and #PRO00008306), the Servicio Nacional Forestal y de Fauna Silvestre in Peru (SERFOR permit

numbers 029-2016-SERFOR-DGGSPFFS, 405-2016-SERFOR- DGGSPFFS and 116-2017-SERFOR-DGGSPFFS), and the Ministerio del Ambiente y los Recursos Naturales in Nicaragua (MARENA permit numbers DGPNB-IC-19-2018 and DGPNB-IC-20-2018). We then measured the mass, length, sex, and age of specimens and dissected the venom gland tissues. Whole specimens and tissues were exported to the University of Michigan. Specimens were accessioned at the University of Michigan Museum of Zoology (UMMZ) and Museo de Historia Natural de la Universidad Nacional Mayor de San Marcos (MUSM; Table 4-1).

4.3.2 MicroCT scanning

To visualise the morphology of the venom system, we scanned a representative specimen from each species using a Nikon Metrology XTH 225ST microCT scanner (Xtect, Tring, UK) at the UMMZ. To enhance soft tissue contrast, we stained specimens in 1.25% Lugol's iodine solution following protocols for diffusible iodine contrast-enhanced computed tomography (diceCT) in snakes [108]. The fang bearing maxillary bone and venom gland were segmented in Volume Graphics Studio Max version 3.2 (Volume Graphics, Heidelberg, Germany) using the 'threshold' and 'draw' tools. We noted the position on the maxilla and extent of grooving on the fangs. Figures were prepared using Adobe Illustrator v24.0.2 (Adobe Inc., San Jose, California, USA).

4.3.3 RNA Extraction and Transcriptome Assembly

We extracted venom gland RNA using recommended protocols of the PureLink RNA mini kit (Life Technologies, Carlsbad, CA, USA). We submitted extracted RNA to the University of Michigan Advanced Genomics Core where quality was assessed, and cDNA libraries were constructed and sequenced on an Illumina NovaSeq 6000. We evaluated the

quality of raw sequences using FastQC v0.11.6 [67]. We used Trimmomatic v.0.36 to remove adapters sequencers and low-quality reads [68,69]. We created a de novo transcript sequence reconstruction from RNA-seq using the Trinity v2.6.6 platform for reference generation and analysis [69,183]

4.3.4 Gene Identification and Abundance Estimate

We used TransDecoder v1.0.3 to find open reading frames and then annotated our sequences using BLASTp against known protein coding genes in the Uniprot database [69,80]. We used CD-HIT to cluster similar sequences [184]. We used RSEM v1.2.28 [185] and Bowtie v2.3.4.1 [83] to estimate the abundance of genes expressed and calculated Fragments Per Kilobase Million (FPKM) and Transcripts Per Million (TPM). We used FPKM to compare transcript abundance within individuals and used TPM to compare transcript abundance among individuals [83,185]. We created a custom nucleotide sequence database by downloading venom gland transcriptomes and snake venom protein sequences from NCBI GenBank [58,70,84,85,87,88]. We used this custom database to identify which sequences were toxins in our transcriptomes via BLASTn [186]. We used R [89] to write a custom script to annotate our nucleotide sequences, identify toxin gene family, and pair sequence identity with abundance values.

4.3.5 Gene tree construction

To evaluate the relationships of non-conventional 3FTx sequences, we aligned our putative 3FTx sequences with 3FTx sequences from other rear-fanged species (Colubrinae, Colubridae) as well as front-fanged families (Elapidae and Viperidae) recovered from GenBank (Accession numbers listed in Table S 4-5). We aligned sequences with the Multiple Sequence

Comparison by Log-Expectation (MUSCLE) [187] via the MUSCLE plugin in Geneious Prime v.2021.0.3 [188]. We checked all alignments for ambiguities by eye and removed erroneous sequences. The nucleotide alignment contained 86 total sequences of which 22 were from our samples. To identify conserved cysteines responsible for the loops indicative of non-conventional 3FTx sequences, we translated nucleotide alignments from this subset using the 'translate' function and MUSCLE plugin in Geneious. The appropriate partitioning schemes and best-fit models were selected using PartitionFinder v2.1.1 [189] with branch lengths linked and utilization of the greedy search algorithm. A gene tree was built using Bayesian analyses with the MrBayes v.3.2.6 [190] plugin in Geneious, which uses Markov chain Monte Carlo methods to estimate the posterior distribution of model parameters. This was run for one million generations, sampling every 1000 generations with the first 10% of sampled trees discarded as burn-in, and lset rates set to 'invgamma'. We also generated a maximum likelihood gene tree using the RAxML plugin in Geneious with 1,000 non-parametric bootstrap replicates. We set a sequence of *Protobothrops mucrosquamatus* as a viper outgroup sequence for the colubrine and elapid sequences.

4.3.6 Selection tests

To test for selection acting on 3FTx transcripts recovered from species examined here, we used the 3FTx MUSCLE alignment and Bayesian gene tree of our sampled 3FTx sequences as inputs for easyCodeML [191], which uses CodeML to test models of selection acting on sites [163,192]. We made comparisons of three pairs of site models to determine if and how selection was acting on specific codon sites. The first comparison was M0 (one ratio), which can determine if selection is acting on a sequence, and M3 (discrete) which is constrained to determining if neutral evolution or negative selection is acting on sites. The second comparison

was of M1 (neutral) and M2 (selection), in which M2 allows for positive selection to act on the sites. Finally, we compared models M7 (beta) and M8 (beta&omega), which are similar to M1 and M2, respectively. However, the two pairs of models differ in that M7 and M8 allow for there to be a distribution of beta values. We also used the webserver datamonkey.org (accessed 12/15/2022) to conduct two analyses implemented in HyPhy 2.5 [165]: Mixed Effects Model of Evolution (MEME; [166]) and Fast Unconstrained Bayesian Approximation (FUBAR; [167]). MEME test for evidence of episodic selection at codon sites while FUBAR test for evidence of passive or purifying evolution at codon sites.

4.4 Results and Discussion

4.4.1 Colubrine venom transcriptomes are dominated by 3FTxs

We used a transcriptomic approach to characterize the venom gland transcriptome of individuals of four Neotropical colubrine species: *C. multiventris, O. aeneus, R. bovallii,* and *S. sulphureus* (Table 4-1). The toxin transcriptomes included putative toxin genes from multiple venom superfamilies: 3FTxs, cysteine-rich secretory proteins (CRiSP), phospholipase A2s (PLA2s), snake venom metalloproteinases (SVMPs), and snake venom serine proteases (SVSPs; Table 4-2). While many toxin transcripts were recovered, non-conventional 3FTxs dominated the expression profiles of the rear-fanged colubrines (Figure 4-1). Total 3FTx expression represents 93-99% of all toxin transcript expression totaled, with the top three to five transcripts expressed being 3FTxs (Figure 4-2). The highest-ranking 3FTx transcript had more than one order of magnitude greater relative abundance (FPKM) than the highest-ranking non-3FTx transcript (Figure 4-2).

The comparatively simple venom gene expression profiles of our study species are consistent with what has been observed for other rear-fanged colubrines [40,58,126,193]. Studies

of *Oxybelis* show a simple venom proteome comprising 3FTXs (including fulgimotoxins from *Oxybelis fulgidus*), SVMP III, L-amino acid oxidase, and CRiSP [194]. While we recover members of those toxin families from *O. aeneus*, the transcriptome of this species was dominated by transcripts of 3FTxs (Figure 4-2). Finally, there are no proteome or toxicity data currently available for the venoms of *C. multiventris*, but their expressed venom composition is very similar to other colubrines [58,59,86,195].

The venom proteome of *R. bovallii* is dominated by 3FTx (86.5%) followed by CRiSP (8.2%) and SVMP III (2.4%) [196]. Previous transcriptome profiling of *S. sulphureus* venom glands also found that 3FTxs were highly expressed, albeit at a lower percentage compared to this study (60% vs 99%), and as a large portion of the total venom gland proteome (92% 3FTx, 6% CRISP, 1% SVMP III; [86]). It is interesting to note that venom profiles of *R. bovallii* and *S. sulphureus* specimens were similar to those of previously analyzed conspecifics [86,196]. Our specimens were collected in Nicaragua and Peru, respectively while the *R. bovallii* analyzed by Calvete et al. [196] were wild caught in Costa Rica and the *S. sulphureus* individual studied by Modahl et al. [86] was a captive purchase originally sourced from Suriname. Geographic variation is a common phenomenon in snake venoms [23] and these comparisons suggest a uniform venom profile within these two species of colubrine snakes. The lack of geographic variation has been found before in elapid species which use neurotoxic venom components [4], though this pattern would need to be confirmed in taxa examined here by studies of individuals sourced throughout their South American range.

3FTxs are functionally diverse across and within species, causing numerous biological effects in prey, such as neurotoxicity or cardiotoxicity, though much of our knowledge is based on elapid snake toxins [86,181,197]. The functional diversity is due to mutations in the binding

site which alter target specificity [19]. Currently, our understanding of the functions of colubrine 3FTxs is that they bind and block skeletal muscle's nicotinic acetylcholine receptors (nAChRs), causing flaccid paralysis [174,175,198]. It is not known if other functional characteristics exist among colubrine 3FTx. Further, individual 3FTxs may prove to become more effective in different taxa, resulting in taxa specific 3FTxs [86,174,175]. These types of toxins have already been described in several colubrine snakes, including *S. sulphureus* [86,174,175]. Given the taxon-specific nature of several *S. sulphureus* 3FTxs [86], functional studies should also be performed on the venoms of colubrine snakes to determine if functional variation exists.

4.4.2 *Parallel evolution of heterodimeric sequence*s

To investigate the evolutionary history of 3FTxs in our specimens, we built gene trees from the putative 3FTx transcripts we recovered and currently available 3FTx sequences from several other colubrines, elapids, and a single viper species as an outgroup (Figure 4-3). The tree topology was generally consistent between Bayesian (Figure 4-3) and maximum likelihood methods (S.Fig. 4-1). Viperid and elapids sequences remained the outgroup to colubrid sequences. Colubrid sequence relationships remained constant between the two trees, with the exception being that the major polytomy found in the Bayesian tree was weakly resolved in the maximum likelihood treed. As the topologies were similar between the two methods, we will only describe the Bayesian tree. As expected from previous studies of 3FTxs [86,119], colubrine sequences occurred uniquely together in a group that clusters with a group of all elapid sequences (Figure 4-3). Our inability to resolve a large polytomy is possibly due to the short sequence length and rapid rate of evolution which has impacted other studies of the evolutionary history of 3FTxs [38,199]. Sequences from the same individual or species often clustered together (Figure 4-3), suggesting that lineage (species) specific duplication events have occurred.

Gene duplication plays an important role in the evolution and diversification of snake toxin gene families [8,38,200], yielding material for mutation and selection to act on to potentially diversify toxin function via neofunctionalization [42,201]. Evidence of gene duplication has been found in several colubrine species. However, these studies have largely been transcriptomic [86,119,202], and rely on the capture of expressed toxin sequences. To determine the full extent of gene duplication in these species, a genomic approach should be taken to uncover the full repertoire of 3FTx loci in genomes.

Tree topology showed interesting patterns of 3FTx evolution as there is evidence for potential homologous 3FTx sequences with known taxon-specific toxicity in other rear-fanged snakes (highlighted blue – Figure 4-3). Some transcripts from our *S. sulphureus* clustered with known taxon-specific 3FTxs previously identified for the same species: sulditoxin A (transcript ID: 3FTx-5) and sulmotoxin-1 (transcript ID: 3FTx-4), but not sulditoxin B (Figure 4-3). It is possible that our individual did not expresses a homolog of sulditoxin B, or we were unable to recover the transcript. None of our sequences clustered near the taxon-specific 3FTxs of *Boiga* cat snakes (i.e., denmotoxin, irditoxin A or irditoxin B). We were unable to recover a transcript from our *O. aeneus* that fell out near fulgimotoxin from closely-related *O. fulgidus*, suggesting that fulgimotoxin may be independently derived, though investigations of more *Oxybelis* species and individuals should be done before making such a conclusion.

To determine the presence of putative heterodimeric 3FTxs in our samples, we translated the complete 3FTx coding sequences and compared translations to amino acid sequences of 3FTxs that have known toxic effects in other studies (i.e., sulditoxin A and B and sulmotoxin-1, irditoxin A and B, and denmotoxin) [86,174,175] (Figure 4-4). The β-stranded loops were annotated based on conserved patterns of occurrence of cysteines that create the five distinctive

loops present in non-conventional 3FTx. The amino acid structure of colubrine 3FTxs appears to be highly conserved across species (Figure 4-4). However, several recovered transcripts lacked a cysteine for Loop I (Figure 4-4). This suggest that these sequences maybe nonfunctional as they lack the ability to form the 3FTx structure.

Sulditoxin A/B and irditoxin A/B possess extra cysteine residues at positions 59 and 86, respectively, which are known to be a part of the heterodimeric interaction between A and B subunits. Some sequences from our *C. multiventris* specimens also have additional cysteines (Figure 4-4) in positions similar to those in both irditoxin and sulditoxin A and B subunits, suggesting that expressed proteins form heterodimers. The heterdimeric *C. multiventris* sequences occur closely with their sulditoxin A and B counterparts in the gene tree (Figure. 4-3), implying that this heterodimeric interaction is conserved among members of this clade of colubrine snakes. Additionally, an *O. aeneus* sequence, which also contains an extra cystine residue, groups closely with sulditoxin A and putative *C. multiventris* heterodimeric sequences. Our results further support the notion that this heterodimeric interaction evolved independently from the similar structure and interaction observed in *B. irregularis* [119,195]. Because the *C. multiventris* sequences also reveal the potential that their expressed products also form heterodimers, these products may taxon-specific toxicity, similar to heterodimeric sulditoxins (which are lizard specific [86]) and irditoxin (which are bird and lizard specific [174]). Functional studies of venoms of *C. multiventris* should be performed to determine if heterodimeric toxins are present and if they are taxon specific.

4.4.3 Colubrine 3FTx toxins are under high positive selection

To determine if codon sites in recovered 3FTx sequences are under selection, we tested several models of evolution on recovered transcripts. All CodeML model comparisons were

found to be significant, indicating these sequences are under positive selection (Table 4-3). We found evidence of positive selection operating at numerous sites, particularly within the second loop (Table 4-4). This finding is similar to what Modahl et al. [86] observed, which suggests that loop II may be a main binding site for target nAChRs. We also found evidence of positive selection within other loops, namely loop V, in which all sites within the loop are under selection.

Xie et al. [119] found two types of N-terminus regions within 3FTx sequences recovered from rear-fanged snakes: a short region restricted to *Boiga* species and a long one that appears to be widespread in colubrine snakes. We did not recover any sequences that would express a short N-terminus region. However, all but two of the transcripts we recovered contained sequences that would express peptides with the long N-terminus. This region may serve a critical function and be under strong selective pressures as it arose and was retained across several species [119]. Interestingly we found several sites under positive selection located in the long N-terminus region. Functional test should be performed before confirming that this region is involved with the envenomation process.

4.5 Conclusion

Our analysis of several colubrid snake venom gland transcriptomes reveled high levels of 3FTx expression across several distantly related species. Further, we identified several putative heterodimeric 3FTx sequences and show evidence of parallel evolution of this interaction. Our selection analysis points to areas within the sequence that show signs of strong positive selection. While our results show evidence of parallel evolution, only protein purification will truly determine if putative heterodimeric sequences form a heterodimer. Further, functional studies of the heterodimeric proteins should be performed on potential prey to determine how effective the

toxin is. Additionally, dietary analysis should be conducted on these species to determine if these species, which have a simple venom, also have small dietary breadths as shown in North American pit vipers [37].Our transcriptomic analysis focused on highly expressed 3FTxs, but it is unknown how other venom components are evolving within these colubrid snakes. Analysis should be performed to detect if selection is acting on lowly expressed toxin sequences.

4.6 Author Contributions

Conceptualization, K.S., P.A.C., J.M.C.R., and A.D.R.; Methodology, K.S., P.A.C., and J.C.R.; Formal Analysis, K.S., P.A.C., and J.C.R.; Investigation, K.S., P.A.C., and J.C.R.; Resources, P.A.C. and A.D.R.; Data Curation, K.S., P.A.C., and J.C.R.; Writing – Original Draft Preparation, K.S., P.A.C., and J.C.R.; Writing – Review & Editing, P.A.C., J.C.R., and A.D.R.; Visualization, K.S., P.A.C., J.C.R. and A.D.R; Supervision, A.D.R., J.C.R.; Project Administration, A.D.R.; Funding Acquisition, P.A.C. and A.D.R. All authors have read and agreed to the published version of the dissertation chapter. K.S. = Kristy Srodawa, P.A.C. = Peter A. Cerda, A.D.R. = Alison Davis Rabosky, J.C.R. = Jenna M. Crowe-Riddell

Table 4-1 Specimen information and basic transcriptomic statistics.

Specimen information and basic transcriptomic statistics. SVL = snout to vent length, MUSM = Museo de Historia Natural, Universidad Nacional Mayor de San Marcos, UMMZ = University of Michigan Museum of Zoology, EBVC = Estación Biológica Villa Carmen, RB = Refugio Bartola, EBLA = Estación Biológica Los Amigos.

Specimen	Field No.	Museum No.	SVL (mm)	Mass (g)	Station Captured
Spilotes sulphureus	RAB562	MUSM 37565	1840	1250	EBVC, Peru
Rhinobothryum bovallii	RAB3018	UMMZ 247120	1030	71.4	RB, Nicaragua
Oxybelis aeneus	RAB3102	UMMZ 247113	865	73.1	Momotombo, Nicaragua
Chironius multiventris A	RAB3332	UMMZ 249111	810	160	EBLA, Peru
Chironius multiventris B	RAB3577	MUSM 39803	739	85	EBLA, Peru

Table 4-2 Total Putative Toxin Contigs per Individual

Total Putative Toxin Contigs per Individual. 3FTx = Three finger toxin, CRiSP = Cysteine rich secretory protein, PLA2 = Phospholipase A2, SVMP = Snake venom metalloproteinase, SVSP = Snake venom serine proteinase.

Taxonomy	3FTx	CRiSP	PLA2	SVMP	SVSP	Other
Spilotes sulphureus	41	16	10	29	55	60
Rhinobothryum bovallii	5	7	3	20	57	58
Oxybelis aeneus	11	6	9	9	23	38
Chironius multiventris A	7	21	8	27	58	53
Chironius multiventris B	10	25	4	34	56	55

Table 4-3 CodeML Selection test comparisons

CodeML Selection test comparisons. LnL = log likelihood, LRT = likelihood ratio test.

Model	Ln L	Model compared	LRT P-value
M0	-2330.419270		
M3	-2249.816621	M0 v. M3	$P = < 0.001$
M1	-2280.078815		
M2	-2250.100414	M1 v. M2	$P = < 0.001$
M7	-2291.574607		
M8	-2249.800082	M8 v. M8	$P = < 0.001$

Table 4-4 Number of sites found under positive selection.

Number of sites found under positive selection across the whole alignment, within the N-terminus region, Loop II, or Loop V. MEME = Mixed Effects Model of Evolution, FUBAR = Fast Unconstrained Bayesian Approximation.

	MEME	FUBAR	Both
Whole alignment	30	21	18
N-terminus region	5	3	3
Loop II	6	3	2
Loop V	4	4	4

Table S 4-5 Accession numbers for three finger toxin sequences gathered form GenBank

Taxa	GenBank Accession No.
T. jacksoni	EU029684.1
T. jacksoni	EU029685.1
A. prasina	KU666926.1
T. jacksoni	EU029682.1
D. typus	EU029683.1
S. sulphureus	MH233095.1
S. sulphureus	MH233110.1
S. sulphureus	MH233117.1
B. irregularis	KU666932.1
B. irregularis	KU666930.1
B. irregularis	KU666931.1
B. irregularis	KU666929.1
T. lambda	KU666935.1
T. lambda	KU666936.1
T. biscutatus	EU029677.1
O. fulgidus	KU666924.1
O. fulgidus Fulgimotoxin	C0HJD3.1
T. lambda	KU666937.1
S. sulphureus	MH233111.1
S. sulphureus Sulditoxin A	MH233113.1
S. sulphureus	MH233113.1

S. sulphureus	MH233093.1
S. sulphureus	MH233116.1
S. sulphureus Sulditoxin B	MH233100.1
S. sulphureus	MH233100.1
S. sulphureus	MH233109.1
S. sulphureus	MH233112.1
S. sulphureus Sulmotoxin-1	MH233112.1
S. sulphureus	MH233094.1
S. sulphureus	MH233115.1
S. sulphureus	MH233114.1
T. dhara	EU029686.1
B. dendrophila Denmotoxin	DQ366293.1
B. irregularis Irditoxin B	DQ304539.1
B. cynodon	KU666927.1
B. nigriceps	KU666934.1
B. irregularis Irditoxin A	DQ304538.1
P. mucrosquamatus	XM15823122.1
M. mipartitus	KY635904.1
M. mipartitus	KY635903.1
M. mipartitus	KY635902.1
M. mipartitus	KY635901.1
M. mipartitus	KY635900.1
N. sputatrix	AF098924.1

N. sputatrix	AF026892.1
N. sputatrix	AF026891.1
O. hannah	DQ273583.1
B. candidus	AY057879.1
B. falviceps	GU190800.1
N. atra	AJ223154.1
H. stephensii	DQ917504.1
P. rossignolii	AB778565.1
M. corallinus	AJ344067.1
B. candidus	AY142323.1
B. multicinctus	X51414.1
B. falviceps	GU190797.1
B. falviceps	GU190799.1
B. falviceps	GU190796.1
B. falviceps	GU190798.1

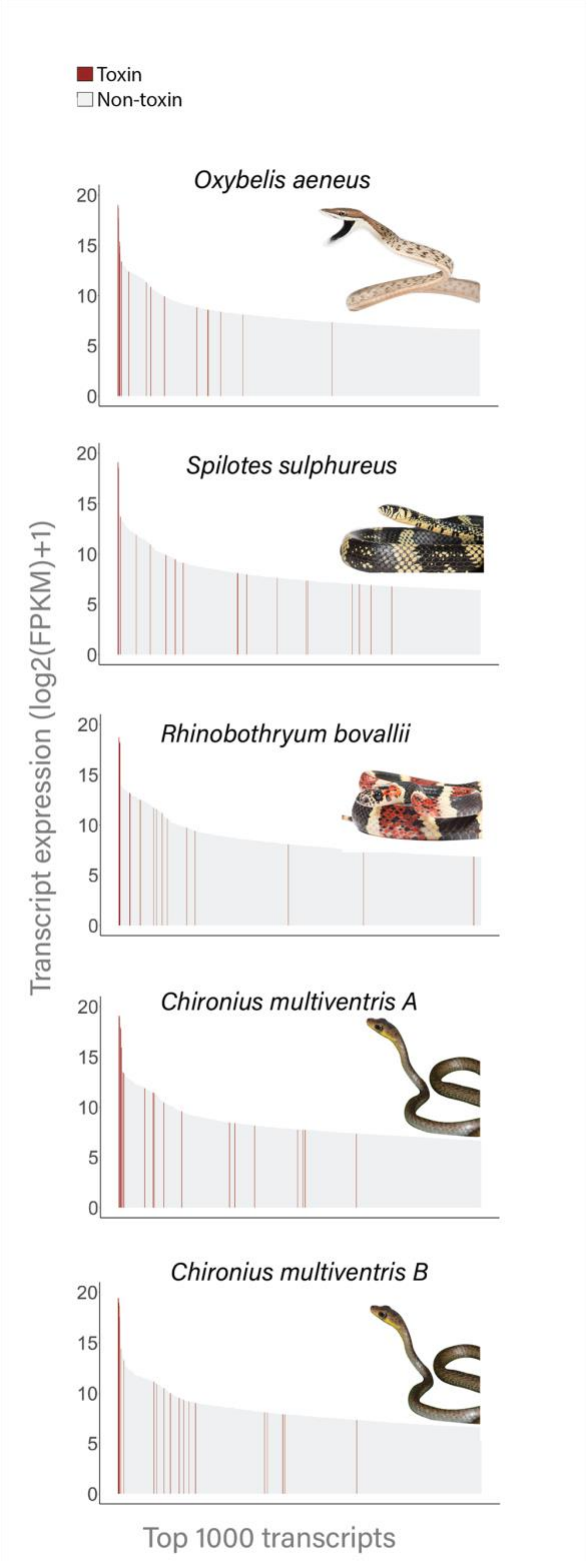
Toxin
Non-toxin
Oxybelis aeneus
Spilotes sulphureus
Rhinobothryum bovallii
Chironius multiventris A
Chironius multiventris B
Transcript expression (log2(FPKM)+1)
0
5
10
15
20
Top 1000 transcripts

Figure 4-1 Venom gland transcriptomes of four colubrine snake species.

Venom gland transcriptomes of four colubrine snake species. Red transcripts indicate toxin sequences. Transcript abundances estimates are in Fragments Per Kilobase Million (FPKM) and only the top 1000 ranked transcripts are shown. Image credits: José G. Martinez-Fonseca (*O. aeneus, S. sulphureus, R. bovallii*) and Consuelo Alarcón Rodríguez (*C. multiventris*).

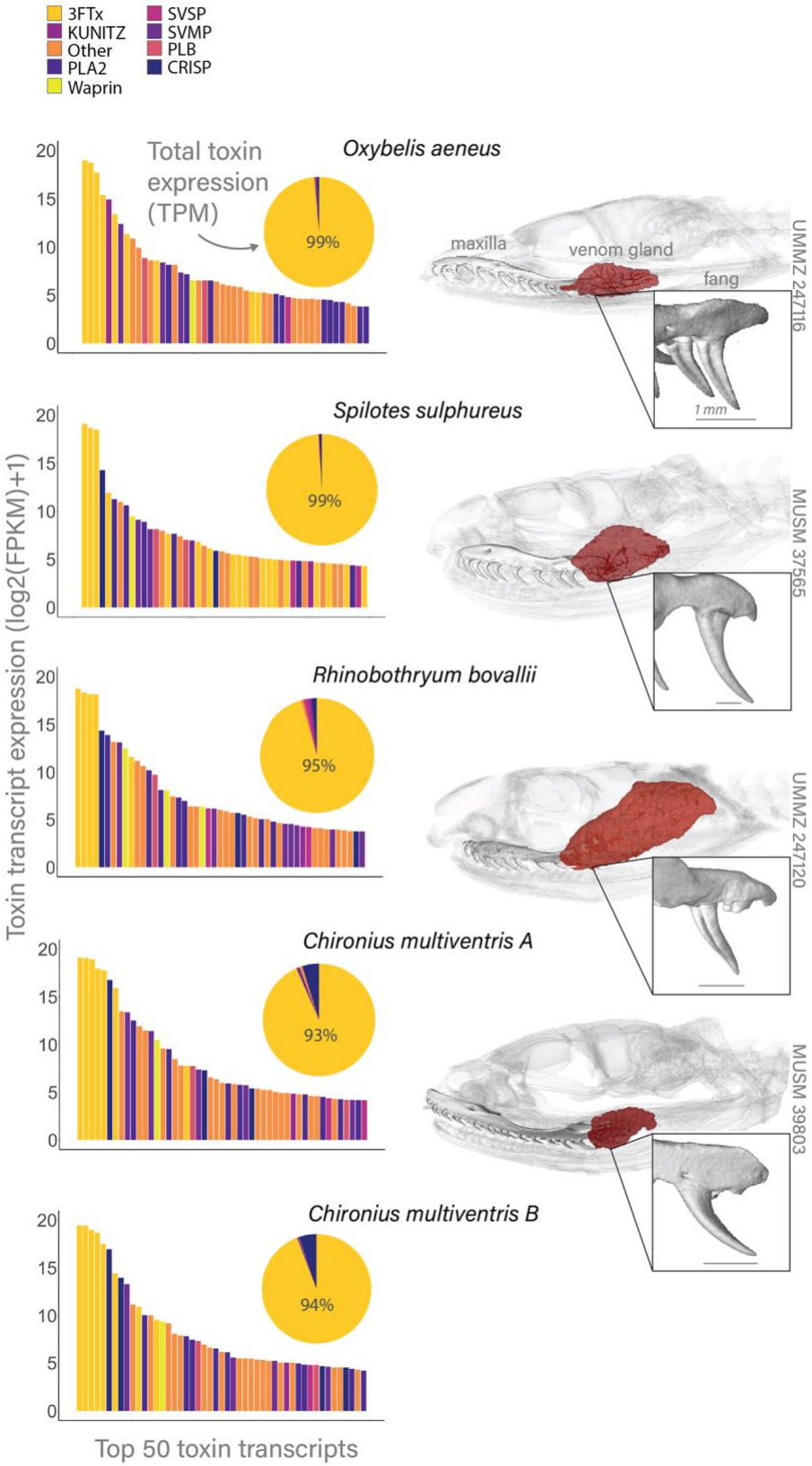
3FTx
KUNITZ
Other
PLA2
Waprin
SVSP
SVMP
PLB
CRISP
Total toxin expression (TPM)
Oxybelis aeneus
99%
maxilla
venom gland
fang
UMMZ 247116
1 mm
Spilotes sulphureus
99%
MUSM 37565
Rhinobothryum bovallii
95%
UMMZ 247120
Chironius multiventris A
93%
MUSM 39803
Chironius multiventris B
94%
Toxin transcript expression (log2(FPKM)+1)
Top 50 toxin transcripts

Figure 4-2 Venom gland toxin expression profile of four species of rear-fanged colubrine snakes.

Venom gland toxin expression profile of four species of rear-fanged colubrine snakes. Bar charts show toxin transcript abundances, measured in Fragments Per Kilobase Million (FPKM), for the top 50 toxin transcripts of each species. Pie charts show proportions of toxin family expression based on Transcripts per Million (TPM). Toxin families are color coded; toxins indicated include three-finger toxins (3FTx), Kunitz-type venom proteins (Kunitz), phospholipase A2 (PLA2), Waprin, snake venom serine proteases (SVSP), snake venom metalloproteinases (SVMP), L-amino acid oxidases (LAAO), and Cysteine-rich secretory proteins (CRiSP). Venom system morphology is shown on the right for each species; obtained from microCT scans of specimens the University of Michigan Museum of Zoology (UMMZ) and Museo de Historia Natural de la Universidad Nacional Major de San Marcos (MUSM).

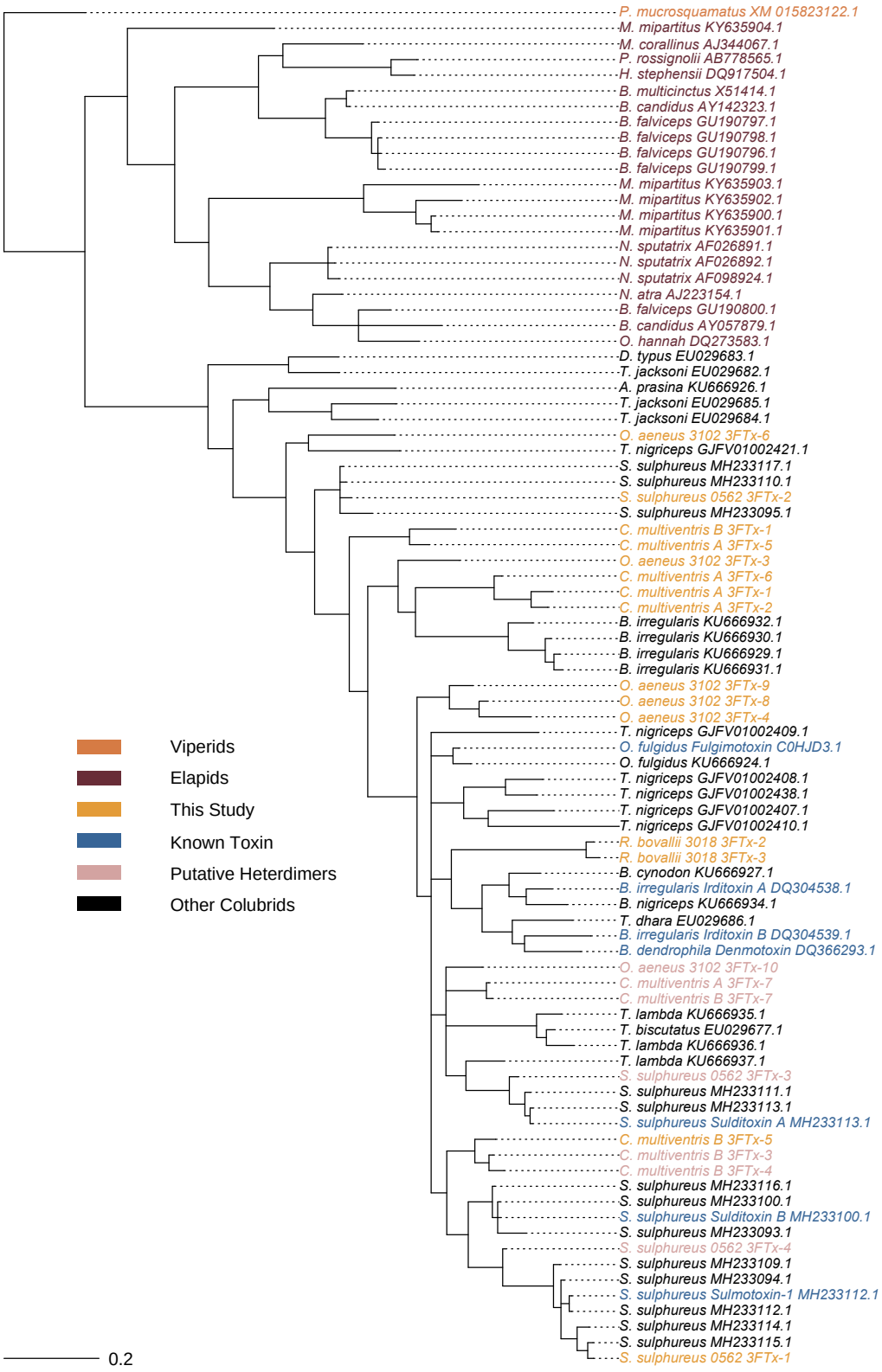
P. mucrosquamatus XM 015823122.1
M. mipartitus KY635904.1
M. corallinus AJ344067.1
P. rossignolii AB778565.1
H. stephensii DQ917504.1
B. multicinctus X51414.1
B. candidus AY142323.1
B. falviceps GU190797.1
B. falviceps GU190798.1
B. falviceps GU190796.1
B. falviceps GU190799.1
M. mipartitus KY635903.1
M. mipartitus KY635902.1
M. mipartitus KY635900.1
M. mipartitus KY635901.1
N. sputatrix AF026891.1
N. sputatrix AF026892.1
N. sputatrix AF098924.1
N. atra AJ223154.1
B. falviceps GU190800.1
B. candidus AY057879.1
O. hannah DQ273583.1
D. typus EU029683.1
T. jacksoni EU029682.1
A. prasina KU666926.1
T. jacksoni EU029685.1
T. jacksoni EU029684.1
O. aeneus 3102 3FTx-6
T. nigriceps GJFV01002421.1
S. sulphureus MH233117.1
S. sulphureus MH233110.1
S. sulphureus 0562 3FTx-2
S. sulphureus MH233095.1
C. multiventris B 3FTx-1
C. multiventris A 3FTx-5
O. aeneus 3102 3FTx-3
C. multiventris A 3FTx-6
C. multiventris A 3FTx-1
C. multiventris A 3FTx-2
B. irregularis KU666932.1
B. irregularis KU666930.1
B. irregularis KU666929.1
B. irregularis KU666931.1
O. aeneus 3102 3FTx-9
O. aeneus 3102 3FTx-8
O. aeneus 3102 3FTx-4
T. nigriceps GJFV01002409.1
O. fulgidus Fulgimotoxin C0HJD3.1
O. fulgidus KU666924.1
T. nigriceps GJFV01002408.1
T. nigriceps GJFV01002438.1
T. nigriceps GJFV01002407.1
T. nigriceps GJFV01002410.1
R. bovallii 3018 3FTx-2
R. bovallii 3018 3FTx-3
B. cynodon KU666927.1
B. irregularis Irditoxin A DQ304538.1
B. nigriceps KU666934.1
T. dhara EU029686.1
B. irregularis Irditoxin B DQ304539.1
B. dendrophila Denmotoxin DQ366293.1
O. aeneus 3102 3FTx-10
C. multiventris A 3FTx-7
C. multiventris B 3FTx-7
T. lambda KU666935.1
T. biscutatus EU029677.1
T. lambda KU666936.1
T. lambda KU666937.1
S. sulphureus 0562 3FTx-3
S. sulphureus MH233111.1
S. sulphureus MH233113.1
S. sulphureus Sulditoxin A MH233113.1
C. multiventris B 3FTx-5
C. multiventris B 3FTx-3
C. multiventris B 3FTx-4
S. sulphureus MH233116.1
S. sulphureus MH233100.1
S. sulphureus Sulditoxin B MH233100.1
S. sulphureus MH233093.1
S. sulphureus 0562 3FTx-4
S. sulphureus MH233109.1
S. sulphureus MH233094.1
S. sulphureus Sulmotoxin-1 MH233112.1
S. sulphureus MH233112.1
S. sulphureus MH233114.1
S. sulphureus MH233115.1
S. sulphureus 0562 3FTx-1
Viperids
Elapids
This Study
Known Toxin
Putative Heterdimers
Other Colubrids
0.2

Figure 4-3 Bayesian gene tree of 3FTx sequences

Bayesian gene tree of 3FTx sequences, including sequences newly generated (gold), previously named and functionally characterized sequences (blue), other colubrine 3FTx sequences from GenBank (black), elapid sequences from GenBank (maroon), and a sequence from a viper species used as an outgroup (orange). For the species *C. multiventris* "A" and "B" were used to differentiate between the two individuals.

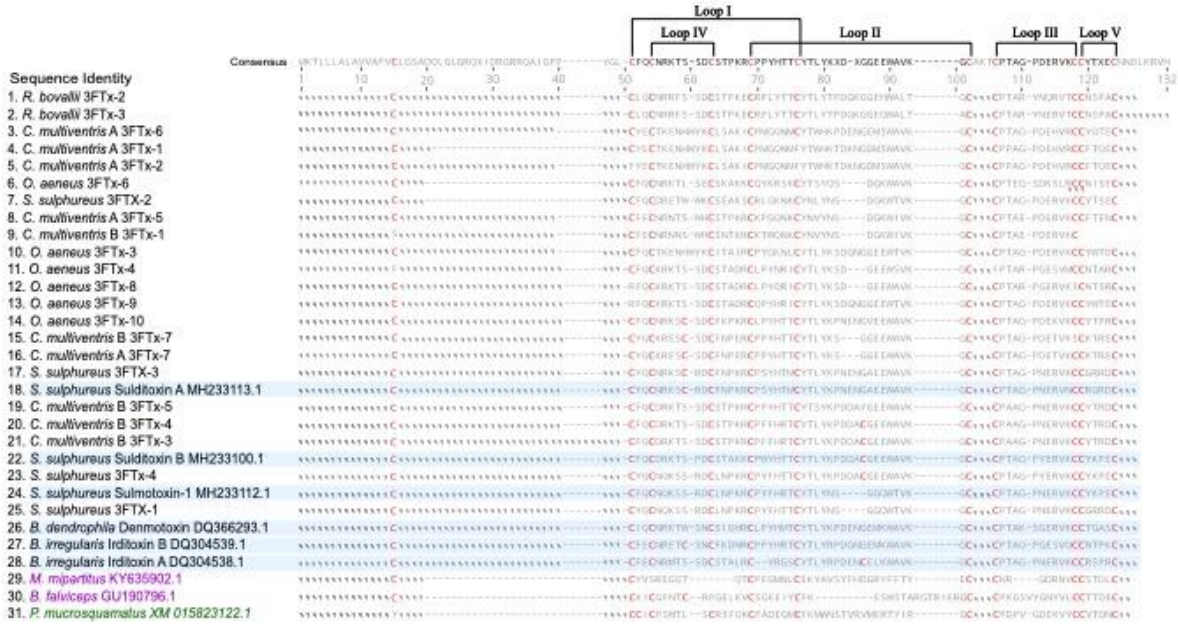

Figure 4-4 Amino acid alignment of non-conventional three finger toxin

Amino acid alignment of non-conventional three finger toxin (3FTx) transcripts generated from this study (black) as well as sequences previously identified and characterized from colubrine snake species (highlighted in blue). We also included two elapid (purple) and one viper (green) sequence to serve as an outgroup. Conserved cysteine residues are highlighted in red; disulfide bonds between cysteine residues result in formation of five loops in non-conventional 3FTx, which are shown with brackets. Loops IV and V are unique to non-conventional 3FTxs.

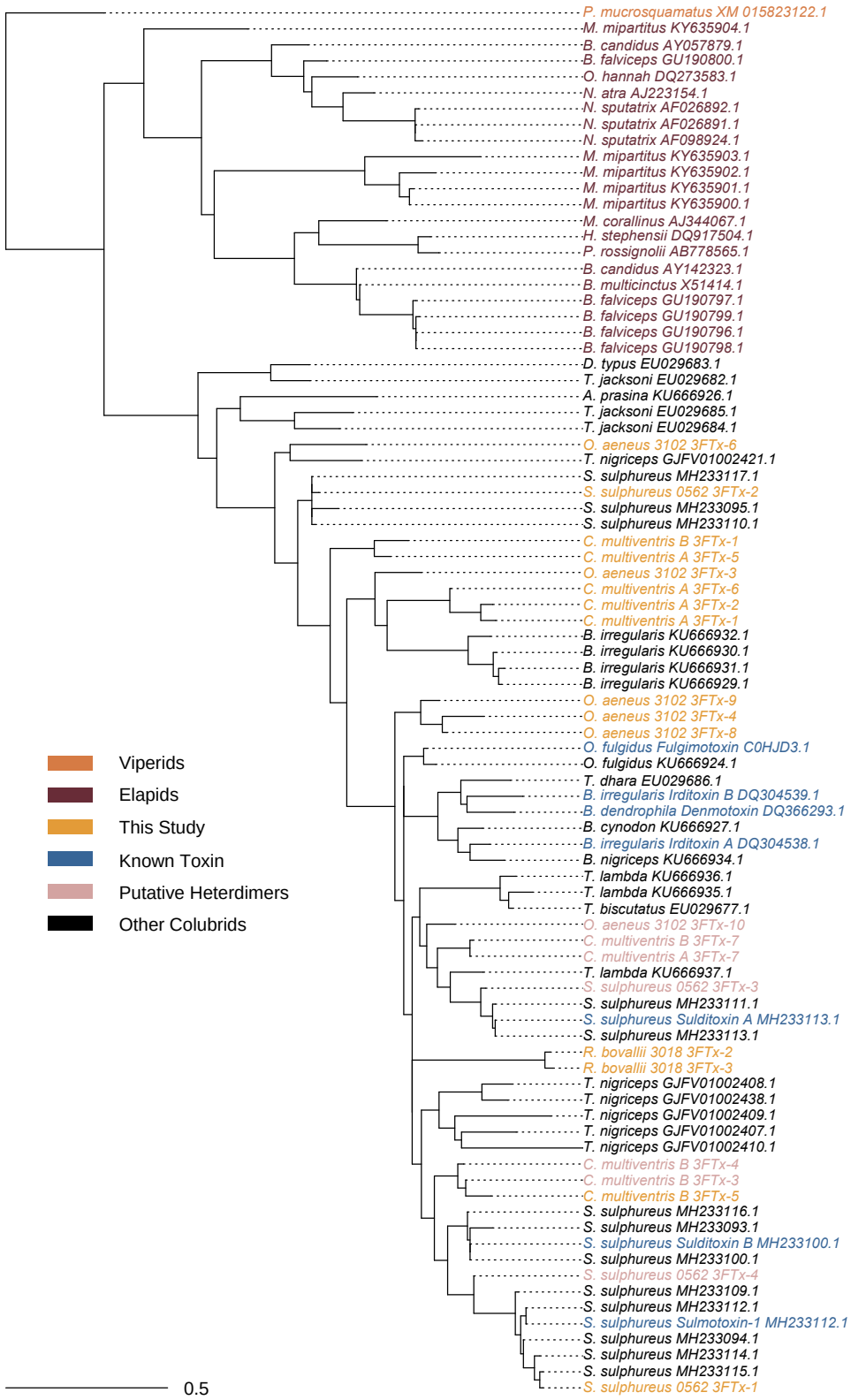
P. mucrosquamatus XM 015823122.1
M. mipartitus KY635904.1
B. candidus AY057879.1
B. falviceps GU190800.1
O. hannah DQ273583.1
N. atra AJ223154.1
N. sputatrix AF026892.1
N. sputatrix AF026891.1
N. sputatrix AF098924.1
M. mipartitus KY635903.1
M. mipartitus KY635902.1
M. mipartitus KY635901.1
M. mipartitus KY635900.1
M. corallinus AJ344067.1
H. stephensii DQ917504.1
P. rossignolii AB778565.1
B. candidus AY142323.1
B. multicinctus X51414.1
B. falviceps GU190797.1
B. falviceps GU190799.1
B. falviceps GU190796.1
B. falviceps GU190798.1
D. typus EU029683.1
T. jacksoni EU029682.1
A. prasina KU666926.1
T. jacksoni EU029685.1
T. jacksoni EU029684.1
O. aeneus 3102 3FTx-6
T. nigriceps GJFV01002421.1
S. sulphureus MH233117.1
S. sulphureus 0562 3FTx-2
S. sulphureus MH233095.1
S. sulphureus MH233110.1
C. multiventris B 3FTx-1
C. multiventris A 3FTx-5
O. aeneus 3102 3FTx-3
C. multiventris A 3FTx-6
C. multiventris A 3FTx-2
C. multiventris A 3FTx-1
B. irregularis KU666932.1
B. irregularis KU666930.1
B. irregularis KU666931.1
B. irregularis KU666929.1
O. aeneus 3102 3FTx-9
O. aeneus 3102 3FTx-4
O. aeneus 3102 3FTx-8
O. fulgidus Fulgimotoxin C0HJD3.1
O. fulgidus KU666924.1
T. dhara EU029686.1
B. irregularis Irditoxin B DQ304539.1
B. dendrophila Denmotoxin DQ366293.1
B. cynodon KU666927.1
B. irregularis Irditoxin A DQ304538.1
B. nigriceps KU666934.1
T. lambda KU666936.1
T. lambda KU666935.1
T. biscutatus EU029677.1
O. aeneus 3102 3FTx-10
C. multiventris B 3FTx-7
C. multiventris A 3FTx-7
T. lambda KU666937.1
S. sulphureus 0562 3FTx-3
S. sulphureus MH233111.1
S. sulphureus Sulditoxin A MH233113.1
S. sulphureus MH233113.1
R. bovallii 3018 3FTx-2
R. bovallii 3018 3FTx-3
T. nigriceps GJFV01002408.1
T. nigriceps GJFV01002438.1
T. nigriceps GJFV01002409.1
T. nigriceps GJFV01002407.1
T. nigriceps GJFV01002410.1
C. multiventris B 3FTx-4
C. multiventris B 3FTx-3
C. multiventris B 3FTx-5
S. sulphureus MH233116.1
S. sulphureus MH233093.1
S. sulphureus Sulditoxin B MH233100.1
S. sulphureus MH233100.1
S. sulphureus 0562 3FTx-4
S. sulphureus MH233109.1
S. sulphureus MH233112.1
S. sulphureus Sulmotoxin-1 MH233112.1
S. sulphureus MH233094.1
S. sulphureus MH233114.1
S. sulphureus MH233115.1
S. sulphureus 0562 3FTx-1
Viperids
Elapids
This Study
Known Toxin
Putative Heterdimers
Other Colubrids
0.5

Figure S 4-5 RAxML tree of putative non-conventional three finger toxin

RAxML tree of putative non-conventional three finger toxin (3FTx) transcripts generated from this study (gold), previously named and functionally characterized sequences (blue), other colubrine 3FTx sequences from GenBank (black), elapid sequences from GenBank (maroon), and a sequence from a viper species used as an outgroup (orange). For the species *C. multiventris* "A" and "B" were used to differentiate between the two individuals.

Chapter 5 Conclusions

In this dissertation, I determined the venom expression patterns and evolution of venom gene families across several clades of snakes. First, I characterized the venom gland transcriptomes of several species of snakes across three families to determine the complexity of venoms across clades. I found that venoms of rear-fanged snakes were comparatively simple, often dominated by a single venom gene family [40]. I leveraged this finding to determine the evolutionary patterns of dominate venom components. I focused on C-type lectins (CTLs), the dominate toxin expression in *Helicops* species. I found both high expression and a novel insertion, which may represent a neofunctionalization event, were associated with high diversification of the gene family. I further determined the patterns of evolution of three finger toxins (3FTxs) in colubrine snakes. I found high rates of selection as well as a putative heterodimeric interaction among three finger toxin proteins. Taken together, my work begins to explore the evolution of rear-fanged snake venoms, a highly variable trait.

This work lays the groundwork for several paths of exploration. First, determining the role dietary preference plays in the evolution of venom is vital, as venom is a trait linked to prey acquisition, and thus species interactions. Prey taxa can evolve resistance to venoms [24,25,49,203], and venom complexity has been linked to dietary breath in some venomous groups [37,47]. As the venom gland transcriptomes tend to be relatively simple, one can hypothesize that their dietary breadth is small. To investigate the evolution of venoms in the absence of organismal ecology will result in an incomplete picture of venom evolution.

While inferences can be made about venom protein function based on family membership, it is important to determine the actual function of venoms on the appropriate prey item, as venom is evolving to interact with those taxa [37]. As much of our information about the functions of venoms comes from a humancentric focus, it is important to determine how these toxins incapacitate their prey. For example, both toxin families investigated in this dissertation (CTLs and 3FTxs) are known to be functionally labile [17,19], and thus a hidden level of functional complexity beyond a coarse grouping of family membership can exist.

The work described here focused on highly expressed toxins recovered from *Helicops* species (Chapter 3) and colubrid species (Chapter 4). However, the evolution of other highly expressed toxin families should be of interest. For example, snake venom metalloproteinase (SVMPs) is highly expressed in *Leptodeira* species (Chapter 2 [40]). The evolutionary histories of SVMPs in dipsadine snakes is largely unknown [119], and it is unclear if *Leptodeira* SVMPs possess novel toxin structures, as shown in *Helicops* CTLs (Chapter 3) or *Chironius multiventris* putative heterodimeric 3FTxs (Chapter 4). Further, lowly expressed toxins recovered here were not investigated, and without understanding how these toxins are evolving, we are missing the full picture of snake venom evolution.

A core theme of my dissertation is the evolution of novelty. In Chapter 2, I described the venom profiles of several genera of snake, one of which showed novel expression patterns. I further investigated this toxin in chapter 3 and showed how a novel insertion within *Helicops* CTL sequences was correlated with high selection and high expression. Finally, in Chapter 4, I recovered putative heterodimeric 3FTxs and showed evidence of parallel evolution of this interaction across colubrid snakes. However, the work here focused on a total of ten rear-fanged snakes, or approximately 1% of all rear-fanged snakes. The ability to recover several novel toxin

structures and expression patterns from a very few snakes suggest that rear-fanged snakes are an exciting avenue for those interested in the evolution of novelty.

Finally, the scientific community can leverage snake venoms as this trait is a potential system to determine the molecular and ecological factors involved in the evolution of complex traits. For example, comparisons with other rapidly evolving gene families can be made to gain insights into how and why some gene families are diversifying while others are relatively conserved. An ideal study would be to determine how rates of evolution among MHC genes and venom genes differ, as both are vital to species interactions.

Bibliography

1. Nam, J.; Nei, M. Evolutionary Change of the Numbers of Homeobox Genes in Bilateral Animals. *Mol Biol Evol* **2005**, *22*, 2386–2394, doi:10.1093/MOLBEV/MSI229.
2. Nei, M. The New Mutation Theory of Phenotypic Evolution. *Proc Natl Acad Sci U S A* **2007**, *104*, 12235–12242, doi:10.1073/PNAS.0703349104/SUPPL_FILE/INDEX.HTML.
3. Nei, M. Molecular Evolutionary Genetics. *Molecular Evolutionary Genetics* **1987**, doi:10.7312/NEI-92038/HTML.
4. Margres, M.J.; McGivern, J.J.; Seavy, M.; Wray, K.P.; Facente, J.; Rokyta, D.R. Contrasting Modes and Tempos of Venom Expression Evolution in Two Snake Species. *Genetics* **2015**, *199*, 165–176, doi:10.1534/genetics.114.172437.
5. Lamichhaney, S.; Han, F.; Berglund, J.; Wang, C.; Almén, M.S.; Webster, M.T.; Grant, B.R.; Grant, P.R.; Andersson, L. A Beak Size Locus in Darwin's Finches Facilitated Character Displacement during a Drought. *Science (1979)* **2016**, *352*, 470–474, doi:10.1126/science.aad8786.
6. Lamichhaney, S.; Berglund, J.; Almén, M.S.; Maqbool, K.; Grabherr, M.; Martinez-Barrio, A.; Promerová, M.; Rubin, C.J.; Wang, C.; Zamani, N.; et al. Evolution of Darwin's Finches and Their Beaks Revealed by Genome Sequencing. *Nature* **2015**, *518*, 371–375, doi:10.1038/nature14181.

7. Mullen, L.M.; Hoekstra, H.E. Natural Selection along an Environmental Gradient: A Classic Cline in Mouse Pigmentation. *Evolution (N Y)* **2008**, *62*, 1555–1570, doi:10.1111/j.1558-5646.2008.00425.x.

8. Wong, E.S.W.; Belov, K. Venom Evolution through Gene Duplications. *Gene* **2012**, *496*, 1–7, doi:10.1016/j.gene.2012.01.009.

9. Duda, T.F.; Palumbi, S.R. Molecular Genetics of Ecological Diversification: Duplication and Rapid Evolution of Toxin Genes of the Venomous Gastropod Conus. *Proc Natl Acad Sci U S A* **1999**, *96*, 6820–6823, doi:10.1073/pnas.96.12.6820.

10. Zhao, W.; Shi, M.; Ye, X.Q.; Li, F.; Wang, X.W.; Chen, X.X. Comparative Transcriptome Analysis of Venom Glands from Cotesia Vestalis and Diadromus Collaris, Two Endoparasitoids of the Host Plutella Xylostella. *Sci Rep* **2017**, *7*, 1–12, doi:10.1038/s41598-017-01383-2.

11. Haney, R.A.; Clarke, T.H.; Gadgil, R.; Fitzpatrick, R.; Hayashi, C.Y.; Ayoub, N.A.; Garb, J.E. Effects of Gene Duplication, Positive Selection, and Shifts in Gene Expression on the Evolution of the Venom Gland Transcriptome in Widow Spiders. *Genome Biol Evol* **2016**, *8*, 228–242, doi:10.1093/gbe/evv253.

12. Zhu, S.; Bosmans, F.; Tytgat, J. Adaptive Evolution of Scorpion Sodium Channel Toxins. *J Mol Evol* **2004**, *58*, 145–153, doi:10.1007/s00239-003-2534-2.

13. Mackessy, S.P. Venom Ontogeny in the Pacific Rattlesnakes Crotalus Viridis Helleri and C. v. Oreganus. *Copeia* **1988**, *1988*, 92, doi:10.2307/1445927.

14. Duda, T.F.; Palumbi, S.R. Gene Expression and Feeding Ecology: Evolution of Piscivory in the Venomous Gastropod Genus Conus. *Proceedings of the Royal Society B: Biological Sciences* **2004**, *271*, 1165–1174, doi:10.1098/rspb.2004.2708.

15. Daltry, J.C.; Wüster, W.; Thorpe, R.S. Diet and Snake Venom Evolution. *Nature* **1996**, *379*, 537–542, doi:10.1038/379537a0.

16. Anne-Isola Nekaris, K.; Moore, R.S.; Johanna Rode, E.; Fry, B.G. Mad, Bad and Dangerous to Know: The Biochemistry, Ecology and Evolution of Slow Loris Venom. *Journal of Venomous Animals and Toxins Including Tropical Diseases* **2013**, *19*, 1–10, doi:10.1186/1678-9199-19-21.

17. Eble, J.A. Structurally Robust and Functionally Highly Versatile—C-Type Lectin (-Related) Proteins in Snake Venoms. *Toxins (Basel)* 2019, *11*, 136.

18. Oyama, E.; Takahashi, H. Structures and Functions of Snake Venom Metalloproteinases (SVMP) from Protobothrops Venom Collected in Japan. *Molecules* **2017**, *22*, doi:10.3390/molecules22081305.

19. Kini, R.M.; Doley, R. Structure, Function and Evolution of Three-Finger Toxins: Mini Proteins with Multiple Targets. *Toxicon* **2010**, *56*, 855–867, doi:10.1016/J.TOXICON.2010.07.010.

20. Santos, A.D.; McIntosh, J.M.; Hillyard, D.R.; Cruz, L.J.; Olivera, B.M. The A-Superfamily of Conotoxins: Structural and Functional Divergence. *Journal of Biological Chemistry* **2004**, *279*, 17596–17606, doi:10.1074/jbc.M309654200.

21. Fry, B.G.; Scheib, H.; van der Weerd, L.; Young, B.; McNaughtan, J.; Ryan Ramjan, S.F.; Vidal, N.; Poelmann, R.E.; Norman, J.A. Evolution of an Arsenal: Structural and Functional Diversification of the Venom System in the Advanced Snakes (Caenophidia). *Molecular and Cellular Proteomics* **2008**, *7*, 215–246, doi:10.1074/mcp.M700094-MCP200.

22. Mackessy, S.P.; Saviola, A.J. Understanding Biological Roles of Venoms among the Caenophidia: The Importance of Rear-Fanged Snakes. *Integr Comp Biol* **2016**, *56*, 1004–1021, doi:10.1093/icb/icw110.

23. Casewell, N.R.; Jackson, T.N.W.; Laustsen, A.H.; Sunagar, K. Causes and Consequences of Snake Venom Variation. *Trends Pharmacol Sci* 2020, *41*, 570–581.

24. Gibbs, H.L.; Sanz, L.; Pérez, A.; Ochoa, A.; Hassinger, A.T.B.; Holding, M.L.; Calvete, J.J. The Molecular Basis of Venom Resistance in a Rattlesnake-Squirrel Predator-Prey System. *Mol Ecol* **2020**, *29*, 2871–2888, doi:10.1111/mec.15529.

25. Robinson, K.E.; Holding, M.L.; Whitford, M.D.; Saviola, A.J.; Yates, J.R.; Clark, R.W. Phenotypic and Functional Variation in Venom and Venom Resistance of Two Sympatric Rattlesnakes and Their Prey. *J Evol Biol* **2021**, *34*, 1447–1465, doi:10.1111/jeb.13907.

26. Mackessy, S.P.; Sixberry, N.M.; Heyborne, W.H.; Fritts, T. Venom of the Brown Treesnake, Boiga Irregularis: Ontogenetic Shifts and Taxa-Specific Toxicity. *Toxicon* **2006**, *47*, 537–548, doi:10.1016/j.toxicon.2006.01.007.

27. Freitas-de-Sousa, L.A.; Nachtigall, P.G.; Portes-Junior, J.A.; Holding, M.L.; Nystrom, G.S.; Ellsworth, S.A.; Guimarães, N.C.; Tioyama, E.; Ortiz, F.; Silva, B.R.; et al. Size Matters: An Evaluation of the Molecular Basis of Ontogenetic Modifications in the Composition of Bothrops Jararacussu Snake Venom. *Toxins (Basel)* **2020**, *12*, 791, doi:10.3390/toxins12120791.

28. Rotenberg, D.; Bamberger, E.S.; Kochva, E. Studies on Ribonucleic Acid Synthesis in the Venom Glands of Vipera Palaestinae (Ophidia, Reptilia). *Biochem J* **1971**, *121*, 609–612, doi:10.1042/bj1210609.

29. Von Reumont, B.M. Studying Smaller and Neglected Organisms in Modern Evolutionary Venomics Implementing RNASeq (Transcriptomics)—A Critical Guide. *Toxins (Basel)* **2018**, *10*, 1–23, doi:10.3390/toxins10070292.

30. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex Cocktails: The Evolutionary Novelty of Venoms. *Trends Ecol Evol* **2013**, *28*, 219–229, doi:10.1016/j.tree.2012.10.020.

31. Junqueira-de-Azevedo, I.L.M.; Campos, P.F.; Ching, A.T.C.; Mackessy, S.P. Colubrid Venom Composition: An -Omics Perspective. *Toxins (Basel)* **2016**, *8*, 1–24, doi:10.3390/toxins8080230.

32. de Oliveira, L.; Scartozzoni, R.R.; de Almeida-Santos, S.M.; Jared, C.; Antoniazzi, M.M.; da Graça Salomão, M. Morphology of Duvernoy's Glands and Maxillary Teeth and a Possible Function of the Duvernoy's Gland Secretion in Helicops Modestus Günther, 1861 (Serpentes: Xenodontinae). *South Am J Herpetol* **2016**, *11*, 54–65, doi:10.2994/SAJH-D-16-00011.1.

33. Westeen, E.P.; Durso, A.M.; Grundler, M.C.; Rabosky, D.L.; Davis Rabosky, A.R. What Makes a Fang? Phylogenetic and Ecological Controls on Tooth Evolution in Rear-Fanged Snakes. *BMC Evol Biol* **2020**, *20*, 1–15, doi:10.1186/s12862-020-01645-0.

34. Sunagar, K.; Undheim, E.A.B.; Scheib, H.; Gren, E.C.K.; Cochran, C.; Person, C.E.; Koludarov, I.; Kelln, W.; Hayes, W.K.; King, G.F.; et al. Intraspecific Venom Variation in the Medically Significant Southern Pacific Rattlesnake (Crotalus Oreganus Helleri): Biodiscovery, Clinical and Evolutionary Implications. *J Proteomics* **2014**, *99*, 68–83, doi:10.1016/j.jprot.2014.01.013.

35. Tan, K.Y.; Tan, C.H.; Chanhome, L.; Tan, N.H. Comparative Venom Gland Transcriptomics of Naja Kaouthia (Monocled Cobra) from Malaysia and Thailand: Elucidating Geographical Venom Variation and Insights into Sequence Novelty. *PeerJ* **2017**, *2017*, doi:10.7717/peerj.3142.

36. Panagides, N.; Jackson, T.N.W.; Ikonomopoulou, M.P.; Arbuckle, K.; Pretzler, R.; Yang, D.C.; Ali, S.A.; Koludarov, I.; Dobson, J.; Sanker, B.; et al. How the Cobra Got Its Flesh-Eating Venom: Cytotoxicity as a Defensive Innovation and Its Co-Evolution with Hooding, Aposematic Marking, and Spitting. *Toxins (Basel)* **2017**, *9*, doi:10.3390/toxins9030103.

37. Holding, M.L.; Strickland, J.L.; Rautsaw, R.M.; Hofmann, E.P.; Mason, A.J.; Hogan, M.P.; Nystrom, G.S.; Ellsworth, S.A.; Colston, T.J.; Borja, M.; et al. Phylogenetically Diverse Diets Favor More Complex Venoms in North American Pitvipers. *Proc Natl Acad Sci U S A* **2021**, *118*, 2015579118, doi:10.1073/pnas.2015579118.

38. Fry, B.G.; Wüster, W.; Kini, R.M.; Brusic, V.; Khan, A.; Venkataraman, D.; Rooney, A.P. Molecular Evolution and Phylogeny of Elapid Snake Venom Three-Finger Toxins. *J Mol Evol* **2003**, *57*, 110–129, doi:10.1007/s00239-003-2461-2.

39. Strickland, J.L.; Smith, C.F.; Mason, A.J.; Schield, D.R.; Borja, M.; Castañeda-Gaytán, G.; Spencer, C.L.; Smith, L.L.; Trápaga, A.; Bouzid, N.M.; et al. Evidence for Divergent Patterns of Local Selection Driving Venom Variation in Mojave Rattlesnakes (Crotalus Scutulatus). *Sci Rep* **2018**, *8*, 1–15, doi:10.1038/s41598-018-35810-9.

40. Cerda, P.A.; Crowe-Riddell, J.M.; Gonçalves, D.J.P.; Larson, D.A.; Duda, T.F.; Davis Rabosky, A.R. Divergent Specialization of Simple Venom Gene Profiles among Rear-

Fanged Snake Genera (Helicops and Leptodeira, Dipsadinae, Colubridae). *Toxins (Basel)* **2022**, *14*, 489, doi:10.3390/toxins14070489.

41. Kordiš, D.; Bdolah, A.; Gubenšek, F. Positive Darwinian Selection in Vipera Palaestinae Phospholipase A2 Genes Is Unexpectedly Limited to the Third Exon. *Biochem Biophys Res Commun* **1998**, *251*, 613–619, doi:10.1006/bbrc.1998.9528.
42. Lynch, V.J. Inventing an Arsenal: Adaptive Evolution and Neofunctionalization of Snake Venom Phospholipase A2 Genes. *BMC Evol Biol* **2007**, *7*, 1–14, doi:10.1186/1471-2148-7-2.
43. Giorgianni, M.W.; Dowell, N.L.; Griffin, S.; Kassner, V.A.; Selegue, J.E.; Carroll, S.B. The Origin and Diversification of a Novel Protein Family in Venomous Snakes. *Proc Natl Acad Sci U S A* **2020**, *117*, 10911–10920, doi:10.1073/pnas.1920011117.
44. Dowell, N.L.; Giorgianni, M.W.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. The Deep Origin and Recent Loss of Venom Toxin Genes in Rattlesnakes. *Current Biology* **2016**, *26*, 2434–2445, doi:10.1016/j.cub.2016.07.038.
45. Margres, M.J.; Bigelow, A.T.; Lemmon, E.M.; Lemmon, A.R.; Rokyta, D.R. Selection To Increase Expression, Not Sequence Diversity, Precedes Gene Family Origin and Expansion in Rattlesnake Venom. *Genetics* **2017**, *206*, 1569–1580, doi:10.1534/genetics.117.202655.
46. Sanz, L.; Lisle Gibbs, H.; Mackessy, S.P.; Calvete, J.J. Venom Proteomes of Closely Related Sistrurus Rattlesnakes with Divergent Diets. *J Proteome Res* **2006**, *5*, 2098–2112, doi:10.1021/pr0602500.

47. Phuong, M.A.; Mahardika, G.N.; Alfaro, M.E. Dietary Breadth Is Positively Correlated with Venom Complexity in Cone Snails. *BMC Genomics* **2016**, *17*, 1–15, doi:10.1186/s12864-016-2755-6.

48. Lyons, K.; Dugon, M.M.; Healy, K. Diet Breadth Mediates the Prey Specificity of Venom Potency in Snakes. *Toxins (Basel)* **2020**, *12*, 74, doi:10.3390/toxins12020074.

49. Barlow, A.; Pook, C.E.; Harrison, R.A.; Wüster, W. Coevolution of Diet and Prey-Specific Venom Activity Supports the Role of Selection in Snake Venom Evolution. *Proceedings of the Royal Society B: Biological Sciences* **2009**, *276*, 2443–2449, doi:10.1098/rspb.2009.0048.

50. Remigio, E.A.; Duda, T.F. Evolution of Ecological Specialization and Venom of a Predatory Marine Gastropod. *Mol Ecol* **2008**, *17*, 1156–1162, doi:10.1111/j.1365-294X.2007.03627.x.

51. Ligabue-Braun, R.; Verli, H.; Carlini, C.R. Venomous Mammals: A Review. *Toxicon* **2012**, *59*, 680–695, doi:10.1016/j.toxicon.2012.02.012.

52. Pekár, S.; Bočánek, O.; Michálek, O.; Petráková, L.; Haddad, C.R.; Šedo, O.; Zdráhal, Z. Venom Gland Size and Venom Complexity—Essential Trophic Adaptations of Venomous Predators: A Case Study Using Spiders. *Mol Ecol* **2018**, *27*, 4257–4269, doi:10.1111/mec.14859.

53. Sanz, L.; Quesada-Bernat, S.; Ramos, T.; Casais-e-Silva, L.L.; Corrêa-Netto, C.; Silva-Haad, J.J.; Sasa, M.; Lomonte, B.; Calvete, J.J. New Insights into the Phylogeographic Distribution of the 3FTx/PLA 2 Venom Dichotomy across Genus Micrurus in South America. *J Proteomics* **2019**, *200*, 90–101, doi:10.1016/j.jprot.2019.03.014.

54. Mackessy, S.P. *Handbook of Venoms and Toxins of Reptiles*; Mackessy, S.P., Ed.; 1st ed.; CRC Press: Boca Raton, 2016; ISBN 9780429186394.

55. Jackson, T.N.W.; Koludarov, I.; Ali, S.A.; Dobson, J.; Zdenek, C.N.; Dashevsky, D.; Op Den Brouw, B.; Masci, P.P.; Nouwens, A.; Josh, P.; et al. Rapid Radiations and the Race to Redundancy: An Investigation of the Evolution of Australian Elapid Snake Venoms. *Toxins (Basel)* **2016**, *8*, 309, doi:10.3390/toxins8110309.

56. Pyron, R.A.; Burbrink, F.T.; Wiens, J.J. A Phylogeny and Revised Classification of Squamata, Including 4161 Species of Lizards and Snakes. *BMC Evol Biol* **2013**, *13*, 1, doi:10.1186/1471-2148-13-93.

57. Jackson, T.N.W.; Young, B.; Underwood, G.; McCarthy, C.J.; Kochva, E.; Vidal, N.; van der Weerd, L.; Nabuurs, R.; Dobson, J.; Whitehead, D.; et al. Endless Forms Most Beautiful: The Evolution of Ophidian Oral Glands, Including the Venom System, and the Use of Appropriate Terminology for Homologous Structures. *Zoomorphology* **2017**, *136*, 107–130, doi:10.1007/s00435-016-0332-9.

58. McGivern, J.J.; Wray, K.P.; Margres, M.J.; Couch, M.E.; Mackessy, S.P.; Rokyta, D.R. RNA-Seq and High-Definition Mass Spectrometry Reveal the Complex and Divergent Venoms of Two Rear-Fanged Colubrid Snakes. *BMC Genomics* **2014**, *15*, 1–18, doi:10.1186/1471-2164-15-1061.

59. Hofmann, E.P.; Rautsaw, R.M.; Mason, A.J.; Strickland, J.L.; Parkinson, C.L. Duvernoy's Gland Transcriptomics of the Plains Black-Headed Snake, Tantilla Nigriceps (Squamata, Colubridae): Unearthing the Venom of Small Rear-Fanged Snakes. *Toxins (Basel)* **2021**, *13*, 336, doi:10.3390/toxins13050336.

60. Lemoine, K.; Girón, M.E.; Aguilar, I.; Navarrete, L.F.; Rodríguez-Acosta, A. Proteolytic, Hemorrhagic, and Neurotoxic Activities Caused by Leptodeira Annulata Ashmeadii (Serpentes: Colubridae) Duvernoy's Gland Secretion. *Wilderness Environ Med* **2004**, *15*, 82–89, doi:10.1580/1080-6032(2004)015[0082:PHANAC]2.0.CO;2.

61. Sánchez, M.N.; Gonzalez, K.Y.; Sciani, J.M.; Gritti, M.A.; Maruñak, S.L.; Tavares, F.L.; Teibler, G.P.; Peichoto, M.E. First Insights into the Biochemical and Toxicological Characterization of Venom from the Banded Cat-Eyed Snake Leptodeira Annulata Pulchriceps. *Comparative Biochemistry and Physiology Part - C: Toxicology and Pharmacology* **2021**, *239*, 108897, doi:10.1016/j.cbpc.2020.108897.

62. Torres-Bonilla, K.A.; Schezaro-Ramos, R.; Floriano, R.S.; Rodrigues-Simioni, L.; Bernal-Bautista, M.H.; Alice da Cruz-Höfling, M. Biological Activities of Leptodeira Annulata (Banded Cat-Eyed Snake) Venom on Vertebrate Neuromuscular Preparations. *Toxicon* **2016**, *119*, 345–351, doi:10.1016/j.toxicon.2016.07.004.

63. Torres-Bonilla, K.A.; Panunto, P.C.; Pereira, B.B.; Zambrano, D.F.; Herrán-Medina, J.; Bernal, M.H.; Hyslop, S. Toxinological Characterization of Venom from Leptodeira Annulata (Banded Cat-Eyed Snake; Dipsadidae, Imantodini): Toxinological Characterization of L. Annulata Venom. *Biochimie* **2020**, *174*, 171–188, doi:10.1016/j.biochi.2020.04.006.

64. Estrella, A.; Sánchez, E.E.; Galán, J.A.; Tao, W.A.; Guerrero, B.; Navarrete, L.F.; Rodríguez-Acosta, A. Characterization of Toxins from the Broad-Banded Water Snake Helicops Angulatus (Linnaeus, 1758): Isolation of a Cysteine-Rich Secretory Protein, Helicopsin. *Arch Toxicol* **2011**, *85*, 305–313, doi:10.1007/s00204-010-0597-6.

65. Estrella, A.; Rodríguez-torres, A.; Serna, L.; Navarrete, L.F. Is the South American Water Snake Helicops Angulatus (Linnaeus, 1758) (Ddipsadidae:Xenodontinae) Venomous? *Herpetotropicos* **2012**, *5*, 79–84.

66. Gutiérrez, J.M.; Calvete, J.J.; Habib, A.G.; Harrison, R.A.; Williams, D.J.; Warrell, D.A. Snakebite Envenoming. *Nat Rev Dis Primers* 2017, *3*, 17063.

67. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available Online at: Http://Www.Bioinformatics.Babraham.Ac.Uk/Projects/Fastqc/ 2010.

68. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120, doi:10.1093/bioinformatics/btu170.

69. Haas, B.J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.D.; Bowden, J.; Couger, M.B.; Eccles, D.; Li, B.; Lieber, M.; et al. De Novo Transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis. *Nat Protoc* **2013**, *8*, 1494–1512, doi:10.1038/nprot.2013.084.

70. Rokyta, D.R.; Lemmon, A.R.; Margres, M.J.; Aronow, K. The Venom-Gland Transcriptome of the Eastern Diamondback Rattlesnake (Crotalus Adamanteus). *BMC Genomics* **2012**, *13*, doi:10.1186/1471-2164-13-312.

71. Holding, M.L.; Margres, M.J.; Mason, A.J.; Parkinson, C.L.; Rokyta, D.R. Evaluating the Performance of de Novo Assembly Methods for Venom-Gland Transcriptomics. *Toxins (Basel)* **2018**, *10*, 1–21, doi:10.3390/toxins10060249.

72. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A. PEAR: A Fast and Accurate Illumina Paired-End ReAd MergeR. *Bioinformatics* **2014**, *30*, 614–620, doi:10.1093/bioinformatics/btt593.

73. Yang, Y.; Smith, S.A. Orthology Inference in Nonmodel Organisms Using Transcriptomes and Low-Coverage Genomes: Improving Accuracy and Matrix Occupancy for Phylogenomics. *Mol Biol Evol* **2014**, *31*, 3081–3092, doi:10.1093/molbev/msu245.

74. Morales-Briones, D.F.; Kadereit, G.; Tefarikis, D.T.; Moore, M.J.; Smith, S.A.; Brockington, S.F.; Timoneda, A.; Yim, W.C.; Cushman, J.C.; Yang, Y. Disentangling Sources of Gene Tree Discordance in Phylogenomic Data Sets: Testing Ancient Hybridizations in Amaranthaceae s.l. *Syst Biol* **2021**, *70*, 219–235, doi:10.1093/sysbio/syaa066.

75. Smith-Unna, R.; Boursnell, C.; Patro, R.; Hibberd, J.M.; Kelly, S. TransRate: Reference-Free Quality Assessment of de Novo Transcriptome Assemblies. *Genome Res* **2016**, *26*, 1134–1144, doi:10.1101/gr.196469.115.

76. Yang, Y.; Smith, S.A. Optimizing de Novo Assembly of Short-Read RNA-Seq Data for Phylogenomics. *BMC Genomics* **2013**, *14*, 328, doi:10.1186/1471-2164-14-328.

77. Davidson, N.M.; Oshlack, A. Corset: Enabling Differential Gene Expression Analysis for de Novo Assembled Transcriptomes. *Genome Biol* **2014**, *15*, 1–14, doi:10.1186/s13059-014-0410-6.

78. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression. *Nat Methods* **2017**, *14*, 417–419, doi:10.1038/nmeth.4197.

79. Http://Transdecoder.Github.Io/.

80. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and Applications. *BMC Bioinformatics* **2009**, *10*, 1–9, doi:10.1186/1471-2105-10-421.

81. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for Clustering the next-Generation Sequencing Data. *Bioinformatics* **2012**, *28*, 3150–3152, doi:10.1093/bioinformatics/bts565.

82. Li, B.; Ruotti, V.; Stewart, R.M.; Thomson, J.A.; Dewey, C.N. RNA-Seq Gene Expression Estimation with Read Mapping Uncertainty. *Bioinformatics* **2009**, *26*, 493–500, doi:10.1093/bioinformatics/btp692.

83. Langmead, B.; Salzberg, S.L. Fast Gapped-Read Alignment with Bowtie 2. *Nat Methods* **2012**, *9*, 357–359, doi:10.1038/nmeth.1923.

84. Rokyta, D.R.; Wray, K.P.; Margres, M.J. The Genesis of an Exceptionally Lethal Venom in the Timber Rattlesnake (Crotalus Horridus) Revealed through Comparative Venom-Gland Transcriptomics. *BMC Genomics* **2013**, *14*, 394, doi:10.1186/1471-2164-14-394.

85. Margres, M.J.; Aronow, K.; Loyacano, J.; Rokyta, D.R. The Venom-Gland Transcriptome of the Eastern Coral Snake (Micrurus Fulvius) Reveals High Venom Complexity in the Intragenomic Evolution of Venoms. *BMC Genomics* **2013**, *14*, doi:10.1186/1471-2164-14-531.

86. Modahl, C.M.; Mrinalini; Frietze, S.; Mackessy, S.P. Adaptive Evolution of Distinct Prey-Specific Toxin Genes in Rear-Fanged Snake Venom. *Proceedings of the Royal Society B: Biological Sciences* **2018**, *285*, 20181003, doi:10.1098/rspb.2018.1003.

87. Vonk, F.J.; Casewell, N.R.; Henkel, C. v.; Heimberg, A.M.; Jansen, H.J.; McCleary, R.J.R.; Kerkkamp, H.M.E.; Vos, R.A.; Guerreiro, I.; Calvete, J.J.; et al. The King Cobra

Genome Reveals Dynamic Gene Evolution and Adaptation in the Snake Venom System. *Proc Natl Acad Sci U S A* **2013**, *110*, 20651–20656, doi:10.1073/pnas.1314702110.

88. Bayona-Serrano, J.D.; Viala, V.L.; Rautsaw, R.M.; Schramer, T.D.; Barros-Carvalho, G.A.; Nishiyama, M.Y.; Freitas-De-Sousa, L.A.; Moura-Da-Silva, A.M.; Parkinson, C.L.; Grazziotin, F.G.; et al. Replacement and Parallel Simplification of Nonhomologous Proteinases Maintain Venom Phenotypes in Rear-Fanged Snakes. *Mol Biol Evol* **2020**, *37*, 3563–3575, doi:10.1093/molbev/msaa192.

89. R Core Team R: A Language and Environment for Statistical Computing. Available online: https://www.r-project.org/.

90. Johnson, M.G.; Gardner, E.M.; Liu, Y.; Medina, R.; Goffinet, B.; Shaw, A.J.; Zerega, N.J.C.; Wickett, N.J. HybPiper: Extracting Coding Sequence and Introns for Phylogenetics from High-throughput Sequencing Reads Using Target Enrichment. *Appl Plant Sci* **2016**, *4*, 1600016, doi:10.3732/apps.1600016.

91. Dong, S.; Kumazawa, Y. Complete Mitochondrial DNA Sequences of Six Snakes: Phylogenetic Relationships and Molecular Evolution of Genomic Features. *J Mol Evol* **2005**, *61*, 12–22, doi:10.1007/s00239-004-0190-9.

92. Kumazawa, Y.; Ota, H.; Nishida, M.; Ozawa, T. The Complete Nucleotide Sequence of a Snake (Dinodon Semicarinatus) Mitochondrial Genome with Two Identical Control Regions. *Genetics* **1998**, *150*, 313–329, doi:10.1093/genetics/150.1.313.

93. Jiang, Z.J.; Castoe, T.A.; Austin, C.C.; Burbrink, F.T.; Herron, M.D.; McGuire, J.A.; Parkinson, C.L.; Pollock, D.D. Comparative Mitochondrial Genomics of Snakes: Extraordinary Substitution Rate Dynamics and Functionality of the Duplicate Control Region. *BMC Evol Biol* **2007**, *7*, doi:10.1186/1471-2148-7-123.

94. Yan, J.; Li, H.; Zhou, K. Evolution of the Mitochondrial Genome in Snakes: Gene Rearrangements and Phylogenetic Relationships. *BMC Genomics* **2008**, *9*, 1–7, doi:10.1186/1471-2164-9-569.

95. Kumazawa, Y.; Nishida, M. Complete Mitochondrial DNA Sequences of the Green Turtle and Blue- Tailed Mole Skink: Statistical Evidence for Archosaurian Affinity of Turtles. *Mol Biol Evol* **1999**, *16*, 784–792, doi:10.1093/oxfordjournals.molbev.a026163.

96. Macey, J.R.; Papenfuss, T.J.; Kuehl, J. V.; Fourcade, H.M.; Boore, J.L. Phylogenetic Relationships among Amphisbaenian Reptiles Based on Complete Mitochondrial Genomic Sequences. *Mol Phylogenet Evol* **2004**, *33*, 22–31, doi:10.1016/j.ympev.2004.05.003.

97. Slater, G.S.C.; Birney, E. Automated Generation of Heuristics for Biological Sequence Comparison. *BMC Bioinformatics* **2005**, *6*, 1–11, doi:10.1186/1471-2105-6-31.

98. Cock, P.J.A.; Antao, T.; Chang, J.T.; Chapman, B.A.; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; et al. Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423, doi:10.1093/bioinformatics/btp163.

99. Li, H.; Durbin, R. Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2009**, *25*, 1754–1760, doi:10.1093/bioinformatics/btp324.

100. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079, doi:10.1093/bioinformatics/btp352.

101. Tange, O. GNU Parallel 20150322 ('Hellwig'). *login: The USENIX Magazine* **2015**, doi:10.5281/ZENODO.16303.

102. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology* **2012**, *19*, 455–477, doi:10.1089/cmb.2012.0021.

103. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol* **2013**, *30*, 772–780, doi:10.1093/molbev/mst010.

104. Brown, J.W.; Walker, J.F.; Smith, S.A. Phyx: Phylogenetic Tools for Unix. *Bioinformatics* **2017**, *33*, 1886–1888, doi:10.1093/bioinformatics/btx063.

105. Nguyen, L.T.; Schmidt, H.A.; Von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol* **2015**, *32*, 268–274, doi:10.1093/molbev/msu300.

106. Shannon, C.E. A Mathematical Theory of Communication. *Bell System Technical Journal* **1948**, *27*, 379–423, doi:10.1002/j.1538-7305.1948.tb01338.x.

107. Bray, J.R.; Curtis, J.T. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecol Monogr* **1957**, *27*, 325–349, doi:10.2307/1942268.

108. Callahan, S.; Crowe-Riddell, J.M.; Nagesan, R.S.; Gray, J.A.; Davis Rabosky, A.R. A Guide for Optimal Iodine Staining and High-Throughput DiceCT Scanning in Snakes. *Ecol Evol* **2021**, *11*, 11587–11603, doi:10.1002/ece3.7467.

109. Lomonte, B.; Rey-Suárez, P.; Fernández, J.; Sasa, M.; Pla, D.; Vargas, N.; Bénard-Valle, M.; Sanz, L.; Corrêa-Netto, C.; Núñez, V.; et al. Venoms of Micrurus Coral Snakes: Evolutionary Trends in Compositional Patterns Emerging from Proteomic Analyses. *Toxicon* **2016**, *122*, 7–25, doi:10.1016/j.toxicon.2016.09.008.

110. Amazonas, D.R.; Portes-Junior, J.A.; Nishiyama-Jr, M.Y.; Nicolau, C.A.; Chalkidis, H.M.; Mourão, R.H.V.; Grazziotin, F.G.; Rokyta, D.R.; Gibbs, H.L.; Valente, R.H.; et al. Molecular Mechanisms Underlying Intraspecific Variation in Snake Venom. *J Proteomics* **2018**, *181*, 60–72, doi:10.1016/j.jprot.2018.03.032.

111. Sanz, L.; Quesada-Bernat, S.; Pérez, A.; de Morais-Zani, K.; SantEänna, S.S.; Hatakeyama, D.M.; Tasima, L.J.; de Souza, M.B.; Kayano, A.M.; Zavaleta, A.; et al. Danger in the Canopy. Comparative Proteomics and Bioactivities of the Venoms of the South American Palm Pit Viper Bothrops Bilineatus Subspecies Bilineatus and Smaragdinus and Antivenomics of B. b. Bilineatus (Rondônia) Venom against the Brazilian Pentabo. *J Proteome Res* **2020**, *19*, 3518–3532, doi:10.1021/acs.jproteome.0c00337.

112. Mackessy, S.P.; Bryan, W.; Smith, C.F.; Lopez, K.; Fernández, J.; Bonilla, F.; Camacho, E.; Sasa, M.; Lomonte, B. Venomics of the Central American Lyre Snake Trimorphodon Quadruplex (Colubridae: Smith, 1941) from Costa Rica. *J Proteomics* **2020**, *220*, 103778, doi:10.1016/j.jprot.2020.103778.

113. Peichoto, M.E.; Tavares, F.L.; Santoro, M.L.; MacKessy, S.P. Venom Proteomes of South and North American Opisthoglyphous (Colubridae and Dipsadidae) Snake Species: A Preliminary Approach to Understanding Their Biological Roles. *Comp Biochem Physiol Part D Genomics Proteomics* **2012**, *7*, 361–369, doi:10.1016/j.cbd.2012.08.001.

114. Modahl, C.M.; Mackessy, S.P. Venoms of Rear-Fanged Snakes: New Proteins and Novel Activities. *Front Ecol Evol* **2019**, *7*, 1–18, doi:10.3389/fevo.2019.00279.

115. Grundler, M.C. SquamataBase: A Natural History Database and R Package for Comparative Biology of Snake Feeding Habits. *Biodivers Data J* **2020**, *8*, e49943-, doi:10.3897/BDJ.8.E49943.

116. Duellman, W.E. *Cusco Amazonico: The Lives of Amphibians and Reptiles in an Amazonian Rainforest*; Comstock Pub. Associates, 2005; Vol. 43; ISBN 9780801439971.

117. Arlinghaus, F.T.; Eble, J.A. C-Type Lectin-like Proteins from Snake Venoms. *Toxicon* **2012**, *60*, 512–519, doi:10.1016/j.toxicon.2012.03.001.

118. Ogawa, T.; Chijiwa, T.; Oda-Ueda, N.; Ohno, M. Molecular Diversity and Accelerated Evolution of C-Type Lectin-like Proteins from Snake Venom. *Toxicon* 2005, *45*, 1–14.

119. Xie, B.; Dashevsky, D.; Rokyta, D.; Ghezellou, P.; Fathinia, B.; Shi, Q.; Richardson, M.K.; Fry, B.G. Dynamic Genetic Differentiation Drives the Widespread Structural and Functional Convergent Evolution of Snake Venom Proteinaceous Toxins. *BMC Biol* **2022**, *20*, 1–31, doi:10.1186/s12915-021-01208-9.

120. Sanz, L.; Pérez, A.; Quesada-Bernat, S.; Diniz-Sousa, R.; Calderón, L.A.; Soares, A.M.; Calvete, J.J.; Caldeira, C.A.S. Venomics and Antivenomics of the Poorly Studied Brazil's Lancehead, Bothrops Brazili (Hoge, 1954), from the Brazilian State of Pará. *Journal of Venomous Animals and Toxins Including Tropical Diseases* 2020, *26*, 20190103.

121. Weese, D.A.; Duda, T.F. Effects of Predator-Prey Interactions on Predator Traits: Differentiation of Diets and Venoms of a Marine Snail. *Toxins (Basel)* **2019**, *11*, doi:10.3390/toxins11050299.

122. Mora-Obando, D.; Salazar-Valenzuela, D.; Pla, D.; Lomonte, B.; Guerrero-Vargas, J.A.; Ayerbe, S.; Gibbs, H.L.; Calvete, J.J. Venom Variation in Bothrops Asper Lineages from

North-Western South America. *J Proteomics* **2020**, *229*, 103945, doi:10.1016/j.jprot.2020.103945.

123. Barua, A.; Mikheyev, A.S. Many Options, Few Solutions: Over 60 My Snakes Converged on a Few Optimal Venom Formulations. *Mol Biol Evol* **2019**, *36*, 1964–1974, doi:10.1093/molbev/msz125.

124. Modahl, C.M.; Frietze, S.; Mackessy, S.P. Transcriptome-Facilitated Proteomic Characterization of Rear-Fanged Snake Venoms Reveal Abundant Metalloproteinases with Enhanced Activity. *J Proteomics* **2018**, *187*, 223–234, doi:10.1016/j.jprot.2018.08.004.

125. Pla, D.; Sanz, L.; Whiteley, G.; Wagstaff, S.C.; Harrison, R.A.; Casewell, N.R.; Calvete, J.J. What Killed Karl Patterson Schmidt? Combined Venom Gland Transcriptomic, Venomic and Antivenomic Analysis of the South African Green Tree Snake (the Boomslang), Dispholidus Typus. *Biochim Biophys Acta Gen Subj* **2017**, *1861*, 814–823, doi:10.1016/j.bbagen.2017.01.020.

126. Schramer, T.D.; Rautsaw, R.M.; Bayona-Serrano, J.D.; Nystrom, G.S.; West, T.R.; Ortiz-Medina, J.A.; Sabido-Alpuche, B.; Meneses-Millán, M.; Borja, M.; Junqueira-de-Azevedo, I.L.M.; et al. An Integrative View of the Toxic Potential of Conophis Lineatus (Dipsadidae: Xenodontinae), a Medically Relevant Rear-Fanged Snake. *Toxicon* **2022**, *205*, 38–52, doi:10.1016/j.toxicon.2021.11.009.

127. Campos, P.F.; Andrade-Silva, D.; Zelanis, A.; Leme, A.F.P.; Rocha, M.M.T.; Menezes, M.C.; Serrano, S.M.T.; Junqueira-De-Azevedo, I.D.L.M. Trends in the Evolution of Snake Toxins Underscored by an Integrative Omics Approach to Profile the Venom of the

Colubrid Phalotris Mertensi. *Genome Biol Evol* **2016**, *8*, 2266–2287, doi:10.1093/gbe/evw149.

128. Ching, A.T.C.; Rocha, M.M.T.; Paes Leme, A.F.; Pimenta, D.C.; de Fátima D. Furtado, M.; Serrano, S.M.T.; Ho, P.L.; Junqueira-de-Azevedo, I.L.M. Some Aspects of the Venom Proteome of the Colubridae Snake Philodryas Olfersii Revealed from a Duvernoy's (Venom) Gland Transcriptome. *FEBS Lett* **2006**, *580*, 4417–4422, doi:10.1016/j.febslet.2006.07.010.

129. Ching, A.T.C.; Paes Leme, A.F.; Zelanis, A.; Rocha, M.M.T.; Furtado, M.D.F.D.; Silva, D.A.; Trugilho, M.R.O.; da Rocha, S.L.G.; Perales, J.; Ho, P.L.; et al. Venomics Profiling of Thamnodynastes Strigatus Unveils Matrix Metalloproteinases and Other Novel Proteins Recruited to the Toxin Arsenal of Rear-Fanged Snakes. *J Proteome Res* **2012**, *11*, 1152–1162, doi:10.1021/pr200876c.

130. Nei, M.; Rooney, A.P. Concerted and Birth-and-Death Evolution of Multigene Families. *Annu Rev Genet* **2005**, *39*, 121–152, doi:10.1146/annurev.genet.39.073003.112240.

131. Yoder, A.D.; Larsen, P.A. The Molecular Evolutionary Dynamics of the Vomeronasal Receptor (Class 1) Genes in Primates: A Gene Family on the Verge of a Functional Breakdown. *Front Neuroanat* **2014**, *8*, 1–9, doi:10.3389/fnana.2014.00153.

132. Arroyo, J.I.; Hoffmann, F.G.; Opazo, J.C. Gene Turnover and Differential Retention in the Relaxin/Insulin-like Gene Family in Primates. *Mol Phylogenet Evol* **2012**, *63*, 768–776, doi:10.1016/j.ympev.2012.02.011.

133. Young, J.M.; Endicott, R.L.M.; Parghi, S.S.; Walker, M.; Kidd, J.M.; Trask, B.J. Extensive Copy-Number Variation of the Human Olfactory Receptor Gene Family. *Am J Hum Genet* **2008**, *83*, 228–242, doi:10.1016/j.ajhg.2008.07.005.

134. Jacquemin, J.; Ammiraju, J.S.S.; Haberer, G.; Billheimer, D.D.; Yu, Y.; Liu, L.C.; Rivera, L.F.; Mayer, K.; Chen, M.; Wing, R.A. Fifteen Million Years of Evolution in the Oryza Genus Shows Extensive Gene Family Expansion. *Mol Plant* **2014**, *7*, 642–656, doi:10.1093/mp/sst149.

135. Zhang, Y.; Feng, X.; Wang, L.; Su, Y.; Chu, Z.; Sun, Y. The Structure, Functional Evolution, and Evolutionary Trajectories of the H+-PPase Gene Family in Plants. *BMC Genomics* **2020**, *21*, 1–17, doi:10.1186/s12864-020-6604-2.

136. Apanius, V.; Penn, D.; Slev, P.R.; Ruff, L.R.; Potts, W.K. The Nature of Selection on the Major Histocompatibility Complex. *Crit Rev Immunol* **1997**, *17*, 179–224, doi:10.1615/CRITREVIMMUNOL.V17.I2.40.

137. Eimes, J.A.; Bollmer, J.L.; Whittingham, L.A.; Johnson, J.A.; van Oosterhout, C.; Dunn, P.O. Rapid Loss of MHC Class II Variation in a Bottlenecked Population Is Explained by Drift and Loss of Copy Number Variation. *J Evol Biol* **2011**, *24*, 1847–1856, doi:10.1111/j.1420-9101.2011.02311.x.

138. Ota, T.; Nei, M. Divergent Evolution and Evolution by the Birth-and-Death Process in the Immunoglobulin V(H) Gene Family. *Mol Biol Evol* **1994**, *11*, 469–482, doi:10.1093/oxfordjournals.molbev.a040127.

139. Innan, H.; Kondrashov, F. The Evolution of Gene Duplications: Classifying and Distinguishing between Models. *Nat Rev Genet* **2010**, *11*, 97–108, doi:10.1038/nrg2689.

140. Perry, G.H.; Dominy, N.J.; Claw, K.G.; Lee, A.S.; Fiegler, H.; Redon, R.; Werner, J.; Villanea, F.A.; Mountain, J.L.; Misra, R.; et al. Diet and the Evolution of Human Amylase Gene Copy Number Variation. *Nat Genet* **2007**, *39*, 1256–1260, doi:10.1038/ng2123.

141. Pajic, P.; Pavlidis, P.; Dean, K.; Neznanova, L.; Romano, R.A.; Garneau, D.; Daugherity, E.; Globig, A.; Ruhl, S.; Gokcumen, O. Independent Amylase Gene Copy Number Bursts Correlate with Dietary Preferences in Mammals. *Elife* **2019**, *8*, doi:10.7554/eLife.44628.

142. Wessinger, C.A.; Rausher, M.D. Lessons from Flower Colour Evolution on Targets of Selection. *J Exp Bot* 2012, *63*, 5741–5749.

143. Kronforst, M.R.; Barsh, G.S.; Kopp, A.; Mallet, J.; Monteiro, A.; Mullen, S.P.; Protas, M.; Rosenblum, E.B.; Schneider, C.J.; Hoekstra, H.E. Unraveling the Thread of Nature's Tapestry: The Genetics of Diversity and Convergence in Animal Pigmentation. *Pigment Cell Melanoma Res* 2012, *25*, 411–433.

144. Khalturin, K.; Anton-Erxleben, F.; Sassmann, S.; Wittlieb, J.; Hemmrich, G.; Bosch, T.C.G. A Novel Gene Family Controls Species-Specific Morphological Traits in Hydra. *PLoS Biol* **2008**, *6*, 2436–2449, doi:10.1371/journal.pbio.0060278.

145. Visscher, P.M.; Wray, N.R.; Zhang, Q.; Sklar, P.; McCarthy, M.I.; Brown, M.A.; Yang, J. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* 2017, *101*, 5–22.

146. Mathieson, I. The Omnigenic Model and Polygenic Prediction of Complex Traits. *Am J Hum Genet* 2021, *108*, 1558–1563.

147. Casewell, N.R.; Harrison, R.A.; Wüster, W.; Wagstaff, S.C. Comparative Venom Gland Transcriptome Surveys of the Saw-Scaled Vipers (Viperidae: Echis) Reveal Substantial Intra-Family Gene Diversity and Novel Venom Transcripts. *BMC Genomics* **2009**, *10*, doi:10.1186/1471-2164-10-564.

148. Casewell, N.R.; Huttley, G.A.; Wüster, W. Dynamic Evolution of Venom Proteins in Squamate Reptiles. *Nat Commun* **2012**, *3*, doi:10.1038/ncomms2065.

149. Juárez, P.; Comas, I.; González-Candelas, F.; Calvete, J.J. Evolution of Snake Venom Disintegrins by Positive Darwinian Selection. *Mol Biol Evol* **2008**, *25*, 2391–2407, doi:10.1093/molbev/msn179.

150. Gibbs, H.L.; Rossiter, W. Rapid Evolution by Positive Selection and Gene Gain and Loss: PLA 2 Venom Genes in Closely Related Sistrurus Rattlesnakes with Divergent Diets. *J Mol Evol* **2008**, *66*, 151–166, doi:10.1007/s00239-008-9067-7.

151. Margres, M.J.; Walls, R.; Suntravat, M.; Lucena, S.; Sánchez, E.E.; Rokyta, D.R. Functional Characterizations of Venom Phenotypes in the Eastern Diamondback Rattlesnake (Crotalus Adamanteus) and Evidence for Expression-Driven Divergence in Toxic Activities among Populations. *Toxicon* **2016**, *119*, 28–38, doi:10.1016/j.toxicon.2016.05.005.

152. Nachtigall, P.G.; Freitas-De-sousa, L.A.; Mason, A.J.; Moura-Da-silva, A.M.; Grazziotin, F.G.; Junqueira-De-azevedo, I.L.M. Differences in PLA2 Constitution Distinguish the Venom of Two Endemic Brazilian Mountain Lanceheads, Bothrops Cotiara and Bothrops Fonsecai. *Toxins (Basel)* **2022**, *14*, 237, doi:10.3390/toxins14040237.

153. Ogawa, T.; Chijiwa, T.; Oda-Ueda, N.; Ohno, M. Molecular Diversity and Accelerated Evolution of C-Type Lectin-like Proteins from Snake Venom. *Toxicon* **2005**, *45*, 1–14, doi:10.1016/j.toxicon.2004.07.028.

154. Tian, H.; Liu, M.; Li, J.; Xu, R.; Long, C.; Li, H.; Mwangi, J.; Lu, Q.; Lai, R.; Shen, C. Snake C-Type Lectins Potentially Contribute to the Prey Immobilization in Protobothrops Mucrosquamatus and Trimeresurus Stejnegeri Venoms. *Toxins (Basel)* **2020**, *12*, 105, doi:10.3390/toxins12020105.

155. Lu, Q.; Navdaev, A.; Clemetson, J.M.; Clemetson, K.J. Snake Venom C-Type Lectins Interacting with Platelet Receptors. Structure-Function Relationships and Effects on Haemostasis. *Toxicon* **2005**, *45*, 1089–1098, doi:10.1016/j.toxicon.2005.02.022.

156. Earl, S.T.H.; Robson, J.; Trabi, M.; de Jersey, J.; Masci, P.P.; Lavin, M.F. Characterisation of a Mannose-Binding C-Type Lectin from Oxyuranus Scutellatus Snake Venom. *Biochimie* **2011**, *93*, 519–527, doi:10.1016/j.biochi.2010.11.006.

157. Nachtigall, P.G.; Rautsaw, R.M.; Ellsworth, S.A.; Mason, A.J.; Rokyta, D.R.; Parkinson, C.L.; Junqueira-de-Azevedo, I.L.M. ToxCodAn: A New Toxin Annotator and Guide to Venom Gland Transcriptomics. *Brief Bioinform* **2021**, *22*, 1–16, doi:10.1093/bib/bbab095.

158. Nachtigall, P.G.; Kashiwabara, A.Y.; Durham, A.M. CodAn: Predictive Models for Precise Identification of Coding Regions in Eukaryotic Transcripts. *Brief Bioinform* **2021**, *22*, 1–11, doi:10.1093/bib/bbaa045.

159. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R.; Teeling, E. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol* **2020**, *37*, 1530–1534, doi:10.1093/molbev/msaa015.

160. Kalyaanamoorthy, S.; Minh, B.Q.; Wong, T.K.F.; von Haeseler, A.; Jermiin, L.S. ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. *Nat Methods* **2017**, *14*, 587–589, doi:10.1038/nmeth.4285.

161. Minh, B.Q.; Nguyen, M.A.T.; von Haeseler, A. Ultrafast Approximation for Phylogenetic Bootstrap. *Mol Biol Evol* **2013**, *30*, 1188–1195, doi:10.1093/molbev/mst024.

162. Jombart, T.; Balloux, F.; Dray, S. Adephylo: New Tools for Investigating the Phylogenetic Signal in Biological Traits. *Bioinformatics* **2010**, *26*, 1907–1909, doi:10.1093/bioinformatics/btq292.

163. Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol Biol Evol* **2007**, *24*, 1586–1591, doi:10.1093/molbev/msm088.

164. Nielsen, R.; Yang, Z. Likelihood Models for Detecting Positively Selected Amino Acid Sites and Applications to the HIV-1 Envelope Gene. *Genetics* **1998**, *148*, 929–936, doi:10.1093/genetics/148.3.929.

165. Kosakovsky Pond, S.L.; Poon, A.F.Y.; Velazquez, R.; Weaver, S.; Hepler, N.L.; Murrell, B.; Shank, S.D.; Magalis, B.R.; Bouvier, D.; Nekrutenko, A.; et al. HyPhy 2.5 - A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Mol Biol Evol* **2020**, *37*, 295–299, doi:10.1093/molbev/msz197.

166. Murrell, B.; Wertheim, J.O.; Moola, S.; Weighill, T.; Scheffler, K.; Kosakovsky Pond, S.L. Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLoS Genet* **2012**, *8*, e1002764, doi:10.1371/journal.pgen.1002764.

167. Murrell, B.; Moola, S.; Mabona, A.; Weighill, T.; Sheward, D.; Kosakovsky Pond, S.L.; Scheffler, K. FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection. *Mol Biol Evol* **2013**, *30*, 1196–1205, doi:10.1093/molbev/mst030.

168. Durban, J.; Sanz, L.; Trevisan-Silva, D.; Neri-Castro, E.; Alagón, A.; Calvete, J.J. Integrated Venomics and Venom Gland Transcriptome Analysis of Juvenile and Adult Mexican Rattlesnakes Crotalus Simus, C. Tzabcan, and C. Culminatus Revealed MiRNA-Modulated Ontogenetic Shifts. *J Proteome Res* **2017**, *16*, 3370–3390, doi:10.1021/acs.jproteome.7b00414.

169. Guimarães-Gomes, V.; Oliveira-Carvalho, A.L.; Inácio, I. de; Denis, D.L.; Pujol-Luz, M.; Castro, H.C.; Ho, P.L.; Zingali, R.B. Cloning, Characterization, and Structural Analysis of a C-Type Lectin from Bothrops Insularis (BiL) Venom. *Arch Biochem Biophys* **2004**, *432*, 1–11, doi:10.1016/j.abb.2004.08.018.

170. Hargreaves, A.D.; Swain, M.T.; Logan, D.W.; Mulley, J.F. Testing the Toxicofera: Comparative Transcriptomics Casts Doubt on the Single, Early Evolution of the Reptile Venom System. *Toxicon* **2014**, *92*, 140–156, doi:10.1016/j.toxicon.2014.10.004.

171. Reyes-Velasco, J.; Card, D.C.; Andrew, A.L.; Shaney, K.J.; Adams, R.H.; Schield, D.R.; Casewell, N.R.; Mackessy, S.P.; Castoe, T.A. Expression of Venom Gene Homologs in Diverse Python Tissues Suggests a New Model for the Evolution of Snake Venom. *Mol Biol Evol* **2015**, *32*, 173–183, doi:10.1093/molbev/msu294.

172. Komori, Y.; Hifumi, T.; Yamamoto, A.; Sakai, A.; Ato, M.; Sawabe, K.; Nikai, T. Comparative Study of Biological Activities of Venom from Colubrid Snakes Rhabdophis Tigrinus (Yamakagashi) and Rhabdophis Lateralis. *Toxins (Basel)* **2017**, *9*, 373, doi:10.3390/toxins9110373.

173. Brahma, R.K.; McCleary, R.J.R.; Kini, R.M.; Doley, R. Venom Gland Transcriptomics for Identifying, Cataloging, and Characterizing Venom Proteins in Snakes. *Toxicon* **2015**, *93*, 1–10, doi:10.1016/j.toxicon.2014.10.022.

174. Pawlak, J.; Mackessy, S.P.; Sixberry, N.M.; Stura, E.A.; Le Du, M.H.; Ménez, R.; Foo, C.S.; Ménez, A.; Nirthanan, S.; Kini, R.M. Irditoxin, a Novel Covalently Linked Heterodimeric Three-Finger Toxin with High Taxon-Specific Neurotoxicity. *FASEB J* **2009**, *23*, 534–545, doi:10.1096/FJ.08-113555.

175. Pawlak, J.; Mackessy, S.P.; Fry, B.G.; Bhatia, M.; Mourier, G.; Fruchart-Gaillard, C.; Servent, D.; Ménez, R.; Stura, E.; Ménez, A.; et al. Denmotoxin, a Three-Finger Toxin from the Colubrid Snake Boiga Dendrophila (Mangrove Catsnake) with Bird-Specific Activity. *Journal of Biological Chemistry* **2006**, *281*, 29030–29041, doi:10.1074/jbc.M605850200.

176. Modahl, C.M.; Brahma, R.K.; Koh, C.Y.; Shioi, N.; Kini, R.M. Omics Technologies for Profiling Toxin Diversity and Evolution in Snake Venom: Impacts on the Discovery of Therapeutic and Diagnostic Agents. *Annu Rev Anim Biosci* 2020, *8*, 91–116.

177. Modahl, C.M.; Mackessy, S.P. Full-Length Venom Protein CDNA Sequences from Venom-Derived MRNA: Exploring Compositional Variation and Adaptive Multigene Evolution. *PLoS Negl Trop Dis* **2016**, *10*, 1–24, doi:10.1371/journal.pntd.0004587.

178. Chang, C.C.; Lee, C.Y. ISOLATION OF NEUROTOXINS FROM THE VENOM OF BUNGARUS MULTICINCTUS AND THEIR MODES OF NEUROMUSCULAR BLOCKING ACTION. *Arch Int Pharmacodyn Ther* **1963**, *144*, 241–257, doi:10.1016/0041-0101(65)90032-2.

179. Utkin, Y.N. Last Decade Update for Three-Finger Toxins: Newly Emerging Structures and Biological Activities. *World J Biol Chem* **2019**, *10*, 17–27, doi:10.4331/wjbc.v10.i1.17.

180. Dashevsky, D.; Fry, B.G. Ancient Diversification of Three-Finger Toxins in Micrurus Coral Snakes. *J Mol Evol* **2018**, *86*, 58–67, doi:10.1007/s00239-017-9825-5.

181. Dashevsky, D.; Rokyta, D.; Frank, N.; Nouwens, A.; Fry, B.G. Electric Blue: Molecular Evolution of Three-Finger Toxins in the Long-Glanded Coral Snake Species Calliophis Bivirgatus. *Toxins (Basel)* **2021**, *13*, 124, doi:10.3390/toxins13020124.

182. Fry, B.G.; Lumsden, N.G.; Wüster, W.; Wickramaratna, J.C.; Hodgson, W.C.; Manjunatha Kini, R. Isolation of a Neurotoxin (α-Colubritoxin) from a Nonvenomous Colubrid: Evidence for Early Origin of Venom in Snakes. *J Mol Evol* **2003**, *57*, 446–452, doi:10.1007/s00239-003-2497-3.

183. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-Length Transcriptome Assembly from RNA-Seq Data without a Reference Genome. *Nat Biotechnol* **2011**, *29*, 644–652, doi:10.1038/nbt.1883.

184. Fu, L.; Niu, B.; Zhu, Z.; Wu, S.; Li, W. CD-HIT: Accelerated for Clustering the next-Generation Sequencing Data. *Bioinformatics* **2012**, *28*, 3150–3152, doi:10.1093/BIOINFORMATICS/BTS565.

185. Li, B.; Dewey, C.N. RSEM: Accurate Transcript Quantification from RNA-Seq Data with or without a Reference Genome. *BMC Bioinformatics* **2011**, *12*, 1–16, doi:10.1186/1471-2105-12-323.

186. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. BLAST+: Architecture and Applications. *BMC Bioinformatics* **2009**, *10*, 1–9, doi:10.1186/1471-2105-10-421/FIGURES/4.

187. Edgar, R.C. MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Res* **2004**, *32*, 1792–1797, doi:10.1093/NAR/GKH340.

188. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data. *Bioinformatics* **2012**, *28*, 1647–1649, doi:10.1093/BIOINFORMATICS/BTS199.

189. Lanfear, R.; Calcott, B.; Ho, S.Y.W.; Guindon, S. Partitionfinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. *Mol Biol Evol* **2012**, *29*, 1695–1701, doi:10.1093/MOLBEV/MSS020.

190. Ronquist, F.; Huelsenbeck, J.P. MrBayes 3: Bayesian Phylogenetic Inference under Mixed Models. *Bioinformatics* **2003**, *19*, 1572–1574, doi:10.1093/BIOINFORMATICS/BTG180.

191. Gao, F.; Chen, C.; Arab, D.A.; Du, Z.; He, Y.; Ho, S.Y.W. EasyCodeML: A Visual Tool for Analysis of Selection Using CodeML. *Ecol Evol* **2019**, *9*, 3891–3898, doi:10.1002/ece3.5015.

192. Yang, Z.; Nielsen, R.; Goldman, N.; Pedersen, A.-M.K. Codon-Substitution Models for Heterogeneous Selection Pressure at Amino Acid Sites. *Genetics* **2000**, *155*, 431–449, doi:10.1093/genetics/155.1.431.

193. Modahl, C.M.; Saviola, A.J.; Mackessy, S.P. Integration of Transcriptomic and Proteomic Approaches for Snake Venom Profiling. *Expert Rev Proteomics* **2021**, *18*, 827–834, doi:10.1080/14789450.2021.1995357.

194. Heyborne, W.H.; Mackessy, S.P. Venoms of New World Vinesnakes (Oxybelis Aeneus and O. Fulgidus). *Toxicon* **2021**, *190*, 22–30, doi:10.1016/j.toxicon.2020.12.002.

195. Dashevsky, D.; Debono, J.; Rokyta, D.; Nouwens, A.; Josh, P.; Fry, B.G. Three-Finger Toxin Diversification in the Venoms of Cat-Eye Snakes (Colubridae: Boiga). *J Mol Evol* **2018**, *86*, 531–545, doi:10.1007/s00239-018-9864-6.

196. Calvete, J.J.; Bonilla, F.; Granados-Martínez, S.; Sanz, L.; Lomonte, B.; Sasa, M. Venomics of the Duvernoy's Gland Secretion of the False Coral Snake Rhinobothryum Bovallii (Andersson, 1916) and Assessment of Venom Lethality towards Synapsid and

Diapsid Animal Models. *J Proteomics* **2020**, *225*, 103882, doi:10.1016/j.jprot.2020.103882.

197. Zhang, Z.Y.; Lv, Y.; Wu, W.; Yan, C.; Tang, C.Y.; Peng, C.; Li, J.T. The Structural and Functional Divergence of a Neglected Three-Finger Toxin Subfamily in Lethal Elapids. *Cell Rep* **2022**, *40*, 111079, doi:10.1016/J.CELREP.2022.111079.

198. Fry, B.G.; Lumsden, N.G.; Wüster, W.; Wickramaratna, J.C.; Hodgson, W.C.; Manjunatha Kini, R. Isolation of a Neurotoxin (α-Colubritoxin) from a Nonvenomous Colubrid: Evidence for Early Origin of Venom in Snakes. *J Mol Evol* **2003**, *57*, 446–452, doi:10.1007/S00239-003-2497-3/FIGURES/4.

199. Sunagar, K.; Jackson, T.N.W.; Undheim, E.A.B.; Ali, S.A.; Antunes, A.; Fry, B.G. Three-Fingered RAVERs: Rapid Accumulation of Variations in Exposed Residues of Snake Venom Toxins. *Toxins (Basel)* **2013**, *5*, 2172–2208, doi:10.3390/TOXINS5112172.

200. Hargreaves, A.D.; Swain, M.T.; Hegarty, M.J.; Logan, D.W.; Mulley, J.F. Restriction and Recruitment-Gene Duplication and the Origin and Evolution of Snake Venom Toxins. *Genome Biol Evol* **2014**, *6*, 2088–2095, doi:10.1093/gbe/evu166.

201. Brust, A.; Sunagar, K.; Undheim, E.A.B.; Vetter, I.; Yang, D.C.; Casewell, N.R.; Jackson, T.N.W.; Koludarov, I.; Alewood, P.F.; Hodgson, W.C.; et al. Differential Evolution and Neofunctionalization of Snake Venom Metalloprotease Domains. *Molecular and Cellular Proteomics* **2013**, *12*, 651–663, doi:10.1074/mcp.M112.023135.

202. Pla, D.; Petras, D.; Saviola, A.J.; Modahl, C.M.; Sanz, L.; Pérez, A.; Juárez, E.; Frietze, S.; Dorrestein, P.C.; Mackessy, S.P.; et al. Transcriptomics-Guided Bottom-up and Top-down Venomics of Neonate and Adult Specimens of the Arboreal Rear-Fanged Brown

Treesnake, Boiga Irregularis, from Guam. *J Proteomics* **2018**, *174*, 71–84, doi:10.1016/J.JPROT.2017.12.020.

203. Davies, E.L.; Arbuckle, K. Coevolution of Snake Venom Toxic Activities and Diet: Evidence That Ecological Generalism Favours Toxicological Diversity. *Toxins (Basel)* **2019**, *11*, 711, doi:10.3390/toxins11120711.