8-2022

The Presence and Distribution of Crotoxin in the Rock Rattlesnake (Crotalus lepidus)

Jade Mellor
njmello@g.clemson.edu

THE PRESENCE AND DISTRIBUTION OF CROTOXIN IN THE ROCK RATTLESNAKE (*CROTALUS LEPIDUS*)

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Biological Sciences

---

by
Nikita Jade Mellor
August 2022

Accepted by:
Christopher L. Parkinson, Committee Chair
Vincent Richards
Christopher Saski

ABSTRACT

Crotoxin and its homologs (hereafter all referred to as CTx) is a highly lethal heterodimeric beta-neurotoxin found in pitvipers (Crotalinae) and is the main driver of neurotoxic venom phenotypes (Type II). In contrast, hemorrhagic venom phenotypes (Type I) are characterized by high snake venom metalloproteinase expression and low toxicity. Although many rattlesnake species have been classified as either Type I or Type II, population level variation in venom phenotype has also been documented in several species. The presence or absence of CTx is the main component of this variation in venom phenotype and has been most widely studied in large-bodied lowland rattlesnakes (*Crotalus scutulatus, C. helleri, and C. horridus*). While it has been suspected to be in *C. lepidus*, a small-bodied montane rattlesnake, there has been no genetic confirmation. We used genomics and transcriptomics to test for the presence, distribution, and evolution of CTx in *C. lepidus*. We genomically and transcriptomically confirmed the presence and expression of CTx in *C. lepidus* and found it in 17 out of 104 samples across their range. CTx presence was not significantly associated with longitude, latitude, subspecies, or elevation. However, we did identify several climatic variables associated with CTx presence, including ones that have been identified in previous studies on CTx expression providing insights on the phylogenetic distribution of CTx across rattlesnakes, the variation in crotoxin expression, and highlighting environments to which CTx may be locally adapted. Our results likely support previous hypotheses of an ancestral origin for crotoxin followed by independent sorting in lineages; therefore, future studies should focus on testing for the presence of CTx in other species of montane rattlesnakes.

DEDICATION

For those that couldn’t be here.

ACKNOWLEDGMENTS

I liked my 8-week trial of Clemson so much I came back for my Master's. First, I would like to thank my advisor and committee members. Dr. Parkinson, thank you for supporting me in this project, and for playing an integral role in my path to research. When I started as a summer REU intern, I wanted to get my DVM. Instead, this fall I will be pursuing a PhD, with a bioinformatic toolbox. I would like to thank my thesis committee, Dr. Richards, and Dr. Saski, for their patience and encouragement during this process. I would also like to thank Dr. Mark Margres, Dr. Darin Rokyta, and Dr. Kenny Wray for donating data to my project so that I could fully investigate these questions. I would also like to thank Dr. Victoria Corbin and Dr. Mike Sears for organizing the REU program in 2018 that served as the preliminary data for my Master's project.

Second, I would like to thank the Parkinson Lab for their incredible support and wealth of knowledge. Dr. Ed Myers, Dr. Rhett Rautsaw, Tristan Schramer, and Ramses Alejandro Rosales-Garcia, your scientific knowledge and emotional support during my time at Clemson have been immeasurable. Thank you for your encouragement during every step of this project. I would also like to thank Dr. Andrew Mason, for showing me the ropes during my REU program and helping to set me on this track.

I would like to thank my friends, both here at Clemson for making these last two years memorable and fun. My mother and brother, who have listened to me talk about snakes for hours, thank you. I would also like to thank my partner, Savannah, for being the best rubber duck a girl could ask for.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

List of Figures (Continued)

Figure Page

List of Figures (Continued)

Figure Page

CHAPTER ONE

INTRODUCTION

Genes and Adaptation

Phenotypes are one of the most visible products of selection. Researchers have been able to watch traits shift in real time as selection acts (Price et al. 1984), they have been ensnared by genetic and physical changes (Labonne and Hendry 2010; Lamichhaney et al. 2018). When unique traits emerge, they become the subjects of interest and scrutiny (Green et al. 2010; Ryder et al. 2021). This leads to more discoveries and a better understanding of the natural world (Slon et al. 2018; Liu et al. 2021; Mahadevan et al. 2021). A trait can also be of particular interest when traits are differentially expressed between or within populations of a single clade. For example, biofluorescence has been found in 105 different genera across the phylogeny of extant ray finned fishes (Sparks et al. 2014). Since the initial discovery of biofluorescent fish (Harvey 1921), more studies have built upon that knowledge, discovering different molecular mechanisms (Guarnaccia et al. 2021) leading to the same phenotype and its different ecological functions (Salih et al. 2000; Tsutsui et al. 2016).

Phenotype is ultimately controlled by genotype. Genes can originate through duplication events (Carretero-Paulet and Fares 2012), be maintained through positive (Chang and Duda 2012) or purifying (Carretero-Paulet and Fares 2012) selection, or lost through deletion events (Dowell et al. 2016). For example, conotoxin, a neurotoxin found in marine cone snails evolved through duplication events and neofunctionalization, offering redundancy and opportunities for beneficial mutations (Chang and Duda 2012). Genes are further controlled by differential genetic architecture, including the number of loci and effect size (Flint and Mackay 2009). The more complex the trait, the more genes involved for a single phenotype, the more likely that each loci

contributes a small amount of phenotypic variation seen (Flint and Mackay 2009). However, there can be a small number of loci that contribute a significant amount to the phenotypic variation (Robertson 1967; Flint and Mackay 2009). Epigenetic controls of expression can also regulate phenotype. Methylation and chromatic accessibility can allow or block access to genes, regulating their expression (Margres, Rautsaw, et al. 2021). Regulating expression allows a gene to still be present in the genome but not expressed, which can be especially useful in response to changing environments (Schlichting and Pigliucci 1995).

Rattlesnakes and Crotoxin

Snake venoms are complex phenotypes composed of varying abundances of many different proteins. An individual species' venom may contain anywhere between 10 to 70+ different toxins (Mackessy 2008; Mackessy 2010; Holding et al. 2021). Variation in snake venom composition is largely due to factors related to diet and predation (Holding et al. 2016; Rokyta et al. 2017; Holding et al. 2021). Venom is frequently categorized into two types in pitvipers: Type I and Type II. Type I venoms are considered highly hemotoxic, featuring high expression of the snake venom metalloproteinase (SVMP) gene family (Mackessy 2010). In short, they cause profuse bleeding and tissues degradation and are less lethal in small quantities. In contrast, Type II venoms have low to no SVMP expression, but high expression of certain phospholipase A2s (PLA2s) which result in highly neurotoxic effects (Mackessy 2010). Comparatively, Type II venoms are more lethal than their Type I counterparts. Other toxin families are also present in varying degrees in either phenotype. Snake venom serine proteinases (SVSPs) interfere with hemostatic systems, and can be present in both venom phenotypes (Mackessy 2010). C-type lectins (CTLs) are more common in Type I phenotypes and target

blood cells and other plasma components (Mackessy 2010). Bradykin-potentiating peptides (BPPs) cause pain and immobilization, and can be present in either phenotype (Mackessy 2010).

Crotoxin (CTx) is a heterodimeric PLA2 that is responsible for the high lethality associated with Type II venoms (Mackessy 2010). Originally characterized in *Crotalus durissus* (Slotta and Fraenkel-Conrat 1939) crotoxin homologs have been found in multiple species, and given a variety of names such as canebrake toxin in *Crotalus horridus* (Rokyta et al. 2013), concolortoxin in *C. concolor* (Aird and Kaiser 1985), and crotoxin in *C. helleri* and *C. tigris* (French et al. 2004; Calvete et al. 2012; Franco-Servín et al. 2021). For the purposes of this thesis, "crotoxin" will refer to all homologs across species. This beta-neurotoxin has two subunits, the acidic subunit (PLA2-gA2) acts as a chaperon for the basic subunit (PLA2-gB2) (Hendon and Fraenkel-Conrat 1971; Faure et al. 2011; Whittington et al. 2018). Individually, the basic subunit is mildly toxic but when found in tandem with the acidic chaperone results in high lethality. Crotoxin works presynaptically to cause immobilization via disruption of nerve cells. Medically, this results in respiratory paralysis, paresthesia, and muscle twitching (Gopalakrishnakone et al. 1980; Massey et al. 2012; Neri-Castro et al. 2019)**.** Crotoxin is most commonly seen in the New World pitvipers. It is hypothesized to be ancient in origin and is found in the Old World pitviper genus *Gloydius* which is the sister group to the New World radiation (Wüster et al. 2008; Yang, Guo, et al. 2015; Alencar et al. 2016; Dowell et al. 2016; Whittington et al. 2018). Crotoxin is further hypothesized to have evolved via PLA2 gene duplication, evolving ancestral basic and acidic subunits. These ancestral subunits originally only weakly interacted, and were thought to be mildly toxic. A single large-effect mutation in the ancestral acidic subunit is hypothesized to allow for proteolytic processing by SVSPs already present in the venom. This single mutation led to episodic, diversifying selection in the ancestral

basic subunit, to capitalize on the new form of the acidic subunit. Upon dimerization, both subunits are hypothesized to have gone through purifying selection (Whittington et al. 2018). In terms of gene expression adaptations, it has been found that transcripts of the acidic subunit are present at double the numbers of the basic subunit, which would allow for the highest utilization of the basic subunit (Whittington et al. 2018). While the PLA2 gene family is conserved in pitvipers, duplications, deletions, and differential expression of individual PLA2 genes facilitates diversification and variation (see Figure 1).

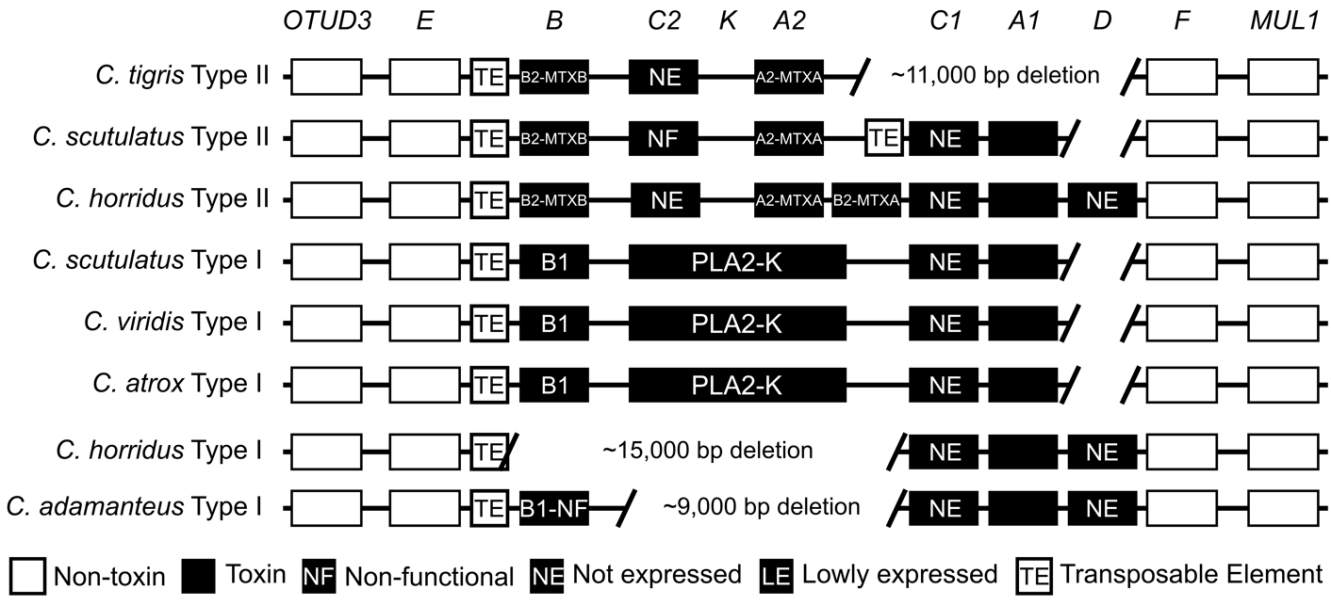

Figure 1. From Margres et. al (2021), shows the PLA2 arrays in different species of *Crotalus*. Comparing Type I and Type II individuals of the same species and shows differing PLA2 haplotypes depending on venom phenotype.

Crotoxin was thought to only be found within the rattlesnakes (genus *Crotalus*) in the New World; however, it has now been described in several other pitviper genera (*Sistrurus* (Calvete et al. 2012)*, Ophryacus* (Neri-Castro et al. 2019)*, Bothriechis* (Fernández et al. 2010), and *Gloydius* (Yang, Yang, et al. 2015)). Additionally, there is inter- and intraspecific variation in crotoxin expression, resulting in a single species being polymorphic for a venom phenotype. This intraspecific variation in crotoxin is best studied in the crotoxin homolog found in *Crotalus scutulatus* called Mojave toxin (Glenn and Straight 1989; Strickland, Mason, et al. 2018; Strickland, Smith, et al. 2018). The northern distribution of *C. scutulatus* has large swaths of a Type II phenotype, with isolated Type I populations. Further south in the distribution, *C. scutulatus* has a mostly Type I expression with isolated Type II. In a third region, the pattern switches again (see Figure 2). Strickland, Smith, et al. (2018) determined that local selective pressures play a role in maintaining both phenotypes within the species. However, *C. scutulatus* is not the only species with variable venom phenotype. *C. horridus* (Rokyta et al. 2013; Margres, Wray, et al. 2021), *C. concolor* (Aird and Kaiser 1985), and *C. helleri* (French et al. 2004; Franco-Servín et al. 2021) are known to have populations that express a crotoxin homolog in a typically Type I expressing species. Much of the research regarding variable phenotypes has been done in large-bodied, lowland clade of rattlesnakes that are abundant and medically important. Venom variation within smaller montane rattlesnake species generally has not been studied, largely because human envenomation is rare.

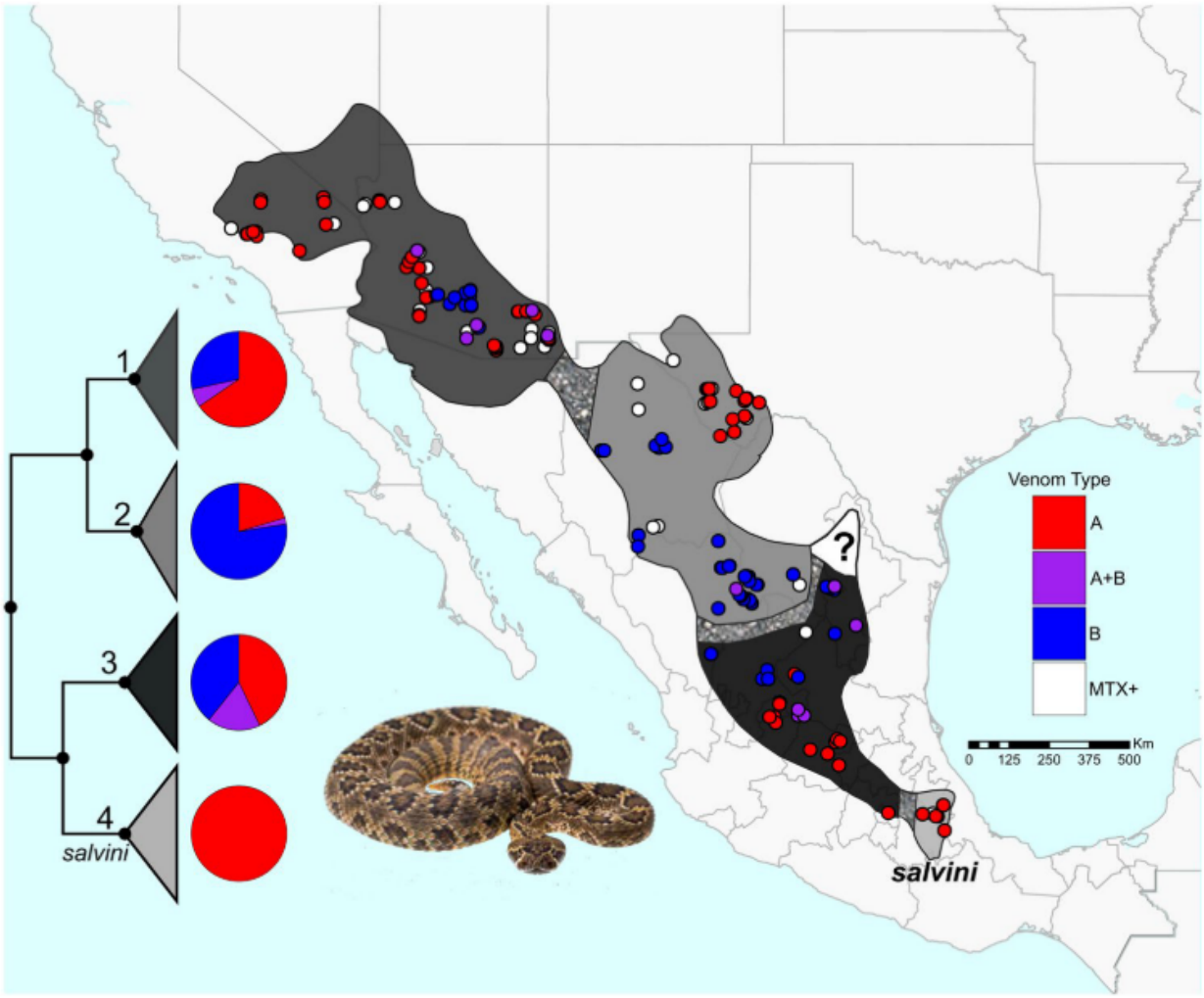

Figure 2. From Strickland, Smith, et. al (2018), shows the distribution of individuals with (Type A) and without (Type B) crotoxin in *C. scutulatus* throughout the range.

Study Species

The Rock Rattlesnake (*Crotalus lepidus*) is a montane rattlesnake found in the American Southwest as well as north-central Mexico. There are three subspecies, *C. l. lepidus, C. l. klauberi,* and *C. l. maculosus* (Figure 3). *Crotalus l. lepidus* makes up the eastern distribution of the species, found in desert rocky outcrops and arroyos (Mata-Silva et al. 2018). *Crotalus l. klauberi* is found on the western half of the distribution, typically being found in the rocky outcrops of pine forest slopes (Holycross et al. 2002). The third subspecies, *C. l. maculosus*, has a small range at the junction of the Mexican states of Durango, Sinaloa, and Nayarit, occupying pine oak forests (Tanner et al. 1972; Armstrong and Murphy 1979). Overall, the species can be found at elevations of 200-2930m (Holycross et al. 2002). In parts of Arizona and New Mexico, in the range of *C. l. klauberi*, the species is found in the Madrean Sky Islands. The Madrean Sky islands are a mountain range of isolated pine forest peaks, separated by lowland desert. This drastic change in habitat type is caused by dramatic changes in elevation. Over the course of 10 miles, the elevation can drop by 1000 feet. Thus, montane species can be physically close to their lowland counterparts.

Crotalus lepidus is widely considered to have Type I venom across all subspecies (Martínez-Romero et al. 2013; Saviola et al. 2017). However, Rael et al. (1992) showed that venom from some individuals of *C. lepidus* were reactive to the crotoxin antibody CSS12, and had significantly lower LD_{50}s, indicating a higher lethality and supporting potential crotoxin expression in this species. These CSS12 reactive individuals were all represented by the subspecies *C. l. klauberi* in Chihuahua, Mexico as well as some individuals in Cochise County, Arizona, and Hidalgo County, New Mexico. Other studies have shown significant venom and

lethality variation by subspecies and locality in *C. lepidus* (Forstner et al. 1997; Martínez-Romero et al. 2013; Rivas et al. 2017).

Figure 3. The distribution of *C. lepidus* subspecies, *C. l. lepidus* (pink), *C. l. klauberi* (green), and *C. l. maculosus* (orange). Made using VenomMaps (Rautsaw et al. 2022).

Objective

Here, I propose to test for the presence, distribution, and origin of a crotoxin homolog in *C. lepidus* using both transcriptomic and genetic data. First, (Q1) I will use venom gland transcriptomes to test for the presence and expression of crotoxin and use phylogenetics to confirm its identity. Second, (Q2) using both transcriptomic and hybrid enrichment genomic data, I will analyze the spatial distribution of crotoxin in *C. lepidus* and test if subspecies, location, or various environmental factors are associated with its presence. Finally, (Q3) I will test if crotoxin's presence in *C. lepidus* is a result of i) independent origin via convergent evolution, ii) evolution of a shared ancestral crotoxin form, or iii) hybridization with another co-distributed species with crotoxin. The spatial distribution and origin analyses will help to further illuminate the evolutionary history of this complicated trait, in addition to having implications for the severity and treatment of snake bites across the range of *C. lepidus*.

CHAPTER TWO

IS CROTOXIN PRESENT AND EXPRESSED IN *CROTALUS LEPIDUS*

Introduction

Phospholipase A2s are a large gene family with five main isoforms present in *Crotalus* (Dowell et al. 2016; Dowell et al. 2018; Whittington et al. 2018). It is hypothesized that each PLA2 paralog arose through a series of duplications with the basic isoforms evolving first, followed by the acidic (Dowell et al. 2016; Whittington et al. 2018). Within the basic isoforms, PLA2-gKs are characterized by a myotoxic Asp to Lys substitution at position 49 known as Asp49Lys (Lomonte et al. 2009; Lomonte and Rangel 2012). Duplications led to the evolution of the basic isoforms PLA2-gB1 and PLA2-gB2, both of which are mildly toxic and the latter being the basic crotoxin subunit (Dowell et al. 2016; Whittington et al. 2018). Additional duplication events have also led to the acidic isoforms, PLA2-gA1 and PLA2-gA2, the latter being the crotoxin chaperone subunit (Whittington et al. 2018). These duplications allowed for mutations to accumulate and neofunctionalization to occur. Crotoxin expression requires both PLA2-gB2 and PLA2-gA2 subunits to be present. Additionally, the key mutation that facilitated proteolytic cleavage of the acidic subunit was a proline to serine substitution at site 127. This substitution led to purifying selection, as the two subunits' interaction resulted in the neurotoxic phenotype we associate with Type II venoms (Whittington et al. 2018). If these two requirements are met, then the presence of crotoxin can be confirmed genetically. Although crotoxin's presence can be confirmed with sequence data, these genes can be nonfunctional, by failing to be transcribed into mRNA or translated into proteins. Therefore, transcriptomic expression data and proteomics are necessary to confirm crotoxin presence and expression in the venom. Our goal was to genomically and transcriptomically confirm the presence of crotoxin in *C. lepidus* through gene

annotation, multiple sequence alignment, and phylogenetic inference. We also used proteomics to confirm expression of crotoxin in the venom.

Materials and Methods

Sampling

For the transcriptomics and genomics work, we collected samples between June and August from 2012 to 2021, following the guidance set by ASIH (Beaupre et al. 2004), University of Central Florida IACUC (#13-17W; #16-17W), and Clemson University Animal Care and Use Committee (Animal Use Protocol 2017-067; 2020-067). We collected snakes from the field while hiking or road-cruising in appropriate habitats under the following permits: Arizona (AZ SCP# SP673390; SP735619), Texas (TX SCP# SPR-0390-029; SPR-0713-098), and Mexico (MX OFICIO NUM. SGPA/DGVS/01090/17; OFICIO NUM. SGPA/DGVS/2190/19). In total, we collected ten samples: six *C. l. klauberi* and four *C. l. lepidus* (Table 1).

Transcriptomic and Genomic Sequencing

To extract venom, we restrained snakes by coaxing them into open-ended tubes. Snakes were then allowed to move through the tube, exposing their head at the opposite end. The snake was then presented with a parafilm-covered cup to bite and deposit venom. We transferred venom to cryotubes, vacuum dehydrated or lyophilized, and stored at -20C. To euthanize snakes, we used sodium pentobarbital or MS222 four days after venom extraction when transcription in the venom gland was at its peak (Rotenberg et al. 1971). We extracted venom glands along with heart, liver, pancreas, muscle, and kidney tissue and immediately transferred them into RNAlater (Thermo Fisher Scientific, Waltham, MA, USA). We briefly stored samples at 4C before long-term storage at -80C. We preserved specimens in formalin and deposited in museums.

Table 1. Sample table for all genome and transcriptome samples including locality information, venom type, and sequencing statistics.

Specimen ID	Museum ID	Subspecies	Venom Type	County	State	Read Pairs	Merged Reads
CLP1932	ASU 36080	*Crotalus lepidus klauberi*	CTx-	Cochise	Arizona	14569348	11951461
CLP1993	ASNHC 14990	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas	10647023	17393420
CLP2201	-	*Crotalus lepidus klauberi*	CTx+	Cochise	Arizona	13333335	10995541
CLP2202	-	*Crotalus lepidus klauberi*	CTx+	Cochise	Arizona	12613354	10732564
CLP2258	CHFCB-0174	*Crotalus lepidus klauberi*	CTx+	Guerrero	Chihuahua	17383757	14707963
CLP2268	CHFCB-0183	*Crotalus lepidus lepidus*	CTx-	Aldama	Chihuahua	11720076	10946206
CLP2346	CHFCB-0258	*Crotalus lepidus klauberi*	CTx-	Torreon	Coahuila	25743450	42195206
KW1748	-	*Crotalus lepidus lepidus*	CTx-	Pecos	Texas	13354418	11040358
MM0245	-	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas	17616278	14847235
CLP2908	-	*Crotalus lepidus klauberi*	CTx+	Santiago Papasquiaro	Durango	24511180	35817354

Specimen ID	Read_Length_Mean	Read_Count	Read_N50_bp	Total_bp	Contigs	Contig_N50_Mb	Scaffolds	Scaffold_N50_Mb
CLP1932	10,447.80	4,412,907	29,279.00	46,105,110,566	3,394	2	3,394	2
CLP2201	13,485.00	2,366,410	13,766.00	31,911,030,911	1,082	15.05	1,082	15.05

We extracted RNA using a standard Trizol extraction protocol following Rokyta, Wray and Margres (2013). Briefly, RNAlater-preserved tissue samples were thawed, minced, and placed in Trizol. We used chloroform to isolate RNA in lysed cells and then purified the samples. RNA was quantified using a Qubit RNA BR Assay (Thermo Fisher Scientific, Waltham, MA, USA) and quality checked on a BioAnalyzer RNA Pico chip (Agilent Technologies, Santa Clara, CA, USA).

For cDNA library preparation, we used the NEBNext Ultra RNA Library Prep Kit (E7530S) (NEB, Ipswich, MA, USA) with the Poly(A) mRNA Magnetic Isolation Module (E7490S) (NEB, Ipswich, MA, USA). We isolated mRNA from equal parts of the left and right venom glands. Following mRNA isolation and cDNA synthesis, we targeted an insert size of 400bp with a fragmentation time of 13.5 minutes and 14 PCR cycles to yield the desired cDNA concentration. We purified cDNA using 1.8X Agencourt AMPure XP Beads and determined concentrations using a Qubit DNA BR Assay (Thermo Fisher Scientific, Waltham, MA, USA). We checked DNA quality using a Bioanalyzer with a DNA HS kit according to the manufacturer's protocols (Agilent Technologies, Santa Clara, CA, USA). Florida State University Molecular Cloning Facility performed KAPA qPCR to determine amplifiable concentrations. KAPA results were used for final concentration and pooling to ensure equal representation of each library. Pooled DNA sample concentration and quality were determined using the Qubit HS DNA Kit and Bioanalyzer, with an additional round of KAPA PCR before sequencing. Libraries were sequenced with 150 bp paired-end reads on an Illumina HiSeq 2500 of Florida State University College of Medicine Translational Science Laboratory.

We sequenced two draft genomes for *C. lepidus*, one CTx- individual (CLP1932) and one CTx+ individual (CLP2201). For the CTx- individual, we extracted DNA from ethanol-preserved blood using a phenol-chloroform extraction and sequenced using a hybrid approach. We prepared short-read Illumina libraries using the TrueSeq DNA PCR-Free Library Prep kit (Illumina, San Diego, CA, USA). We then sequenced 150 bp PE reads on a NovaSeq 6000 at the Translational Science Laboratory in the College of Medicine at Florida State University. The University of Delaware Sequencing and Genotyping Center performed library preparation and PacBio Continuous Long Read (CLR) Sequencing on four Sequel Single-Molecule Real-Time (SMRT) cells. For the CTx+ individual, we extracted DNA using a Monarch HMW DNA Extraction Kit (NEB, Ipswich, MA, USA). The University of Delaware Sequencing and Genotyping Center performed PacBio HiFi Sequencing on 1.5 Sequel II SMRT Cells.

Transcriptome Processing and Analysis

We processed venom gland transcriptomes following the ToxCodAn guide (Nachtigall et al. 2021). Briefly, we first used Trim Galore v0.6.6 (https://github.com/FelixKrueger/TrimGalore) to trim reads and PEAR v0.9.10 (Zhang et al. 2014) to merge paired-end reads. Next, we used Extender (Rokyta et al. 2012), SeqMan NGen v14.0 (DNAStar, Inc., Madison, WI, USA), and Trinity v2.9 (Haas et al. 2013) with default settings to assemble the transcriptomes following recommendations by Holding *et al.*, (2018). We removed redundancy by clustering assemblies at 100% sequence identity using cd-hit-est v.4.8.1 (Li and Godzik 2006; Fu et al. 2012). We then used ToxCodAn v1.0 (Nachtigall et al. 2021) to identify and annotate toxin sequences in the assembly and CodAn v1.0 (Nachtigall et al. 2021) to identify additional nontoxin coding sequences followed by blast to annotate these sequences from the UniProt animal database (Bateman et al. 2021). ChimeraKiller v. 0.7.3

(https://github.com/masonaj157/ChimeraKiller) was used to remove chimeric sequences. All samples were checked for internal stop codons and then clustered using cd-hit-est v.4.8.1 at 98% to reduce redundancy caused by allelic variation. To make a consensus transcriptome for *C. lepidus*, we combined all individual's transcriptomes and clustered at 97%.

We used Bowtie2 v2.4.1 (Langmead and Salzberg 2012; Langmead et al. 2019) and RSEM v1.3.3 (Li and Dewey 2011) to calculate the expression of genes in the venom gland. The data were then imported and plotted in R v4.0.2 (R CoreTeam 2021) using ToxCodAn plotting functions. We used DESeq2 v1.30.0 (Love et al. 2014) to test for differential expression across subspecies, longitude, latitude, crotoxin positivity, elevation, and the 19 WorldClim variables (Fick and Hijmans 2017). Additionally, we calculated total SVMP, PLA2, and CTL expression for each sample and then tested for total expression differences between CTx+ and CTx- individuals using a Student's T-test.

Genome Assembly and Analysis

We performed a hybrid Illumina + PacBio assembly for the CTx- genome using MaSuRCA v. 3.2.8 with default settings (Zimin et al. 2017). For the CTx+ genome sequenced with PacBio HiFi, we used hifiasm v.0.16.1-r375 with default settings (Cheng et al. 2021). For both genomes, we used EDTA v1.9.9 for transposable element annotation and repeat masking (Xu and Wang 2007; Gremme et al. 2013; Xiong et al. 2014; Smith and Hubley 2015; Ou and Jiang 2018; Ou et al. 2019; Shi and Liang 2019; Su et al. 2019; Zhang et al. 2022). We annotated each genome using Funannotate v1.8.9 (https://github.com/nextgenusfs/funannotate). Briefly, we used the Funannotate 'train' module to perform a genome-guided Trinity v.2.5.1 assembly of RNA-Seq data generated from the heart, muscle, pancreas, liver, and kidney tissues and followed this with a PASA assembly with a maximum intron length set to 30,000 bp (Haas et al. 2003;

Campbell et al. 2006; Grabherr et al. 2011; Haas et al. 2013). Next, we used the Funannotate 'predict' module to predict genes with Augustus v3.2.1 (Stanke et al. 2008) and GeneMark-ES/ET (Ter-Hovhannisyan et al. 2008), followed by consensus generation with EvidenceModeler (Haas et al. 2008). We used the training results and our *de novo* assembled consensus venom-gland transcriptome as transcript evidence for gene prediction evidence and the UniProt/SwissProt database (04/2021) as protein evidence. We refined gene and UTR annotations using the Funannotate `update` module, which re-performs genome-guided Trinity v2.5.1 assembly of the RNA-Seq data and PASA assembly. Kallisto (Bray et al. 2016) is used to estimate expression and Funannotate, then uses this information to filter the most likely PASA gene models. Finally, we performed functional annotation with InterProScan5 (Jones et al. 2014), followed by the Funannotate 'annotate' module with annotations pulled from PFAM (Mistry et al. 2021), InterProScan5, EggNog (Huerta-Cepas et al. 2019), UniProtKB, MEROPS (Rawlings et al. 2018), CAZyme (Cantarel et al. 2009), and GO ontology (Ashburner et al. 2000; Carbon et al. 2021).

PLA2 Gene Family Evolution

We estimated the expression of annotated genes using HiSat2 v2.2.1 (Kim et al. 2019) and StringTie v2.2.1 (Pertea et al. 2015). We defined the region between OTUD3 and MUL1 as the putative PLA2 array and manually checked Funannotate annotations by mapping our consensus transcriptome with minimap2 (Li 2018; Li 2021). We manually curated the PLA2 region by identifying exon boundaries following the minimap2 alignments. To identify the genes in the PLA2 region, we extracted and translated the coding DNA sequence (CDS). We aligned the translated CDS regions with sequences from Dowell et al. (2016) using mafft v7.47 (Katoh and Standley 2013), trimmed with trimal v1.4.1 (Capella-Gutiérrez et al. 2009), and then re-

aligned. We then used IQ-TREE v2.2.0 to infer a maximum likelihood phylogeny with the substitution model selected via ModelFinder and 1000 Ultrafast bootstrap replicates to test nodal support (Minh et al. 2020).

We aligned the crotoxin subunits identified in Whittington, Mason and Rokyta (2018) with the *C. lepidus* subunits in Geneious 2020.2.4 (https://www.geneious.com) using a translation alignment with ClustalW (Thompson et al. 1994) to test for the serine mutation at position 127. We used a custom R script to graph the hydrophobicity differences between crotoxin subunits, using hydrophobicity values as designated by Geneious.

High-Performance Liquid Chromatography

We used reversed-phase High-Performance Liquid Chromatography (HPLC) to test for proteomic confirmation of crotoxin. The crotoxin subunits have identifiable peaks, especially when compared to other taxa with a known presence of crotoxin, like *C. tigris* (Margres, Rautsaw, et al. 2021). HPLCs were run on a Shimadzu Prominence HPLC system, according to Margres et al. (2021). 15ug of venom protein was placed into an Aeris 3.6 um C18 Column (Phenomenex, Torrace, CA, USA) using a detection wavelength of 220nm. All samples had a flow rate of 0.2nL min for 125 minutes. HPLCs were compared between *C. lepidus* individuals to see if there were different profiles between CTx+ and CTx- individuals. *C. tigris* and *C. scutulatus* were used for comparisons to check for crotoxin presence, as crotoxin peaks should align between species.

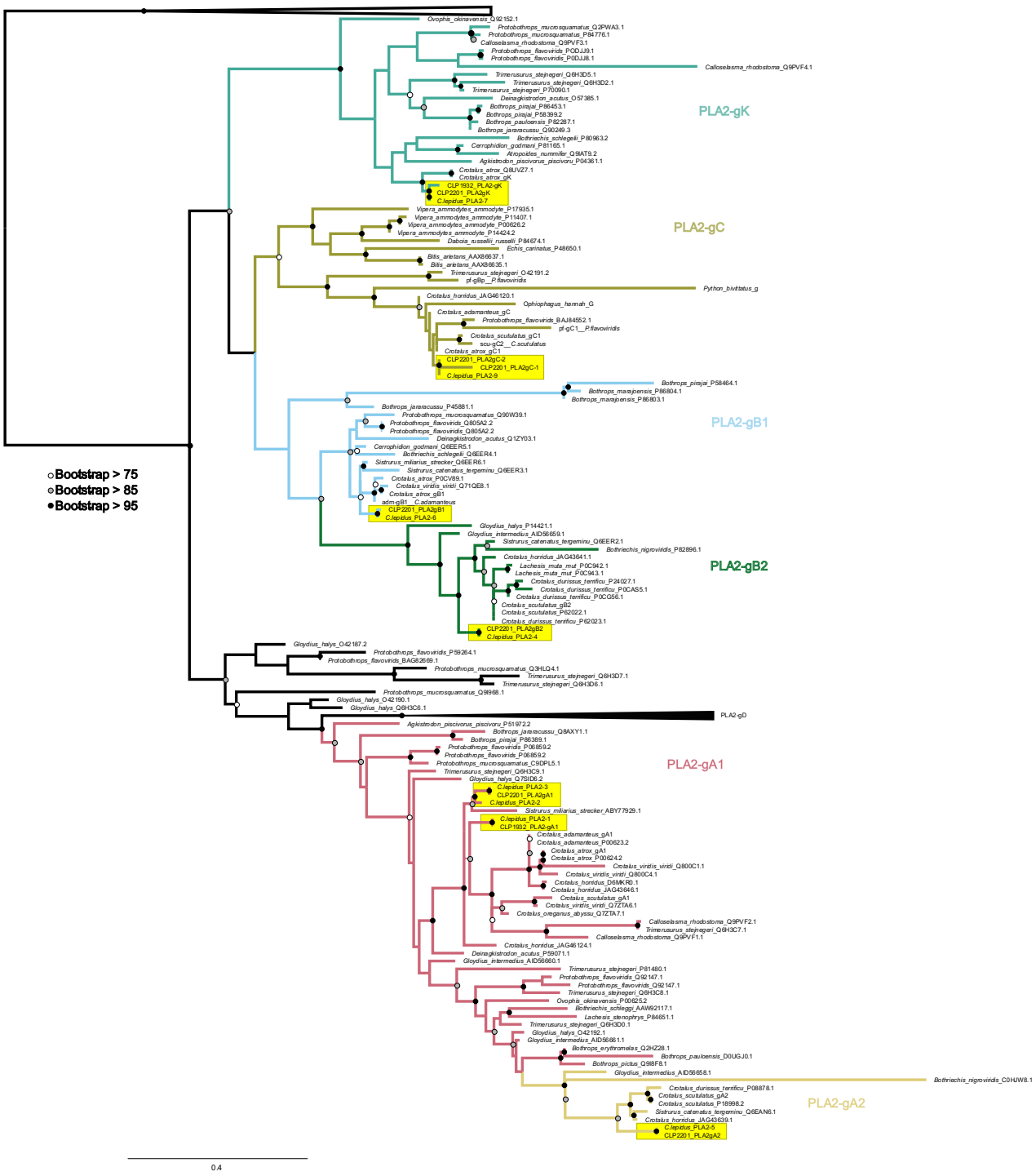

Figure 4. Phylogeny of PLA2-g gene family using sequences found in Dowell et al. (2016) made in IQTree. Light blue represents PLA2-gB1, turquoise represents PLA2-gK, dark green PLA2-gB2, dark yellow PLA2-gC, light yellow PLA2-gA2, and pink PLA2-gA1. Highlighted sequences are *C. lepidus* samples, those of which without a sample ID are from the consensus transcriptome. *C. lepidus* samples with a sample ID are from that sample's genome. We genomically and transcriptomically confirm the presence of both crotoxin subunits (PLA2-gA2 and PLA2-gB2), in addition to three alleles of PLA2-gA1, and one PLA2-gK and PLA2-gB1.

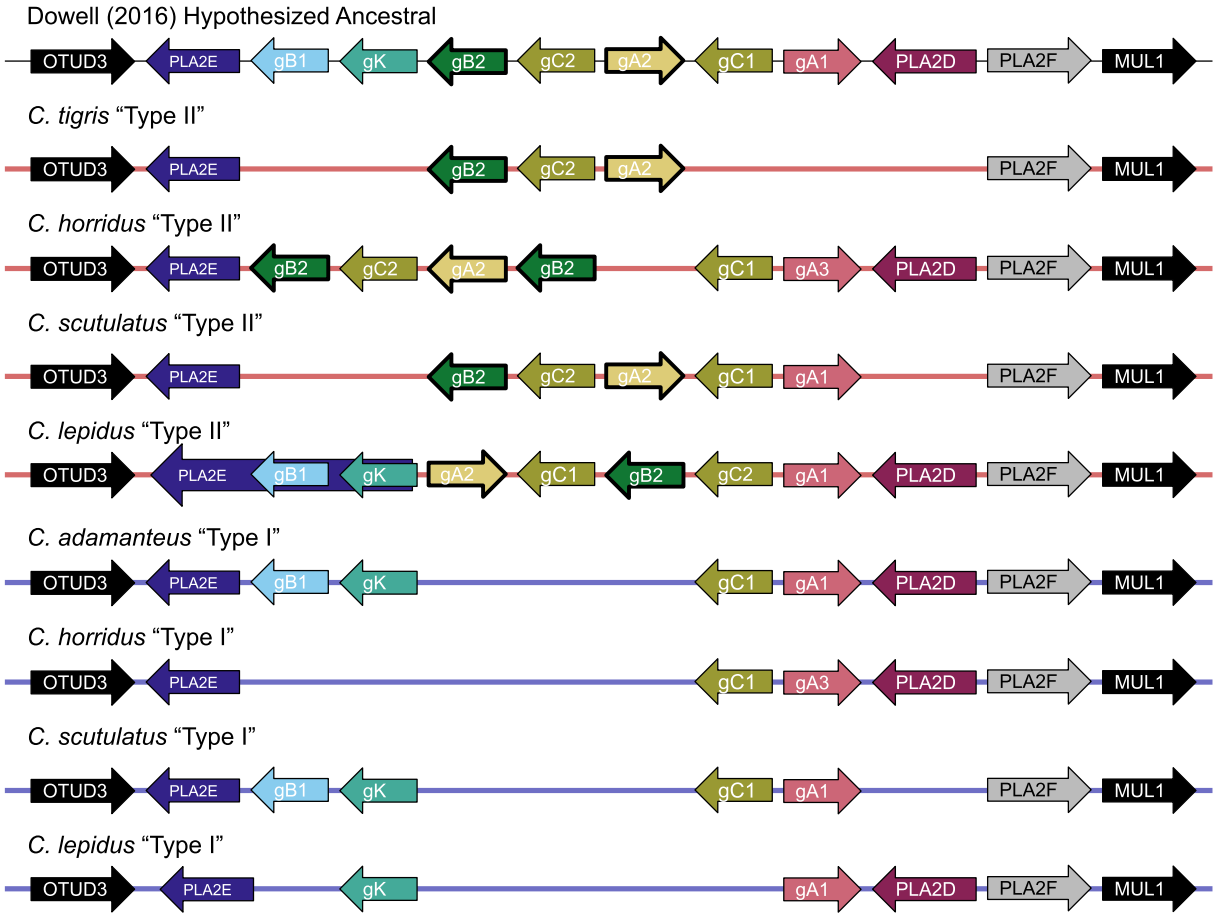

Figure 5. Full PLA2 region arrays comparing regions from Dowell et al. (2016; 2018). Black represents the conserved OTUD3 and MUL1 regions bracketing the PLA2 region. Dark blue represents PLA2-E, light blue represents PLA2-gB1, turquoise represents PLA2-gK, dark green PLA2-gB2, dark yellow PLA2-gC, light yellow PLA2-gA2, pink PLA2-gA1, burgundy PLA2-D, and grey PLA2-F. Red lines represent Type II individuals, blue lines type I. Bolded arrows are the crotoxin subunits. The first array is the Dowell et al. (2016) hypothesized ancestral pitviper. Comparatively, Type II *C. lepidus* individuals are very similar to the hypothesized ancestral pitviper and contain a unique arrangement of the PLA2-E/PLA2-gB1/PLA2-gK. Type I *C. lepidus* individuals differ from other Type I species through loss of the PLA2-gB1 and PLA2-gC.

Results

We assembled two *C. l. klauberi* genomes; one individual from the Huachuca mountains which was CTx- (CLP1932), and one from the Chiricahuas which was CTx+ (CLP2201). For the CTx- individual we sequenced four PacBio Sequel CLR SMRT cells totaling >4.4 M reads and 46.1 Gbp (29x coverage) as well as >345 M 150bp reads on an Illumina NovaSeq S4 flow cell (32x coverage). The resulting genome had a N50 of 2 Mbp and 3,394 contigs (Table 1) The CTx+ genome was generated using 1.5 PacBio Sequel II HiFi SMRT cells totaling >2.3 M reads and 31.9 Gbp (20x coverage) with a final assembly N50 of 15 Mbp from only 1,082 contigs (Table 1).

Both genomes were annotated and the PLA2 region was defined as the section between OTUD3 and MUL1 (Dowell et al. 2016). Minimap2 alignments of the respective transcriptomes facilitated additional manual curation of gene annotations. These CDS regions were then extracted and used to infer a phylogeny of the entire PLA2 gene family, including all data from Dowell et al. (2016), which allowed us to identify and annotate each PLA2 homolog (Figure 4). We confirm the genomic presence of both crotoxin subunits in the *C. lepidus* genome of sample CLP2201 and determine that CTx+ and CTx- individuals have distinct gene composition and gene order (Figure 5). Additionally, the composition and order are different compared to other *Crotalus* species that have variable expression of crotoxin (Figure 5). The PLA2 arrangement in the CTx+ individual contained a translocation of the acidic and basic crotoxin unit, in addition to the retention of the PLA2-gK (Figure 5). The CTx+ individual also had novel gene structure, with the PLA2-gK and PLA2-E sharing their second exon, resulting in an embedded parallel structure (Figure 5). The PLA2-E and PLA2-gB1 had a nested parallel structure, with the PLA2E

surrounding the PLA2-gB1 (Figure 5). The CTx- individual is missing both crotoxin subunits, PLA2-gB1, and a PLA2-gC (Figure 5).

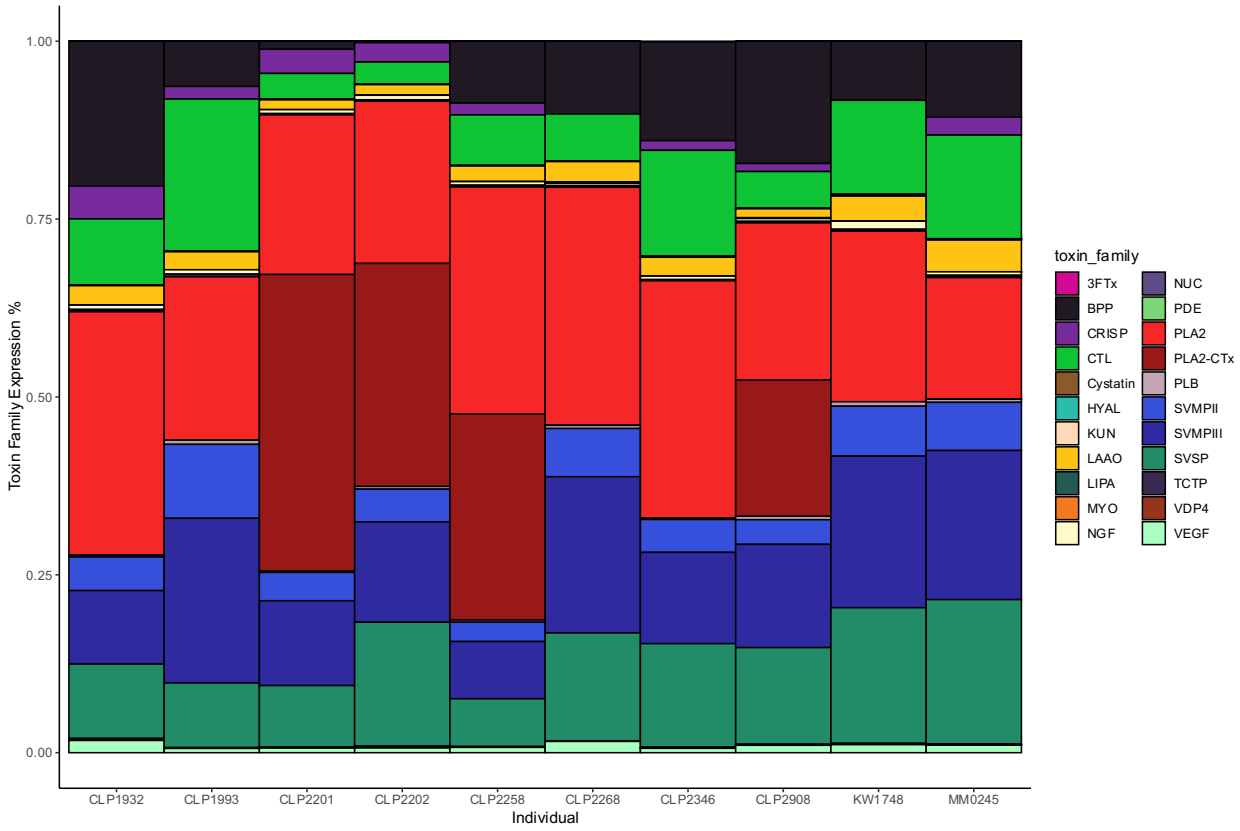

Figure 6. Barplot showing comparative toxin family expression RSEM results across transcriptome individuals. Overall, PLA2s were among the highest if not the highest expressed toxin family. In individuals that expressed crotoxin, crotoxin made up half of the total PLA2 expression. Other highly expressed toxin families included CTLs, SVMPs, and SVSPs.

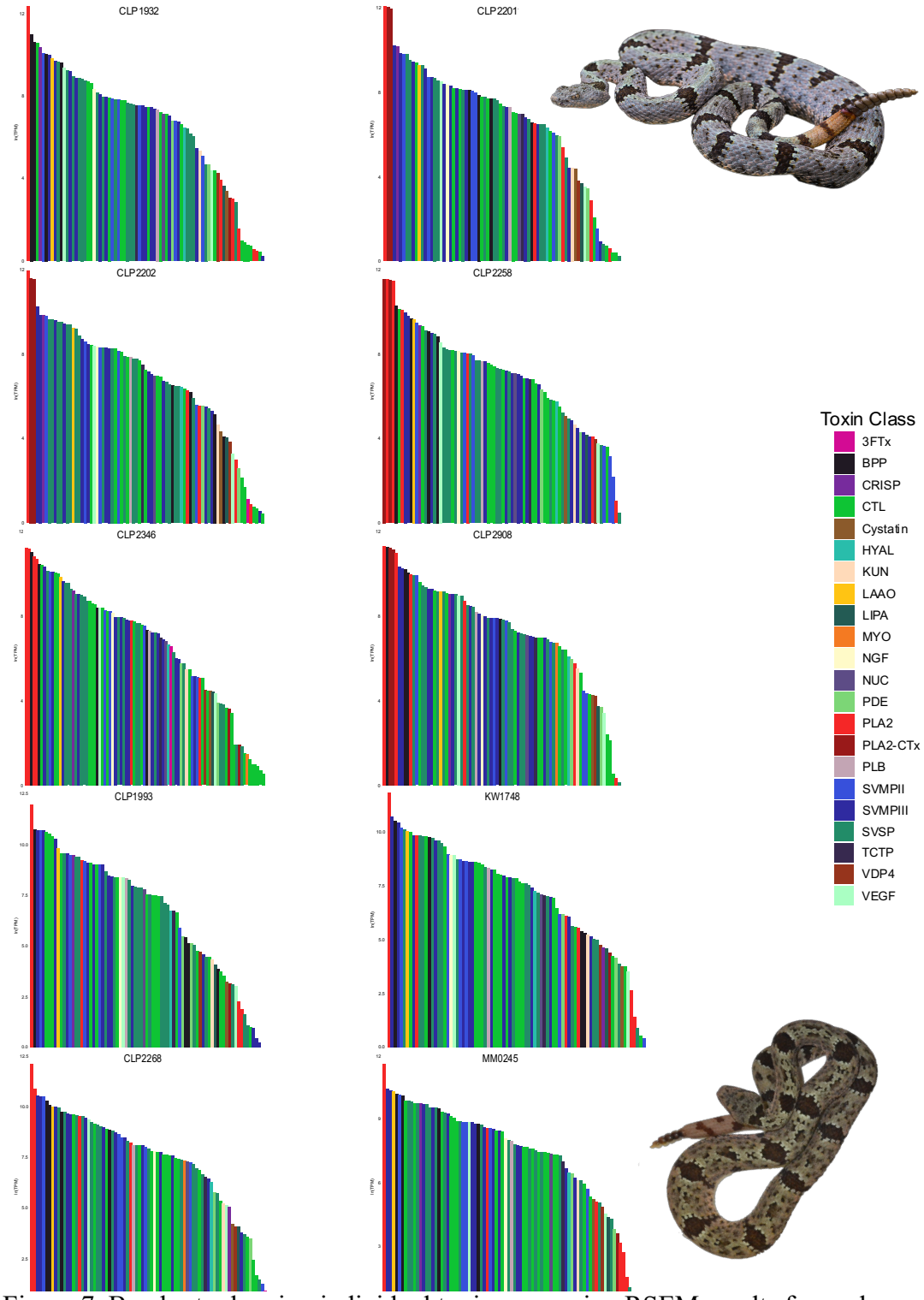

Figure 7. Barcharts showing individual toxin expression RSEM results for each transcriptome sample. First three rows are *C. l. klauberi* (pictured top), bottom two rows are *C. l. lepidus* (pictured bottom). For all samples, PLA2s were the highest single expressed toxin. Samples that expressed crotoxin expressed each subunit in near equal amounts at high rates. Other high, individually expressed toxins include BPPs, CTLs, and SVMPs.

We assembled 10 venom gland transcriptomes from six *C. l. klauberi* and four *C. l. lepidus* found throughout the American Southwest and Mexico (Figures 6 and 7). Only four *C. l. klauberi* individuals expressed crotoxin (CLP2201, CLP2202, CLP2258, CLP2908) while the remaining six samples did not (CLP1932, CLP1993, CLP2268, CLP2346, KW1748, MM0245). PLA2s were the highest expressed toxin family for eight individuals and a PLA2 was the most highly expressed gene in all ten individuals. For *C. l. lepidus*, the highest expressed toxin families were PLA2s, SVMPIIIs, and SVSPs. In addition to these toxin families, *C. l. klauberi* also highly expressed BPPs and CTLs (Figures 6 and 7). Importantly, if crotoxin was found in *C. l. klauberi*, then PLA2 expression made up ~50% of the total venom expression on average, with crotoxin specifically being 50 – 66% of the total PLA2 expression. In individuals without crotoxin, PLA2s were still the highest expressed toxin family, accounting for ~33% of the overall venom expression.

In addition to crotoxin, several other PLA2s were also expressed including three alleles of PLA2-gA1, one PLA2-gK, and one PLA2-gB1. Individuals who were CTx+ had expression of a specific PLA2-gA1 (PLA2-3), another gA1 was associated with CTx- individuals. Differential expression results showed that crotoxin expression was significantly associated with *C. l. klauberi* (PLA2-gA2 p = 3.7e-04; PLA2-gB2 p = 3.5e-07). Expression was not, however, associated with elevation, latitude, or longitude (Table 2). When we tested to see if there was differential expression of crotoxin with environmental variables, only Minimum Temperature of Coldest Month (BIO6) was significant for PLA2-gB2 (p = 0.047). Looking at differential expression of all toxins, several CTLs, SVMPs, and SVSPs were associated with *C. l. klauberi*, and two CTLs were associated with *C. l. lepidus (***S.Table 1)**. Longitude was not a significant axis, and latitude was only significant for SVSP-4.

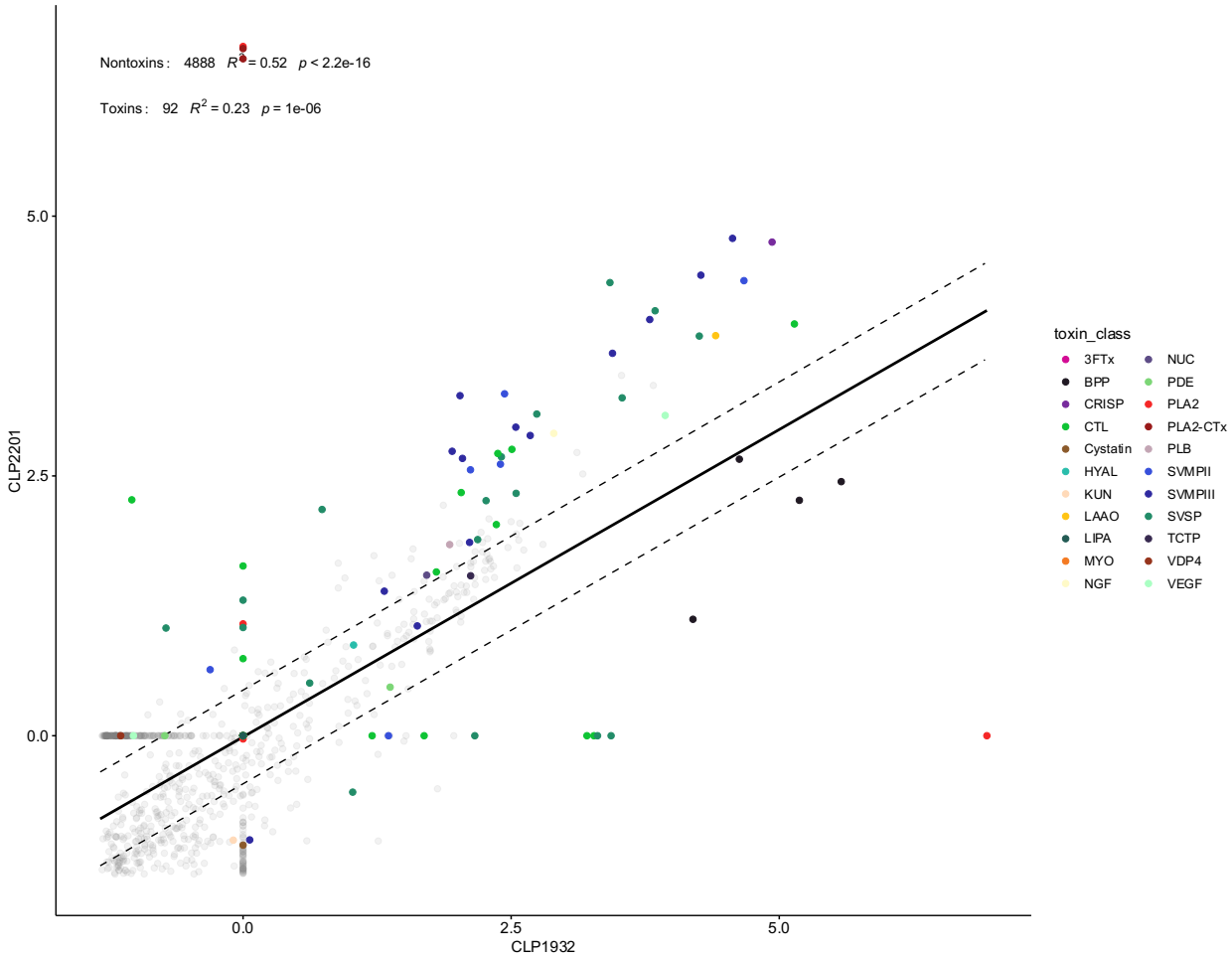

Figure 8. Comparative toxin expression between genome samples CLP1932 (CTx-) and CLP2201 (CTx+). Gray points show nontoxin expression with the solid and dashed lines representing the linear regression and 95% confidence intervals, respectively. The CTx+ individual has higher expression of crotoxin, SVMPs and CTLs. The CTx- individual has higher expression of separate CTLs and separate BPPs.

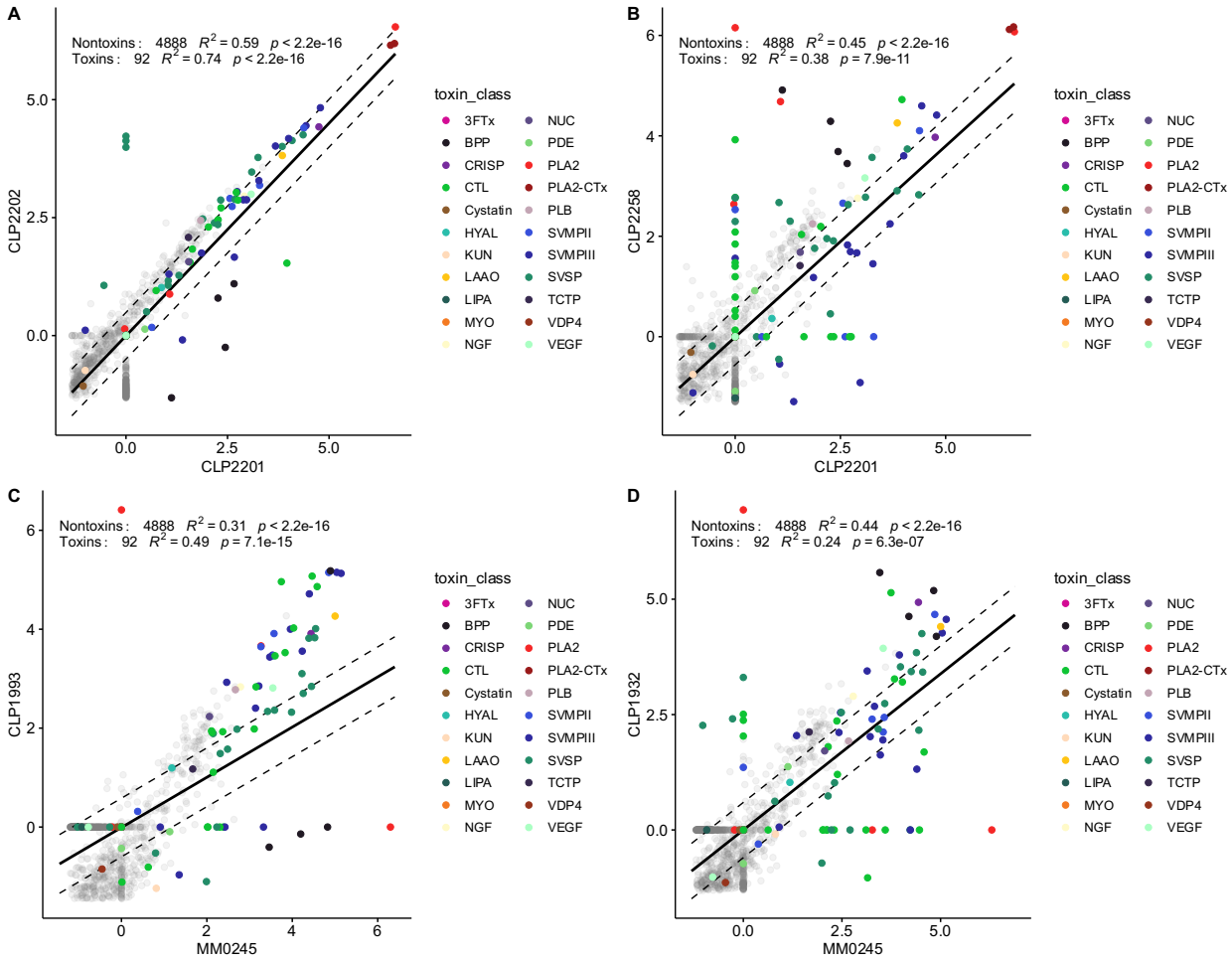

Figure 9. Comparative toxin expression between samples, with dashed lines representing 95% confidence intervals for linear regressions. Gray points show nontoxins. Panel A shows CTx+ samples from the same locality CLP2201 and CLP2202. Panel B shows CTx+ samples from different localities, CLP2201 and CLP2258. Panel C shows CTx- samples from same localities, CLP1993 and MM0245. Panel 4 shows CTx- samples from different localities, CLP1932 and MM0245. Overall samples from the same locality had more similar venom profiles than samples with the same venom type from different localities.

Table 2. Differential Expression results of PLA2-gA2 and PLA2-gB2 along environmental and geographic variables. Values =< 0.05 were considered significant and bolded.

Independent Variable	Expression		Presence	
	CTx Associated with:	pvalue (gA2 / gB2)	Slope:	pvalue
Latitude	-	0.96 / 0.95	-	0.29
Longitude	-	0.72 / 0.23	-	0.13
Subspecies	*C. l. klauberi*	**3.7e-04 / 3.5e-07**	-	0.74
Elevation	-	0.99 / 0.99	-	0.10
Bio_3	-	0.84 / 0.76	0.077	**0.02**
Bio_4	-	0.99 / 0.99	-0.003	**0.05**
Bio_5	-	0.99 / 0.72	-0.189	**0.04**
Bio_6	- / Negative	0.28 / **0.047**	0.020	0.83
Bio_8	-	0.99 / 0.99	-0.228	**0.03**
Bio_10	-	0.99 / 0.99	-0.246	**0.01**
Bio_18	-	0.99 / 0.99	0.011	**0.02**
Bio_19	-	0.70 / 0.29	0.027	**0.00**

Table 3. Toxins that were significantly differentially expressed along the axis of CTx+/-. Toxins in red were more highly expressed in CTx+ individuals, and beige in CTx- individuals. We see that each venom phenotype had an associated PLA2-gA1.

Gene	Expression		log**Δ**DESeq2	padj
	Avg Type I	Avg Type II		
CTL-10	2589	0	-29.04	1.0E-17
CTL-20	0.530	643	10.63	1.3E-02
CTL-23	0.548	265	9.30	8.3E-03
CTL-4	10852	123	-6.26	3.4E-03
CTL-5	9778	0	-30.00	1.9E-21
PLA2-1 (gA1)	103244	0.555	-17.46	5.9E-13
PLA2-3 (gA1)	95.7	121406	10.89	3.3E-12
PLA2-4 (gB2)	8.32	114338	13.92	5.1E-32
PLA2-5 (gA2)	25.3	109293	12.66	8.1E-19

Other toxins were significantly co-expressed with crotoxin (Table 3), for example there was higher expression of two CTLs, and PLA2-3 in individuals with crotoxin, and higher expression of three different CTLs and PLA2-1without crotoxin. PLA2-1 and PLA2-3 are both PLA2-gA1s, implying haplotypes specific to crotoxin positivity**.** The third PLA2-gA1 (PLA-2) was mainly seen in *C. l. klauberi* individuals in Mexico. As a whole, in individuals without crotoxin, there is significantly higher expression of SVMPs ($p = 0.04758$) and CTLs ($p = 0.01048$), and in CTx+ individuals there is significantly higher expression of PLA2s ($p = 0.00483$).

We compared the transcriptomes of both genome specimens (Figure 8**)** and found that both expressed the same number of toxins ($n = 55$). As expected, based off the crotoxin differential expression analyses, the CTx+ individual had more PLA2s expressed, and there were differences in which PLA2s were present in the venom gland transcriptome. The CTx+ individual expressed more SVMPs and CTLs, whereas the CTx- individual expressed more, although different, CTLs. We also performed comparisons for six select individuals (Figure 9). In Figure 9A which features CTx+ samples from the same region, nearly all toxins are expressed at similar rates, with few exceptions. In Figure 9B which features CTx+ samples from different regions, the highest divergence is in CTL and SVMP expression. Additionally, the individual from Mexico expressed more BPPs, and had higher expression of the dual expressed CTx. Looking at the CTx- individuals in Figure 9C (same region) and Figure 9D (different region), most of the venom variation centers on difference in SVMP, CTL, and SVSP expression. In Figure 9C with individuals from the same region, the only toxin that was expressed highly in CLP1993, but not expressed in MM0245 was a PLA2. Compared to the CTx+ individuals from the same area, there was more significant venom variation in the CTx- individuals, with

generally higher expression of the toxin in CLP1993. In Figure 9D, there is a large amount of venom variation, similar with the CTx+ individuals from different regions. Venom variation was largely due to differences in SVSP, CTL, and BPP expression. Overall, venom expression was more similar among individuals from the same area and CTx+ individuals have less variation.

The amino acid alignment of the acidic subunit (PLA2-gA2) shows the key serine mutation at 127 for neurotoxicity. High performance liquid chromatography (HPLC) (Figure 10**)** reveals unique peaks within the CTx+ samples that are absent in the CTx- samples. While peaks do not directly align with *C. tigris* peaks, mutations that change the type of amino acid could change how the venom would elude in a column. Reverse phase HPLCs rely on primarily hydrophobicity, and changes in the hydrophobicity of the amino acid could change that elution time (Hanai and Smith 1999). Van der Waals forces can also play an important role in retention and elution. The *C. lepidus* subunits consist of amino acids that are largely less hydrophobic. Figures 11 and 12 show how the hydrophobicity changes for nonsynonymous mutations in the acidic and basic crotoxin subunit respectively.

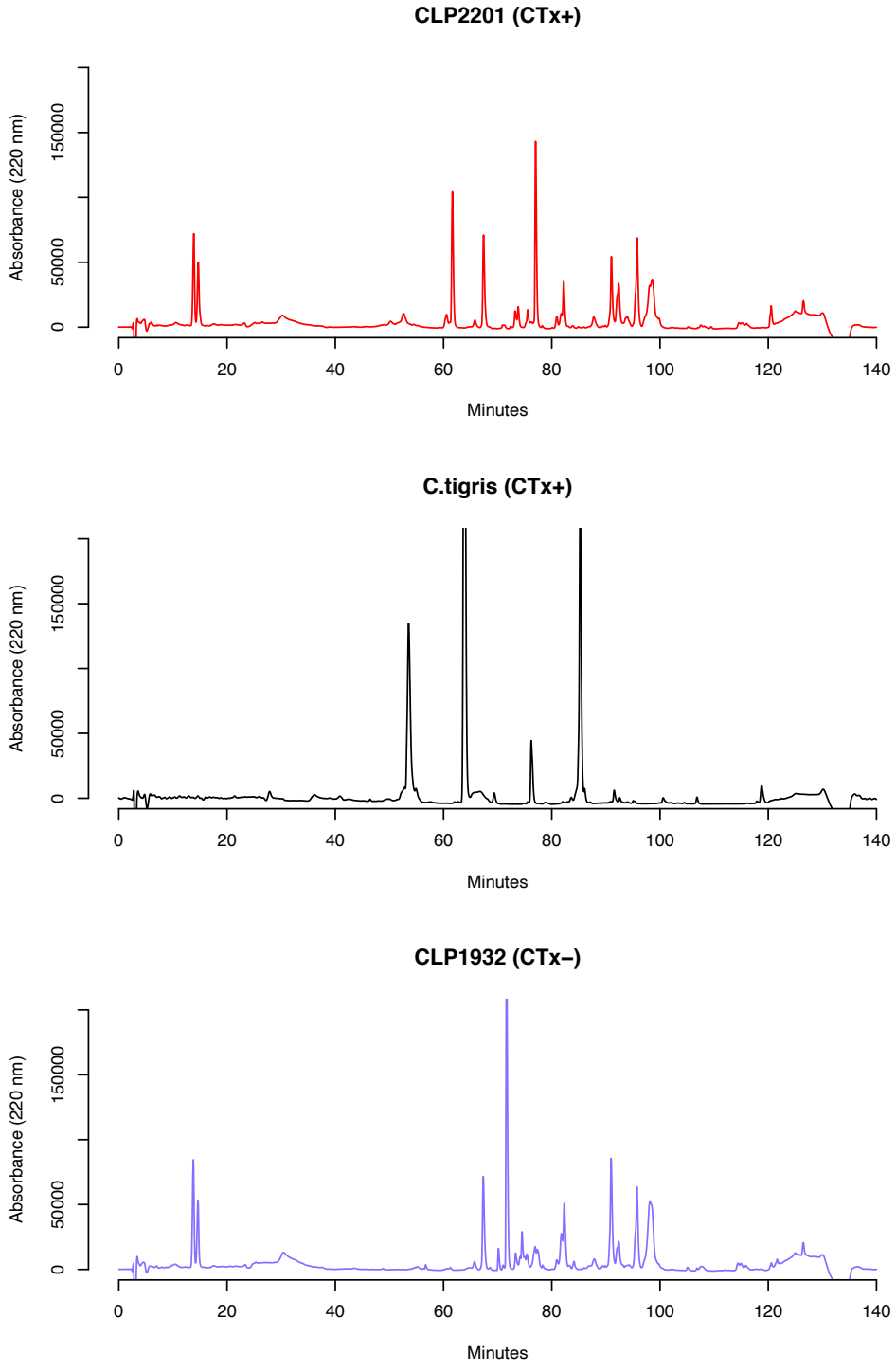

Figure 10. High Performance Liquid Chromatography (HPLC) comparisons of the CTx+ *C. lepidus, C. tigris* (CTx+ from Margres et. al (2021)), and CTx- *C. lepidus*. There are two peaks in the CTx+ *C. lepidus*, just have 60 minutes and before 80 minutes, that do not aign with the CTx- *C. lepidus* sample. Due to the amino acid hydrophobicity changes, they could be the crotoxin subunits, even if they do not align with *C. tigris* peaks.

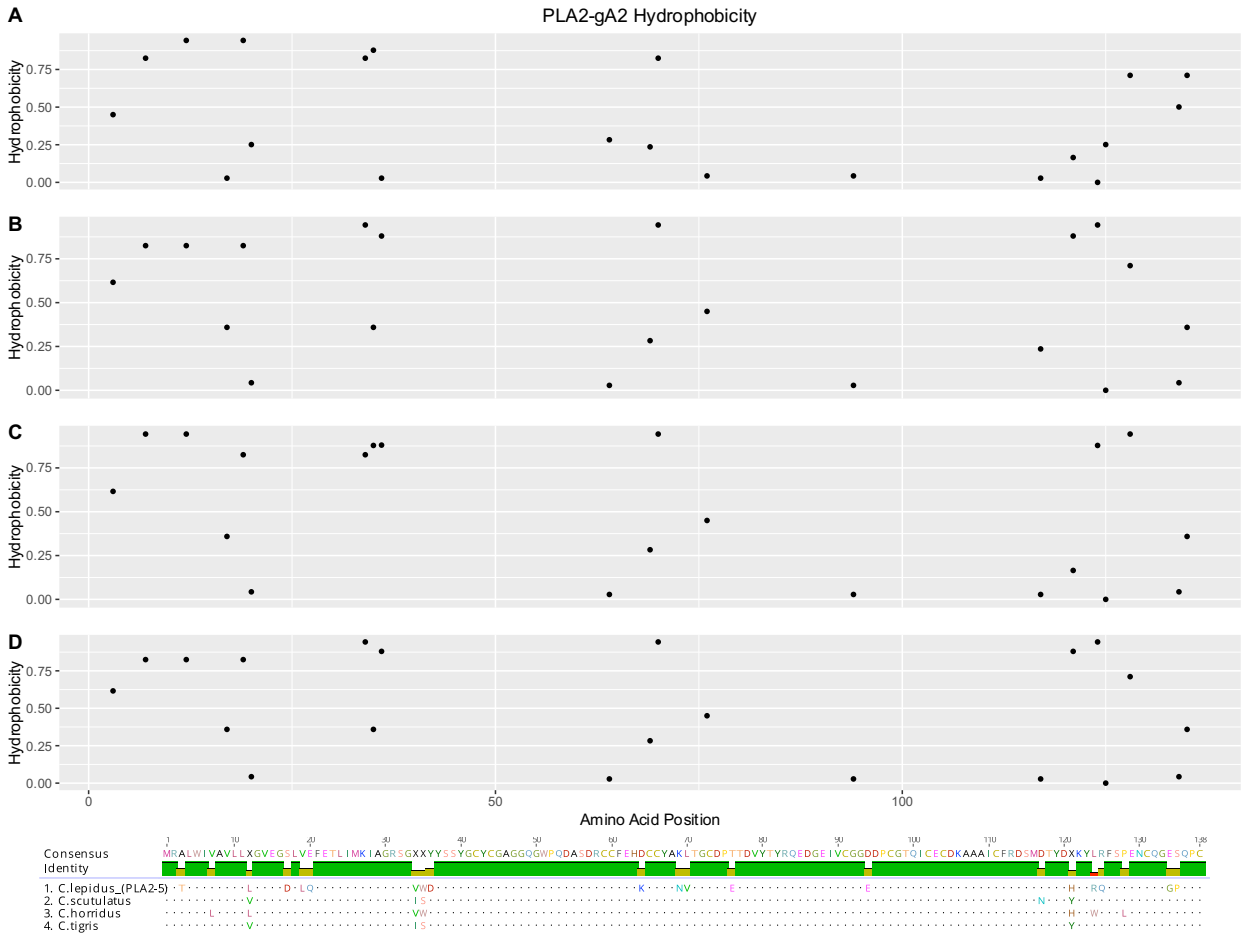

Figure 11. The hydrophobicity of PLA2-gA2 alignments in (A) *C. lepidus,* (B) *C. scutulatus,* (C) *C. horridus,* and (D) *C. tigris*. Top shows hydrophobicity values at nonsynonymous sites, with a score of 1 meaning very hydrophobic, and a score of 0 meaning hydrophilic. Underneath is the nonsynonymous amino acid residues in alignment against the consensus. Overall, the *C. lepidus* sequence is less hydrophobic than other identified crotoxin acidic subunits.

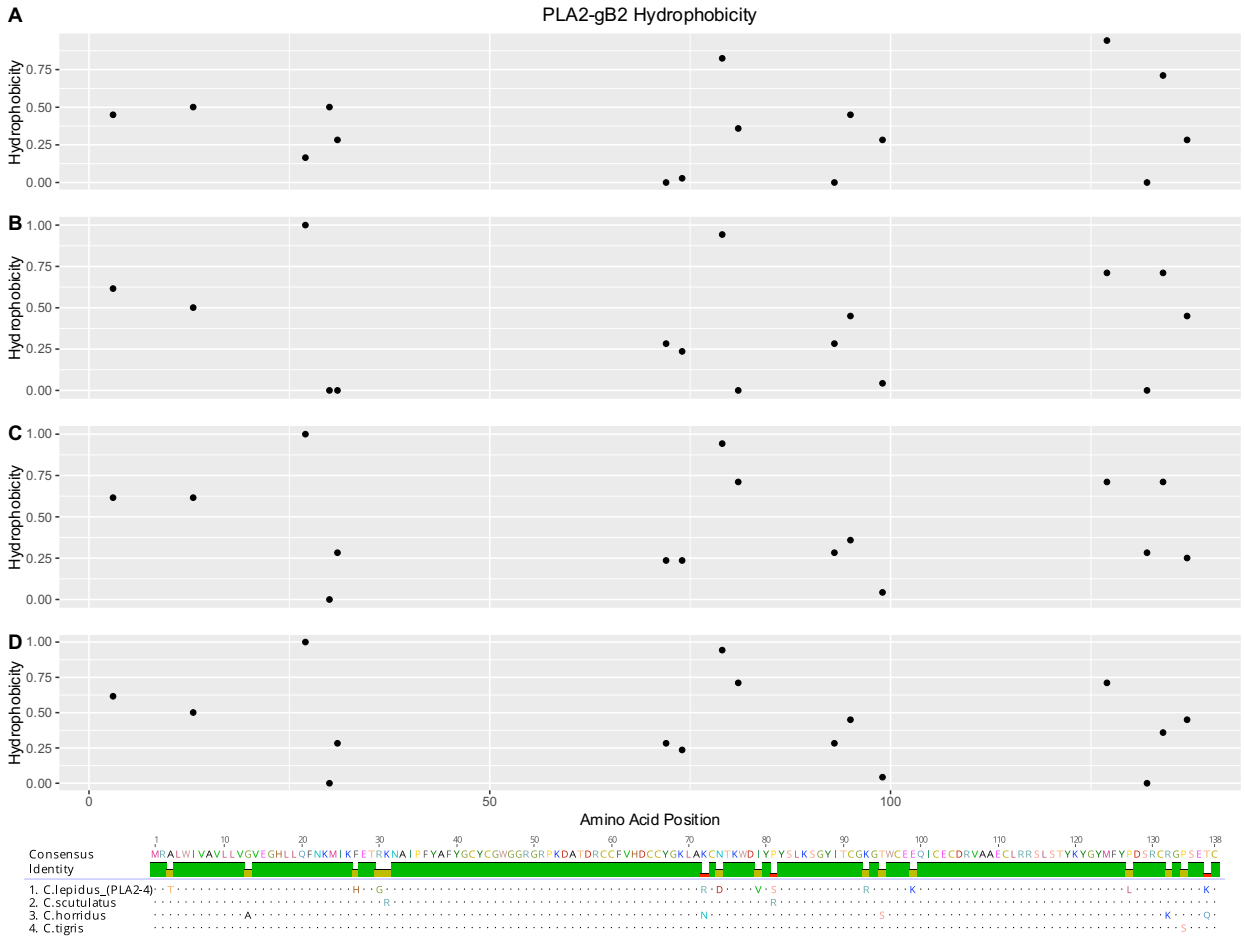

Figure 12. The hydrophobicity of PLA2-gB2 alignments in (A) *C. lepidus,* (B) *C. scutulatus,* (C) *C. horridus,* and (D) *C. tigris*. Top shows hydrophobicity values at nonsynonymous sites, with a score of 1 meaning very hydrophobic, and a score of 0 meaning hydrophilic. Underneath is the nonsynonymous amino acid residues in alignment against the consensus. Overall, the *C. lepidus* sequence is less hydrophobic in the first 50 residues, and slightly more hydrophobic in the remaining residues.

Discussion

Using genome sequencing and transcriptomics we find that crotoxin is present and expressed in several *C. lepidus* individuals. This is the first genomic and transcriptomic confirmation of crotoxin in a montane rattlesnake. We also recover unique PLA2 arrays for *C. lepidus* further illustrating the evolutionary history of this toxin family in the species. This expands the range of the trait geographically and taxonomically. This confirmation opens further exploration of the types of environments crotoxin could be a local adaptation to, and the opportunity to test for selection across different taxa and biomes. The phylogenetic expansion allows further opportunities to test existing hypotheses of crotoxin origin, like the ancestral pitviper hypothesis.

Our main goal was to identify which PLA2s were present in *C. lepidus*, both in the genome and transcriptome. We accomplished this using genomic and transcriptomic sequencing, and phylogenetic inference using previously identified PLA2 sequences. We detected both PLA2 subunits (PLA2-gA2 and PLA2-gB2) necessary to form crotoxin, placed basal to previously characterized *Crotalus* crotoxin subunits in the PLA2 phylogeny.

Surprisingly, the subunits of crotoxin are expressed in near equal proportion, instead of in a 2:1 basic to acidic ratio (Table 3). The 2:1 ratio was initially hypothesized to exist to ensure there was enough PLA2-gB2 chaperone for the active PLA2-gA2 to bind to, maximizing neurotoxic effects (Whittington et al. 2018). While our study does not go into the neurotoxic effects of *C. lepidus* venom, previous studies have shown low LD50 values in individuals that were reactive to a Mojave Toxin antibody (average 0.395 ug/g (Rael et al. 1992)). Individuals that were not reactive to the Mojave Toxin antibody had an average LD50 of 9.50 ug/g (Rael et al. 1992). In comparison, *C. scutulatus* CTx+ samples have LD50s ranging from 0.34 mg/kg to

2.6 mg/kg (Massey et al. 2012). CTx- samples had LD50s greater than 2.9 mg/kg. *C. lepidus* samples from Rael *et al*. (1992) had a lethality comparable to *C. scutulatus*, which suggests no loss of efficiency. Like in Strickland, Mason, *et al.*, (2018) we also identified CTLs significantly differentially expressed between CTx+ and CTx- individuals. Venom evolves as a response to local adaptation (Gibbs and Mackessy 2009; Holding et al. 2016). There are similarities between CTx+ *C. lepidus* and *C. scutulatus* venoms (mainly PLA2 expression), and similarities between CTx- *C. lepidus* and *C. scutulatus* venoms (mainly SVMP and CTL expression) that could suggest a similar selection pressure acting on both. Further research into both the ecology and selection pressures of the two species, and the evolutionary histories of each toxin family would be needed to find any overlap in circumstances.

The PLA2-gE in the CTx+ genome offers a novel opportunity to examine newly discovered parallel nested gene structure. We are confident that the gene is present in the genome and is not a chimeric effect of the assembly or annotation process. The gene was expressed in the transcriptome at an average TPM of 653.377. When examining the alignment between different species, the first two CDS regions, accounting for 39% of the gene, are more similar to PLA2-gKs. The last two CDS regions are conserved throughout several species. It was surprising that the PLA2-gK was retained in the CTx+ individual as well and would mean that a myotoxic PLA2 may also be expressed in addition to crotoxin. The retention could be related due to divergent evolutionary histories when large and small-bodied rattlesnakes diverged, or due to the unique genetic structure of the CTx+ array. We propose two possible hypotheses regarding how the PLA2-E/PLA2-gB1/PLA2-gK translocations could have evolved: transposable element rearrangement or recombination. There are several transposable elements between the PLA2-E and PLA2-gB1 in other species (Dowell et al. 2016). There could have been a rearrangement,

taking the region right before the start of the PLA2-E and transferring it to the other side of the PLA2-gB1. Non-allelic homologous recombination (NAHR) between misaligned regions could have also contributed to the arrangement seen in *C. lepidus*. NAHRs can cause structural variations such as inversions and translocations through the process of repairing double stranded breaks using the homologous region of a sister chromatid (Robberecht et al. 2013; Parks et al. 2015). The repetitive tandem duplications of the PLA2 region could, therefore, induce errors during recombination, allowing for novel rearrangements.

For both hypotheses, additional sequencing of more individuals would help confirm the CTx+ and CTx- PLA2 genomic arrays. Specifically, the presence of PLA2-gK throughout CTx+ individuals, and the position of PLA2-gB1 in CTx- individuals, since our CTx- genome sample was unique in its' absence of that gene. The retrotransposons identified in other *Crotalus* species (Dowell et al. 2016), especially those between the PLA2-E and PLA2-gB1 could play a role in disrupting genes, immediately making them nonfunctional. Disruption of regulatory elements of toxin genes could put the gene under relaxed selection pressures (Feschotte 2008). Transposable elements have been hypothesized to have caused co-opted regulation of venom genes (Perry et al. 2022), so disruption of gene regulatory elements is a possibility. To further investigate the potential role of NAHR, a program like detect-NAHR (Parks et al. 2015) could be used to identify signals of NAHR in all variable expression species. This would allow for genus-wide comparisons of variational points, which are indicative of NAHR. NAHRs are a particularly attractive candidate as they allow for gene conversions, which could explain the embedded parallel structure of PLA2 genes in *C. lepidus* in addition to deletions and duplications (Parks et al. 2015). NAHR has resulted in deletions and hybrid genes in inbred soybean lines (Cho et al. 2019). In human IgG receptors, NAHR is responsible for a high copy number variation, and a

rare chimeric gene variant (Nagelkerke et al. 2015). To see the role of NAHR in viperids would be a difficult feat, as delineating breakpoints would likely require inbred lines or a large sample size, however if there is consistency across species with the same gene arrays in intronic SNPs, that could be a region potentially containing a breakpoint.

CHAPTER THREE

SPATIAL DISTRIBUTION AND ORIGINS OF CROTOXIN IN *C. LEPIDUS*

Introduction

Within a species, crotoxin expression can be variable among populations. *Crotalus scutulatus* is a prominent example of how a single species can have variable expressions of a specific venom phenotype. Although *C. scutulatus* is thought to be a Type II species, there are Type I and hybrid individuals across its range (Glenn and Straight 1989; Strickland, Mason, et al. 2018; Strickland, Smith, et al. 2018; Zancolli et al. 2019). Strickland, Smith, et al. (2018) hypothesized that local selection favored one venom phenotype over another in specific sub-populations, resulting in a species wide variable expression phenotype. Crotoxin expressing (CTx+) and crotoxin absent (CTx-) individuals occupied different niches, specifically along axes of temperature and rainfall. These environmental variables could be correlated with prey types or other circumstances that could facilitate directional selection. *Crotalus horridus* is another species with variable expression of crotoxin and CTx+ individuals are more commonly found along the western and southern parts their range (Glenn et al. 1994; Rokyta et al. 2013; Margres, Wray, et al. 2021). While local selection drivers have not been studied in relation to crotoxin's expression in *C. horridus*, it is known to be present in only certain parts of the distribution. Similar phenomena could be occurring in *C. lepidus*. Spatial analyses will identify the distribution of this trait across the species' range. If crotoxin is localized or distributed throughout, this would have implications on the origin of crotoxin in the species in addition to the severity of snakebites in the area. To understand the distribution of crotoxin in *C. lepidus* we will combine our transcriptome dataset (Q1) with a large hybrid enrichment or sequence capture dataset. This will increase the breadth of my data set, and we can test specifically for genomic

presence of crotoxin across their range. With both the sequence capture and transcriptome data, we will perform logistic regressions along latitude, longitude, elevation, and environment to test for correlations with presence/absence of crotoxin.

It is hypothesized that crotoxin is an ancestral trait in pitviper venoms (Dowell et al. 2016; Whittington et al. 2018). However, it is rarely documented in montane rattlesnake species (Rael et al. 1992; Martínez-Romero et al. 2013; Saviola et al. 2017), and our previous chapter was the first to establish the genomic presence of crotoxin. Through a combination of the PLA2 phylogeny, spatial distributions, and a Patterson's D statistic, we tested to see if crotoxin is a result of i) independent origin, ii) hybridization with crotoxin-harboring congeners *C. scutulatus* or *C. tigris,* or iii) ancestral origin.

Materials and Methods

Sampling

Collaborators Mark Margres, Kenny Wray, Darin Rokyta, and members of the Parkinson lab collected ninety-six *C. lepidus* samples throughout Texas (Permit #SPR-0713-098), New Mexico (Permit #3571), Arizona (Permit #SP677954; #SP706039; #SP735619), and Mexico (OFICIO NUM. SGPA/DGVS/03562/15) following the guidance set by Florida State University's IACUC (protocols #0924 and #1333) and University of Central Florida's IACUC (#13-17W; #16-17W) to generate the sequence capture data. Before individuals were released, they collected blood/tissue (Table 4). One sample, KW1748, was also collected for transcriptome processing.

Table 2. Sequence Capture samples with locality information and venom type.

Sample ID	Subspecies	Venom Type	County	State
KW1748	*Crotalus lepidus lepidus*	CTx-	Pecos	Texas
KW2059	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas
KW2073	*Crotalus lepidus klauberi*	CTx+	Hidalgo	New Mexico
KW2074	*Crotalus lepidus klauberi*	CTx+	Hidalgo	New Mexico
KW2075	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2076	*Crotalus lepidus klauberi*	CTx+	Hidalgo	New Mexico
KW2077	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2078	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2079	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2080	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas
KW2081	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas
KW2083	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2084	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2085	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2086	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2087	*Crotalus lepidus klauberi*	CTx+	Hidalgo	New Mexico
KW2088	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2089	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2090	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2092	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2093	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2094	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2095	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2096	*Crotalus lepidus klauberi*	CTx-	El Paso	Texas
KW2097	*Crotalus lepidus klauberi*	CTx-	El Paso	Texas
KW2098	*Crotalus lepidus klauberi*	CTx-	Hudspeth	Texas
KW2099	*Crotalus lepidus*	CTx-	*not included in enviornmental analyses	
KW2100	*Crotalus lepidus klauberi*	CTx-	Hudspeth	Texas
KW2101	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2102	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2103	*Crotalus lepidus lepidus*	CTx-	Terrell	Texas
KW2104	*Crotalus lepidus lepidus*	CTx-	Terrell	Texas
KW2105	*Crotalus lepidus klauberi*	CTx-	Dona Ana	New Mexico
KW2106	*Crotalus lepidus klauberi*	CTx-	Dona Ana	New Mexico
KW2107	*Crotalus lepidus klauberi*	CTx-	Pinos	Zacatecas
KW2108	*Crotalus lepidus klauberi*	CTx-	Pinos	Zacatecas
KW2109	*Crotalus lepidus klauberi*	CTx-	Pinos	Zacatecas
KW2110	*Crotalus lepidus klauberi*	CTx-	Mezquital	Durango
KW2111	*Crotalus lepidus klauberi*	CTx-	Mezquitic	Jalisco
KW2112	*Crotalus lepidus klauberi*	CTx-	Ahualulco de Mercado	Jalisco
KW2113	*Crotalus lepidus klauberi*	CTx-	Ahualulco de Mercado	Jalisco
KW2114	*Crotalus lepidus lepidus*	CTx-	Aramberri	Nuevo Leon
KW2115	*Crotalus lepidus lepidus*	CTx+	Catorce	San Luis Potosi
KW2116	*Crotalus lepidus lepidus*	CTx+	Catorce	San Luis Potosi
KW2117	*Crotalus lepidus lepidus*	CTx-	Catorce	San Luis Potosi
KW2118	*Crotalus lepidus lepidus*	CTx+	Catorce	San Luis Potosi
KW2119	*Crotalus lepidus lepidus*	CTx-	Catorce	San Luis Potosi
KW2120	*Crotalus lepidus lepidus*	CTx-	Catorce	San Luis Potosi
KW2121	*Crotalus lepidus lepidus*	CTx+	Catorce	San Luis Potosi
KW2122	*Crotalus lepidus lepidus*	CTx-	Edwards	Texas
KW2123	*Crotalus lepidus lepidus*	CTx-	Edwards	Texas
KW2124	*Crotalus lepidus lepidus*	CTx-	Presidio	Texas
KW2125	*Crotalus lepidus lepidus*	CTx-	Pecos	Texas
KW2126	*Crotalus lepidus lepidus*	CTx-	Pecos	Texas
KW2127	*Crotalus lepidus lepidus*	CTx-	Hudspeth	Texas
KW2128	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2129	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2130	*Crotalus lepidus klauberi*	CTx-	Luna	New Mexico
KW2131	*Crotalus lepidus klauberi*	CTx-	Dona Ana	New Mexico
KW2132	*Crotalus lepidus klauberi*	CTx-	Dona Ana	New Mexico
KW2133	*Crotalus lepidus klauberi*	CTx-	Grant	New Mexico
KW2135	*Crotalus lepidus klauberi*	CTx-	Grant	New Mexico
KW2137	*Crotalus lepidus klauberi*	CTx-	Socorro	New Mexico
KW2138	*Crotalus lepidus lepidus*	CTx-	Hudspeth	Texas
KW2139	*Crotalus lepidus klauberi*	CTx-	Sierra	New Mexico
KW2140	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas
KW2142	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2143	*Crotalus lepidus klauberi*	CTx-	Dona Ana	New Mexico
KW2144	*Crotalus lepidus klauberi*	CTx-	Dona Ana	New Mexico
KW2145	*Crotalus lepidus klauberi*	CTx-	Dona Ana	New Mexico
KW2146	*Crotalus lepidus klauberi*	CTx-	Luna	New Mexico
KW2147	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2151	*Crotalus lepidus klauberi*	CTx-	Grant	New Mexico
KW2152	*Crotalus lepidus klauberi*	CTx-	Sierra	New Mexico
KW2154	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas
KW2198	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas
KW2200	*Crotalus lepidus klauberi*	CTx-	Grant	New Mexico
KW2216	*Crotalus lepidus lepidus*	CTx-	Presidio	Texas
KW2217	*Crotalus lepidus lepidus*	CTx-	Presidio	Texas
KW2218	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2219	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2220	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2221	*Crotalus lepidus klauberi*	CTx+	Hidalgo	New Mexico
KW2222	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2223	*Crotalus lepidus lepidus*	CTx-	Val Verde	Texas
KW2224	*Crotalus lepidus klauberi*	CTx-	Hidalgo	New Mexico
KW2226	*Crotalus lepidus klauberi*	CTx-	Gomez Palacio	Durango
KW2227	*Crotalus lepidus klauberi*	CTx-	Viesca	Coahuila
KW2229	*Crotalus lepidus klauberi*	CTx-	Matamoros	Coahuila
KW2230	*Crotalus lepidus klauberi*	CTx-	Cochise	Arizona
KW2232	*Crotalus lepidus lepidus*	CTx-	Jeff Davis	Texas
KW2235	*Crotalus lepidus klauberi*	CTx+	Cochise	Arizona
KW2236	*Crotalus lepidus klauberi*	CTx-	Cochise	Arizona
KW2237	*Crotalus lepidus klauberi*	CTx+	Cochise	Arizona
MM0104	*Crotalus lepidus klauberi*	CTx+	Cochise	Arizona
MM0105	*Crotalus lepidus klauberi*	CTx+	Cochise	Arizona

Probe Design and Sequence Capture

Sequence capture probes and hybrid enrichment data used herein were generated by Mark Margres and Darin Rokyta (FSU) and generously shared with us for this work. This probe set design and methods were described in Margres, Bigelow, et al. (2017). In short, they designed 120 bp probes for toxins (94 long exons, 112 short exons; see below), nontoxins (n =200), anchored loci (n = 348) and anonymous (short n = 829; long n = 240) loci, which included probes for the main toxin families. The probes were designed using sequences from venom-gland transcriptomes and low coverage genomes of other pitviper species (*C. adamanteus, Agkistrodon piscivorus, Sisturus miliarius, C. horridus* (one with Type I venom, one with Type II)). This probe set specifically included probes that hybridized to PLA2 gene regions and were used to capture the PLA2 genes from the 96 *C. lepidus* samples. Although toxins (including PLA2s) evolve rapidly (Casewell et al. 2013; Dowell et al. 2016; Holding et al. 2016), previous studies have shown that the flanking regions around PLA2s in the genome consist of conserved domains facilitating the capture of these loci (Dowell et al. 2016; Margres, Bigelow, et al. 2017). These conserved domains are what the sequence capture probes were based on. The Center for Anchored Phylogenomics at Florida State University performed the target capture (Lemmon et al. 2012; Prum et al. 2015; Margres, Wray, et al. 2017; Margres, Bigelow, et al. 2017).

Briefly, The Center for Anchored Phylogenomics at FSU used an Omega Bio-tek E.Z.N.A Tissue DNA Kit (Omega Bio-Tek Inc, Norcross, GA, USA) to extract DNA from blood or tissue samples. The Center for Anchored Phylogenomics at FSU used a Covaris E220 Focused-ultrasonicator with Covaris microTUBES (Covaris, Woburn, MA, USA) to sonicate each DNA sample to the desired length of 175-325bp. The Center for Anchored Phylogenomics at FSU used a Beckman-Coulter Biomek FXp liquid handling robot to perform library

preparation and indexing with SPRIselect beads (0.9x ratio of bead to sample volume; Beckman-Coulter, Brea, CA, USA) for a size-selecting step after the repairing of blunt-ends (Margres, Bigelow, et al. 2017). An Agilent Custom SureSelect Kit (Agilent Technologies, Santa Clara, CA, USA), utilizing the 57,292 probes, was used to isolate and enrich regions of interest. These libraries were pooled in equal amounts. Florida State University College of Medicine Translational Science Laboratory sequenced the libraries with 150bp paired-end reads on an Illumina HiSeq 2500. Raw sequencing reads were provided to us for the *C. lepidus* samples and we used Trim Galore v0.6.6 to trim reads and PEAR v0.9.10 to merge paired-end reads (https://github.com/FelixKrueger/TrimGalore) (Zhang et al. 2014). We used BWA-MEM and GATK following the Best Practices For Variant Calling With The GATK (see https://github.com/njm-data/GatkSNPS) using HaplotypeCaller with the PLA2s from the *C. lepidus* consensus transcriptome (generated in Chapter Two) as the reference (Li and Durbin 2009; McKenna et al. 2010). Genotyping was done using GenotypeGVCF, variants were filtered out with a filter of "QD < 2.0 || FS > 60.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0" (https://github.com/njm-data/GatkSNPS). If a sequence capture sample had alignment matches to both PLA2-gA2 and PLA2-gB2 (the genes that code for both subunits of CTx), it was considered positive for the presence of crotoxin.

Spatial Analyses

We examined the spatial relationship of crotoxin presence to test for a relationship between crotoxin presence and latitude or longitude to see if there was directionality in the trait and as a proxy for subspecies. Strickland, Smith, *et al.* (2018) previously hypothesized that isolated crotoxin populations arise due to local adaptation. To see if local adaptation played a role, we tested to see if crotoxin presence is associated with a specific elevation or environmental

variables and compared them to the environments of other variable expression species (Fick and Hijmans 2017). We removed one sequence capture sample that did not have location data. We used the R packages "rgdal" (Bivand et al. 2021), "raster" (Hijmans 2021), and "tidyjson" (Stanley and Arendt 2020) to extract and organize environmental raster data for each *C. lepidus* sample, and used the "stats" packages to run all binary logistic regressions (R CoreTeam 2021). All figures were plotted in ggplot2 (Wickham 2016).

Patterson's D

We tested for introgression between *C. lepidus* and *C. scutulatus,* and *C. tigris* to assess the possibility of crotoxin being present in *C. lepidus* via introgression. Patterson's D, an introgression metric, was calculated in R using admixr (Petr et al. 2019) using transcriptomes of *C. lepidus*, *C. scutulatus, C. tigris*, and *Sistrurus* **(S.Table 2)** (Green et al. 2010; Strickland, Smith, et al. 2018; Strickland, Mason, et al. 2018; Holding et al. 2021; Margres, Rautsaw, et al. 2021). Specifically, we aligned RNA-Seq data from all species to nontoxins found in the *C. lepidus* transcriptome as a reference using BWA-MEM. We then called SNPs using the previously described GATK pipeline. We used eigensoft to convert the SNP vcf file to the eigenstrat format (Patterson et al. 2006; Price et al. 2006). The eigenstrat formatted SNPs were used as the input for admixr, with three separate tests being performed. Tests one and two examined whether CTx+ or CTx- *C. scutulatus* samples were hybridizing with *C. lepidus* samples. Test three examined if *C. tigris* was hybridizing in place of *C. scutulatus*. In all three tests, *Sisturus* acted as the outgroup.

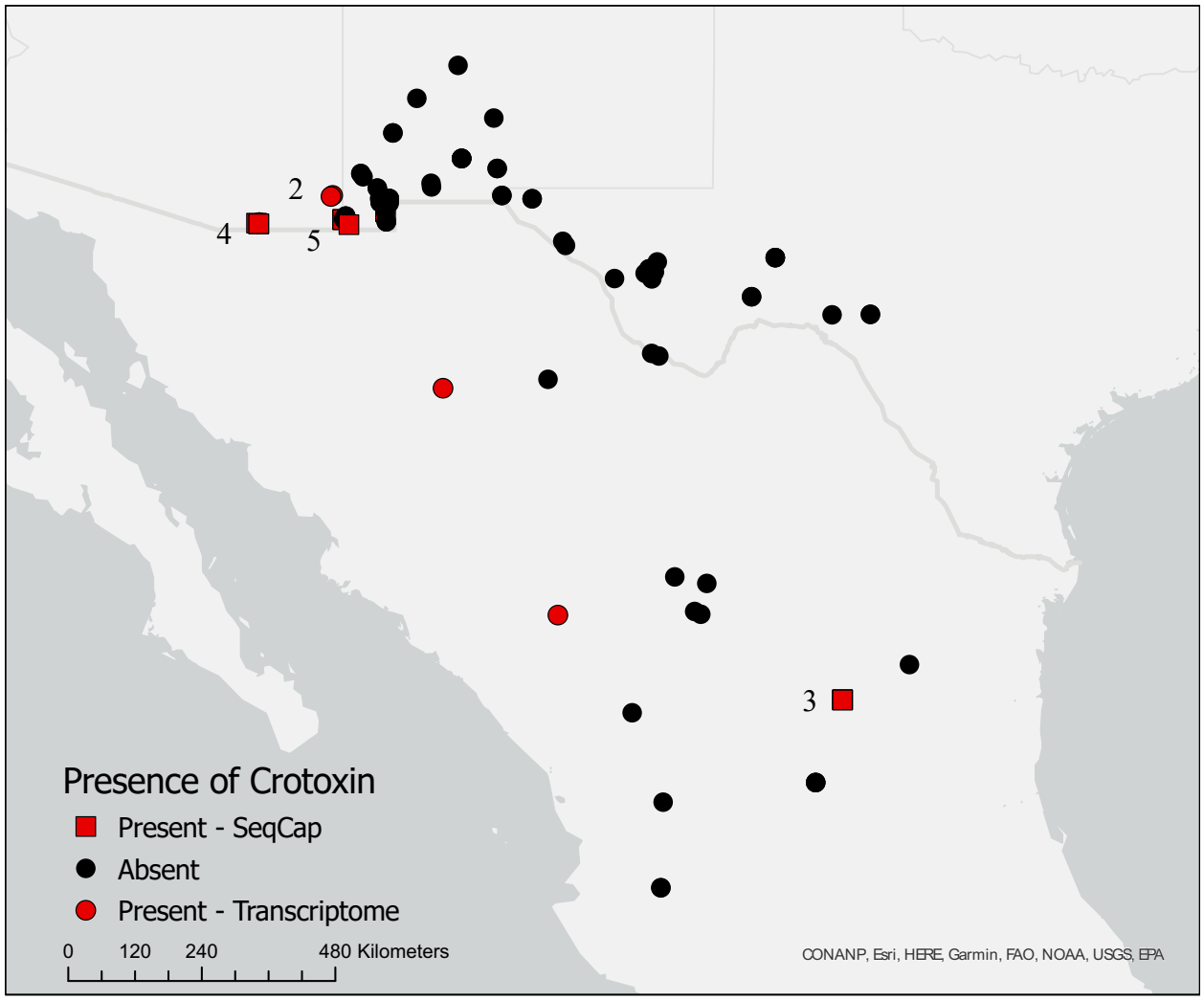

Figure 13. Map of all *C. lepidus* samples. Black dots represent transcriptomic and SeqCap samples that were CTx-. Red symbols represent CTx+ samples, with transcriptome samples being circles, and SeqCap squares. Numbers represent the number of CTx+ samples in each locality.

Results

Spatial Analyses

Of the 96 SeqCap samples, it was determined that 13 possessed crotoxin (Table 4). We combined the SeqCap data with the transcriptomic data to perform the following spatial analyses, bringing our total samples to 105, and 17 with crotoxin present (Figure 13).

Binary logistic regressions showed no significant association between crotoxin presence and latitude, longitude, elevation, or subspecies (Table 2). Regressions using WorldClim data yield significant results for BIO3 (Isothermality, pvalue = 0.0235), BIO4 (Temperature Seasonality, pvalue = 0.0463), BIO5 (Max Temp Warmest Month, pvalue =0.0440), BIO8 (Mean Temp Wettest Quarter), BIO10 (Mean Temp Warmest Quarter, pvalue = 0.0127), BIO18 (Precipitation Warmest Quarter, pvalue = 0.0199), and BIO19 (Precipitation Coldest Quarter, pvalue = 0.00193) (S.Table1).

Evolutionary origins

Table 3. Shows how different data types support different origin hypotheses.

Data Type	Convergent Evolution	Ancestral Trait	Hybridization
PLA2 Phylogeny	*C. lepidus* crotoxin cluster independently	*C. lepidus* crotoxin cluster basal to existing crotoxin	*C. lepidus* crotoxin clusters more closely to another species'
Spatial	CTx+ present regardless of distribution of other species. Extremely localized	CTx+ present regardless of distribution of other species. More widespread	CTx+ present at lower elevations and/or overlapping in distribution with to *C. scutulatus*/*C. tigris*
Patterson's D	No gene flow	No gene flow	Gene flow

In the case of an independent origin, or convergent evolution, crotoxin would require *Crotalus* to have lost crotoxin in the montane lineage (Table 5). The montane and lowland lineages split ~11 million years ago (Holding et al. 2021), and crotoxin has not been previously confirmed in the montane species. We would expect the *C. lepidus* crotoxin subunits to form a monophyletic cluster separate from the crotoxin subunits from the lowland rattlesnakes in the PLA2 phylogeny. Spatially CTx+ individuals will be present regardless of the distribution of the lowland species and will be extremely localized to account for the mutation occurring in the recent past. While the *C. lepidus* crotoxin subunits were generally basal to *Crotalus* identified subunits, the spatial distribution does not support this hypothesis. Crotoxin is wildly distributed throughout the range, not localized to a small area.

Our second hypothesis on crotoxin origin in *C. lepidus* was by a hybridization event with either *C. scutulatus* or *C. tigris*, taxa known to possess crotoxin and are co-distributed. Both lowland species express crotoxin and are found in the desert surrounding the Madrean Sky Islands. While *C. lepidus* is found in the pine forests at higher elevation, the distance between the habitat is short. In the PLA2 phylogeny, the *C. lepidus* crotoxin subunits would be expected to be sister to the subunits from the species they introgressed from. Introgression has been hypothesized as the source of crotoxin in *C. helleri* originating from *C. scutulatus*, and *Sistrurus catenatus* from *C. horridus* (Rokyta et al. 2015; Dowell et al. 2018). We would have also expected crotoxin to only occur in the lower elevations, along potential contact zones, and be dependent on neighboring *C. scutulatus* or *C. tigris* ranges. *Crotalus lepidus* crotoxin was basal within *Crotalus* in the maximum likelihood phylogeny, and its overall presence was not tied to a specific elevation range. We performed a Patterson's D to assess gene flow between populations by testing for deviations in incomplete lineage sorting. Using previously published

transcriptomes from *Sistrurus tergeminus*, *S. catenatus, S. miliarius, C. scutulatus,* and *C. tigris* (see table S.Table 2 (Holding et al. 2021)) We assessed gene flow using *Sistrurus* as the outgroup, testing for historical admixture between CTx+/CTx- *C. lepidus* individuals and CTx+ *C. scutulatus*, CTx- *C. scutulatus,* and *C. tigris*. No gene flow was found (Table 6).

Table 4. D-statistic values for each introgression trial, first looking at CTx+ *C. scutulatus*, CTx- *C. scutulatus*, and *C. tigris*. A Z score of |3| would have an alpha of 0.05. All trials have near equal ABBA vs. BABA site patterns, indicating no gene flow.

W	X	Y	Outgroup	Zscore	BABA	ABBA	nsnps
CTx+ *C. lepidus*	CTx- *C. lepidus*	CTx+ *C. scutulatus*	*Sisturus*	-0.951	250	261	85800
CTx+ *C. lepidus*	CTx- *C. lepidus*	CTx- *C. scutulatus*	*Sisturus*	-0.26	249	252	85784
CTx+ *C. lepidus*	CTx- *C. lepidus*	*C. tigris*	*Sisturus*	0.036	251	251	85750

Our final hypothesis was an ancestral origin of crotoxin. For an ancestral origin we would have expected the *C. lepidus* crotoxin subunits to be basal to existing Crotaline subunits to match the ancestral PLA2 array recovered in *C.lepidus*, which they were. Spatially we would have expected crotoxin to be distributed regardless of other species and be more widespread than when compared to the convergent evolution hypothesis. Since the species has such a large range, and venom is locally adaptive and subject to local selection pressures, there should be multiple opportunities for the trait to be under selection. Lack of gene flow would add additional support to the ancestral origin hypothesis. With the Patterson's D statistic, we found no support for gene flow, and the near equal amounts of ABBA vs. BABA site patterns support ILS.

Discussion

Our goal was to investigate the spatial distribution of crotoxin throughout the range of *C. lepidus*. Through spatial analyses and tests of gene flow, we are able to make inferences on the evolutionary history of crotoxin in montane rattlesnakes. We confirm the presence of crotoxin throughout the range of *C. lepidus* and identify environmental variables associated with the genomic presence of the trait (Figure 13, Table 2, Supp table 1). We found no evidence of gene

flow between *C. lepidus* and codistributed CTx+ species, thus we hypothesize that ancestral gene segregation and selection are responsible for maintaining crotoxin in *C. lepidus* and the distinct venom phenotypes across *Crotalus*.

While the *C. lepidus* Type I and Type II venom types do not group as distinct clusters as in Strickland, Smith, et al. (2018)**,** there are clusters of Type I and Type II individuals throughout the range of *C. lepidus*. The three individuals from San Luis Potosi are the only *C. l. lepidus* individuals that are CTx+, and the farthest to the south and east. While the transcriptome dataset had crotoxin being expressed at significantly higher levels in *C. l. klauberi*, that association is not true in our larger joint transcriptome and SeqCap analyses. This is due to the transcriptome dataset being only ten individuals and representing a smaller part of the overall species range. Additionally, this exhibits the importance of combining datatypes for the largest extent to better understand the evolutionary trends.

In previous studies with *C. scutulatus* BIO1; BIO4; BIO8, BIO9; BIO11, BIO12, BIO15, BIO17, and BIO19 were significant indicators of crotoxin presence (Strickland, Smith, et al. 2018). In *C. horridus* BIO4; BIO6; BIO8; and BIO18 were significantly correlated with crotoxin presence. Both studies hypothesized that CTx+ individuals had different niches and hypothesized that CTx presence could be a local adaptation to prey. *Crotalus scutulatus* and *C. horridus* predominantly prey on birds and small mammals (Platt et al. 2001; Zancolli et al. 2019; Margres, Wray, et al. 2021), and have been documented preying on lizards (Hamilton and Pollack 1955; Zancolli et al. 2019). *Crotalus lepidus* preys predominately on small mammals and lizards, while neonates also eat centipedes (Beaupre 1995; Holycross et al. 2002). Venom composition and prey type are closely intertwined (Holding et al. 2016; Holding et al. 2021), prey type is also closely associated with gape size, as snakes are gape limited predators (Shine 1991). While it is

highly likely that prey type plays a role in the local selection factors that select for crotoxin presence, it is unlikely to be the same exact species due to the influence of gape size.

Environment plays a major role in local adaption as a driver of selection (Li et al. 2018; Tobler et al. 2018; Albecker et al. 2021 Mar 2). The BIO4 (Temperature Seasonality) and BIO8 (Mean Temperature of Wettest Quarter) WorldClim values for *C. scutulatus* and *C. horridus* with known Type I vs Type II venom types were added to the *C. lepidus* data for analysis, t-tests revealed that both were significant (p = 2.3xe-4 and 1.84xe-14 respectively). When we expanded our analysis to all WorldClim variables, 13 of the 20 variables were significant, including variables that were not significant in any single species study. This could suggest an overarching environment factor that selects for the presence of crotoxin across *Crotalus*. A meta-analysis of climatic and prey type availability across variable expression species could provide useful insight in what factors could be exhibiting a selection pressure for crotoxin. While we have identified significantly correlated variables, how they are acting as selective pressures is still unknown.

Ancestral origin is supported by the literature, as the ancestral North American pitviper is hypothesized to have had crotoxin (Dowell et al. 2016; Whittington et al. 2018). In the PLA2 phylogeny, the *C. lepidus* crotoxin subunits would be basal to the lowland clade and would be present throughout the distribution. This wide distribution of crotoxin would be irrespective of the lowland species and would be consistent with crotoxin being an ancestral trait. The data so far support both parts of this hypothesis, the *C. lepidus* subunits are basal in the phylogeny, and can be found in both *C. l. lepidus* and *C. l. klauberi*. The lack of hybridization with *C. tigris* or *C. scutulatus* is additional data supporting this hypothesis.

Our CTx+ *C. lepidus* has a PLA2 array most similar to Dowell et al's (2016) hypothesized ancestral pitviper. Further genomic research into other montane species would help

illustrate their PLA2 composition, shed light on the small-bodied lineage in comparison with the large-bodied. Body size alone would impact the selection pressures each group faces due to gape size caused differences in diet, in addition to environment. As most small-bodied rattlesnakes are associated with montane habitats, and large-bodied rattlesnakes are associated with lowlands.

We hypothesize that crotoxin has been maintained through a combination of ancestral gene segregation and local selection. Ancestral gene segregation would have facilitated crotoxin to have persisted in New World ancestral pitvipers as they diverged from Old World pitvipers ~25MYA (Wüster et al. 2008; Alencar et al. 2016), and in subsequent radiations throughout the New World. Out of 22 genera of pitvipers, there are six genera that have crotoxin in at least one species (*Bothriechis* (Fernández et al. 2010), *Crotalus* (Hendon and Fraenkel-Conrat 1971), *Gloydius* (Yang, Yang, et al. 2015), *Lachesis* (Damico et al. 2005; Damico et al. 2012), *Mixcoatlus* (Neri-Castro et al. 2020), *Ophryacus* (Neri-Castro et al. 2019), and *Sistrurus* (Calvete et al. 2012)). From there, local selection factors would have maintained crotoxin as a venom component. Venom is highly adaptive and responsive to selection pressures (Gibbs and Rossiter 2008). Individual toxin genes are lost if not expressed (Strickland et al. *in prep*), indicating an emphasis on a stream-lined venome. Independent loss events are hypothesized to explain the varied presence of crotoxin in *Crotalus* (Dowell et al. 2016), they could also explain variation across the pitviper phylogeny as well. Future studies that thoroughly investigate the PLA2 presence across New World Pitvipers in addition to local selection pressure where crotoxin is and is not found can demystify both the evolution of how crotoxin originated and initially spread in this clade, and how it continues to be maintained.

REFERENCES

Aird SD, Kaiser II. 1985. Comparative studies on three rattlesnake toxins. Toxicon. 23(3):361–374. doi:10.1016/0041-0101(85)90020-0.

Albecker MA, Stuckert AMM, Balakrishnan CN, McCoy MW. 2021 Mar 2. Molecular mechanisms of local adaptation for salt-tolerance in a treefrog. Mol Ecol.:mec.15867. doi:10.1111/mec.15867. [accessed 2021 Mar 8]. https://onlinelibrary.wiley.com/doi/10.1111/mec.15867.

Alencar LRV V, Quental TB, Grazziotin FG, Alfaro ML, Martins M, Venzon M, Zaher H. 2016. Molecular Phylogenetics and Evolution Diversification in vipers : Phylogenetic relationships , time of divergence and shifts in speciation rates. Mol Phylogenet Evol. 105:50–62. doi:10.1016/j.ympev.2016.07.029. [accessed 2022 Jun 6]. http://dx.doi.org/10.1016/j.ympev.2016.07.029.

Armstrong BL, Murphy JB. 1979. The Natural History of Mexican Rattlesnakes. Wiley EO, Collins JT, editors. Lawrence, Kansas: University of Kansas Publications Museum of Natural History.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. 2000. Gene Ontology: tool for the unification of biology. Nat Genet 2000 251. 25(1):25–29. doi:10.1038/75556. [accessed 2022 May 16]. https://www.nature.com/articles/ng0500_25.

Bateman A, Martin MJ, Orchard S, Magrane M, Agivetova R, Ahmad S, Alpi E, Bowler-Barnett EH, Britto R, Bursteinas B, et al. 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. 49(D1):D480–D489. doi:10.1093/NAR/GKAA1100. [accessed 2022 May 16]. https://academic.oup.com/nar/article/49/D1/D480/6006196.

Beaupre SJ. 1995. Comparative Ecology of the Mottled Rock Rattlesnake , Crotalus lepidus , in Big Bend National Park Author (s): Steven J . Beaupre Published by : Allen Press on behalf of the Herpetologists ' League Stable URL : https://www.jstor.org/stable/3892784 REFER. Herpetologica. 51(1):45–56.

Beaupre SJ, Jacobson ER, Lillywhite HB, Zamudio K. 2004. GUIDELINES FOR USE OF LIVE AMPHIBIANS AND REPTILES IN FIELD AND LABORATORY RESEARCH. 2nd Editio. Lawrence, KS.

Bivand R, Keitt T, Rowlingson B. 2021. rgdal. https://cran.r-project.org/package=rgdal.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 2016 345. 34(5):525–527. doi:10.1038/nbt.3519. [accessed 2022 Apr 18]. https://www.nature.com/articles/nbt.3519.

Calvete JJ, Pérez A, Lomonte B, Sánchez EE, Sanz L. 2012. Snake venomics of Crotalus tigris: The minimalist toxin arsenal of the deadliest neartic rattlesnake venom. Evolutionary clues for generating a pan-specific antivenom against crotalid type II venoms. J Proteome Res. 11(2):1382–1390. doi:10.1021/pr201021d.

Campbell MA, Haas BJ, Hamilton JP, Mount SM, Robin CR. 2006. Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. BMC Genomics. 7(1):1–17. doi:10.1186/1471-2164-7-327/TABLES/5. [accessed 2022 Apr 18]. https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-7-327.

Cantarel BI, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. 37(Database issue):D233. doi:10.1093/NAR/GKN663. [accessed 2022 May 16]. /pmc/articles/PMC2686590/.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25(15):1972–1973. doi:10.1093/BIOINFORMATICS/BTP348. [accessed 2022 Apr 13]. https://academic.oup.com/bioinformatics/article/25/15/1972/213148.

Carbon S, Douglass E, Good BM, Unni DR, Harris NL, Mungall CJ, Basu S, Chisholm RL, Dodson RJ, Hartline E, et al. 2021. The Gene Ontology resource: enriching a GOld mine. Nucleic Acids Res. 49(D1):D325–D334. doi:10.1093/NAR/GKAA1113. [accessed 2022 May 16]. https://academic.oup.com/nar/article/49/D1/D325/6027811.

Carretero-Paulet L, Fares MA. 2012. Evolutionary Dynamics and Functional Specialization of Plant Paralogs Formed by Whole and Small-Scale Genome Duplications. Mol Biol Evol. 29(11):3541–3551. doi:10.1093/MOLBEV/MSS162. [accessed 2021 Nov 13]. https://academic.oup.com/mbe/article/29/11/3541/1155189.

Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG. 2013. Complex cocktails: the evolutionary novelty of venoms. Trends Ecol Evol. 28(4):219–229. doi:10.1016/J.TREE.2012.10.020.

Chang D, Duda TF. 2012. Extensive and Continuous Duplication Facilitates Rapid Evolution and Diversification of Gene Families. Mol Biol Evol. 29(8):2019–2029. doi:10.1093/MOLBEV/MSS068. [accessed 2021 Nov 13]. https://academic.oup.com/mbe/article/29/8/2019/1041756.

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods 2021 182. 18(2):170–175. doi:10.1038/s41592-020-01056-5. [accessed 2022 Apr 13]. https://www.nature.com/articles/s41592-020-01056-5.

Cho YB, Jones SI, Vodkin LO. 2019. Nonallelic homologous recombination events responsible for copy number variation within an RNA silencing locus. Plant Direct. 3(8):e00162. doi:10.1002/PLD3.162. [accessed 2022 Jun 3]. https://onlinelibrary.wiley.com/doi/full/10.1002/pld3.162.

Damico DCS, Bueno LGF, Rodrigues-Simioni L, Marangoni S, Da Cruz-Höfling MA, Novello JC. 2005. Neurotoxic and myotoxic actions from Lachesis muta muta (surucucu) whole venom on the mouse and chick nerve–muscle preparations. Toxicon. 46(2):222–229. doi:10.1016/J.TOXICON.2005.04.011.

Damico DCS, Vassequi-Silva T, Torres-Huaco FD, Nery-Diez ACC, de Souza RCG, Da Silva SL, Vicente CP, Mendes CB, Antunes E, Werneck CC, et al. 2012. LmrTX, a basic PLA2 (D49) purified from Lachesis muta rhombeata snake venom with enzymatic-related antithrombotic and anticoagulant activity. Toxicon. 60(5):773–781. doi:10.1016/J.TOXICON.2012.06.010.

Dowell NL, Giorgianni MW, Griffin S, Kassner VA, Selegue JE, Sanchez EE, Carroll SB. 2018. Extremely Divergent Haplotypes in Two Toxin Gene Complexes Encode Alternative Venom Types within Rattlesnake Species. Curr Biol. 28(7):1016-1026.e4. doi:10.1016/j.cub.2018.02.031. [accessed 2021 Jan 29]. https://doi.org/10.1016/j.cub.2018.02.031.

Dowell NL, Giorgianni MW, Kassner VA, Selegue JE, Sanchez EE, Carroll SB. 2016. The Deep Origin and Recent Loss of Venom Toxin Genes in Rattlesnakes. Curr Biol. 26(18):2434–2445. doi:10.1016/j.cub.2016.07.038. [accessed 2020 Feb 23]. https://linkinghub.elsevier.com/retrieve/pii/S0960982216307898.

Faure G, Xu H, Saul FA. 2011. Crystal Structure of Crotoxin Reveals Key Residues Involved in the Stability and Toxicity of This Potent Heterodimeric β-Neurotoxin. J Mol Biol. 412(2):176–191. doi:10.1016/J.JMB.2011.07.027.

Fernández J, Lomonte B, Sanz L, Angulo Y, Gutiérrez JM, Calvete JJ. 2010. Snake venomics of bothriechis nigroviridis reveals extreme variability among palm pitviper venoms: Different evolutionary solutions for the same trophic purpose. J Proteome Res. 9(8):4234–4241. doi:10.1021/PR100545D/ASSET/IMAGES/LARGE/PR-2010-00545D_0001.JPEG. [accessed 2022 Jun 6]. https://pubs.acs.org/doi/full/10.1021/pr100545d.

Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. Nat Rev Genet 2008 95. 9(5):397–405. doi:10.1038/nrg2337. [accessed 2022 Jun 3]. https://www.nature.com/articles/nrg2337.

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol. 37(12):4302–4315. doi:10.1002/JOC.5086. [accessed 2021 Nov 18]. https://onlinelibrary.wiley.com/doi/full/10.1002/joc.5086.

Flint J, Mackay TFC. 2009. Genetic architecture of quantitative traits in mice, flies, and humans. Genome Res. 19(5):723–733. doi:10.1101/gr.086660.108. [accessed 2022 Jun 7]. http://www.genome.org/cgi/doi/10.1101/gr.086660.108.

Forstner MRJ, Hilsenbeck RA, Scudday JF. 1997. Geographic Variation in Whole Venom Profiles from the Mottled Rock Rattlesnake (Crotalus lepidus lepidus) in Texas. J Herpetol. 31(2):277. doi:10.2307/1565397. [accessed 2019 Oct 16]. https://www.jstor.org/stable/1565397?origin=crossref.

Franco-Servín C, Neri-Castro E, Bénard-Valle M, Alagón A, Rosales-García RA, Guerrero-Alba R, Poblano-Sánchez JE, Silva-Briano M, Guerrero-Barrera AL, Sigala-Rodríguez JJ. 2021. Biological and Biochemical Characterization of Coronado Island Rattlesnake (Crotalus helleri caliginis) Venom and Antivenom Neutralization. Toxins 2021, Vol 13, Page 582. 13(8):582. doi:10.3390/TOXINS13080582. [accessed 2021 Oct 13]. https://www.mdpi.com/2072-6651/13/8/582/htm.

French WJ, Hayes WK, Bush SP, Cardwell MD, Bader JO, Rael ED. 2004. Mojave toxin in venom of Crotalus helleri (Southern Pacific Rattlesnake): molecular and geographic characterization. Toxicon. 44(7):781–791. doi:10.1016/J.TOXICON.2004.08.008.

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: Accelerated for clustering the next-generation sequencing data. Bioinformatics. 28(23):3150–3152. doi:10.1093/bioinformatics/bts565. [accessed 2021 Feb 15]. https://pubmed.ncbi.nlm.nih.gov/23060610/.

Gibbs HL, Mackessy SP. 2009. Functional basis of a molecular adaptation: Prey-specific toxic effects of venom from Sistrurus rattlesnakes. Toxicon. 53(6):672–679. doi:10.1016/j.toxicon.2009.01.034. [accessed 2021 Jul 8]. http://dx.doi.org/10.1016/j.toxicon.2009.01.034.

Gibbs HL, Rossiter W. 2008. 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. 66(2):151–166. doi:10.1007/S00239-008-9067-7/FIGURES/6. [accessed 2022 Jun 9]. https://link.springer.com/article/10.1007/s00239-008-9067-7.

Glenn JL, Straight RC. 1989. Intergradation of two different venom populations of the Mojave rattlesnake (Crotalus scutulatus scutulatus) in Arizona. Toxicon. 27(4):411–418. doi:10.1016/0041-0101(89)90203-1.

Glenn JL, Straight RC, Wolt TB. 1994. Regional variation in the presence of canebrake toxin in Crotalus horridus venom. Comp Biochem Physiol Part C Pharmacol Toxicol Endocrinol. 107(3):337–346. doi:10.1016/1367-8280(94)90059-0.

Gopalakrishnakone P, Hawgood BJ, Holbrooke SE, Marsh NA, Santana S, Sa DE, Tu AT. 1980. SITES OF ACTION OF MOJAVE TOXIN ISOLATED FROM THE VENOM OF THE MOJAVE RATTLESNAKE. [accessed 2018 Jul 25]. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2044271/pdf/brjpharm00658-0070.pdf.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 2011 297. 29(7):644–652. doi:10.1038/nbt.1883. [accessed 2022 Apr 18]. https://www.nature.com/articles/nbt.1883.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MHY, et al. 2010. A draft sequence of the neandertal genome. Science (80-). 328(5979):710–722. doi:10.1126/science.1188021. [accessed 2021 May 25]. http://science.sciencemag.org/.

Gremme G, Steinbiss S, Kurtz S. 2013. GenomeTools: A Comprehensive Software Library for Efficient Processing of Structured Genome Annotations. IEEE/ACM Trans Comput Biol Bioinforma. 10(03):645–656. doi:10.1109/TCBB.2013.68.

Guarnaccia AM, Krivoshik SR, Sparks JS, Gruber DF, Gaffney JP. 2021. Discovery and Characterization of a Bilirubin Inducible Green Fluorescent Protein From the Moray Eel Gymnothorax zonipectis. Front Mar Sci. 8:1265. doi:10.3389/FMARS.2021.678571/BIBTEX.

Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD, et al. 2003. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res. 31(19):5654–5666. doi:10.1093/NAR/GKG770. [accessed 2022 Apr 18]. https://academic.oup.com/nar/article/31/19/5654/1092988.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc 2013 88. 8(8):1494–1512. doi:10.1038/nprot.2013.084. [accessed 2021 Oct 19]. https://www.nature.com/articles/nprot.2013.084.

Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Robin CR, Wortman JR. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. 9(1):1–22. doi:10.1186/GB-2008-9-1-R7/TABLES/1. [accessed 2022 Apr 18]. https://genomebiology.biomedcentral.com/articles/10.1186/gb-2008-9-1-r7.

Hamilton W, Pollack J. 1955. The food of some crotalid snakes from Fort Benning, Georgia. Nat Hist Misc. 140:1–4.

Hanai TT, Smith RM. 1999. HPLC. Cambridge: Royal Society of Chemistry (RSC chromatography monographs).

Harvey EN. 1921. Studies on Bioluminescence. Biol Bull. 41(5):280–287. doi:10.2307/1536528. [accessed 2021 Nov 13]. https://www.journals.uchicago.edu/doi/abs/10.2307/1536528.

Hendon RA, Fraenkel-Conrat H. 1971. Biological Roles of the Two Components of Crotoxin. Proc Natl Acad Sci. 68(7):1560–1563. doi:10.1073/PNAS.68.7.1560. [accessed 2021 Nov 19]. https://www.pnas.org/content/68/7/1560.

Hijmans RJ. 2021. raster. https://cran.r-project.org/package=raster.

Holding ML, Biardi JE, Gibbs HL. 2016. Coevolution of venom function and venom resistance in a rattlesnake predator and its squirrel prey. Proc R Soc B Biol Sci. 283(1829). doi:10.1098/rspb.2015.2841.

Holding ML, Margres MJ, Mason AJ, Parkinson CL, Rokyta DR. 2018. Evaluating the Performance of De Novo Assembly Methods for Venom-Gland Transcriptomics. Toxins (Basel). 10(6). doi:10.3390/toxins10060249. [accessed 2018 Jul 26]. http://www.ncbi.nlm.nih.gov/pubmed/29921759.

Holding ML, Strickland JL, Rautsaw RM, Hofmann EP, Mason AJ, Hogan MP, Nystrom GS, Ellsworth SA, Colston TJ, Borja M, et al. 2021. Phylogenetically diverse diets favor more complex venoms in North American pitvipers. Proc Natl Acad Sci. 118(17):e2015579118. doi:10.1073/pnas.2015579118. http://www.pnas.org/lookup/doi/10.1073/pnas.2015579118.

Holycross AT, Painter CW, Prival DB, Swann DE, Schroff MJ, Edwards T, Schwalbe CR. 2002. Diet of Crotalus lepidus klauberi (banded rock rattlesnake). J Herpetol. 36(4):589–597. doi:10.1670/0022-1511(2002)036[0589:DOCLKB]2.0.CO;2.

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, et al. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. 47(D1):D309–D314. doi:10.1093/NAR/GKY1085. [accessed 2022 May 16]. https://academic.oup.com/nar/article/47/D1/D309/5173662.

Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics. 30(9):1236–1240. doi:10.1093/BIOINFORMATICS/BTU031. [accessed 2022 May 16]. https://academic.oup.com/bioinformatics/article/30/9/1236/237988.

Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 30(4):772–780. doi:10.1093/MOLBEV/MST010. [accessed 2022 Apr 13]. https://academic.oup.com/mbe/article/30/4/772/1073398.

Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 2019 378. 37(8):907–915. doi:10.1038/s41587-019-0201-4. [accessed 2022 Apr 13]. https://www.nature.com/articles/s41587-019-0201-4.

Labonne J, Hendry AP. 2010. Natural and sexual selection giveth and taketh away reproductive barriers: Models of population divergence in guppies. Am Nat. 176(1):26–39. doi:10.1086/652992/SUPPL_FILE/51341APA.PDF. [accessed 2021 Nov 18]. https://www.journals.uchicago.edu/doi/abs/10.1086/652992.

Lamichhaney S, Han F, Webster MT, Andersson L, Grant BR, Grant PR. 2018. Rapid hybrid speciation in Darwin's Finches. Science (80-). 359(6372):224–228. doi:10.1126/SCIENCE.AAO4593/SUPPL_FILE/PAP.PDF. [accessed 2021 Nov 18]. https://www.science.org/doi/abs/10.1126/science.aao4593.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012 94. 9(4):357–359. doi:10.1038/nmeth.1923. [accessed 2022 Feb 3]. https://www.nature.com/articles/nmeth.1923.

Langmead B, Wilks C, Antonescu V, Charles R. 2019. Scaling read aligners to hundreds of threads on general-purpose processors. Bioinformatics. 35(3):421–432. doi:10.1093/BIOINFORMATICS/BTY648. [accessed 2022 Feb 3]. https://academic.oup.com/bioinformatics/article/35/3/421/5055585.

Lemmon AR, Emme SA, Lemmon EM. 2012. Anchored Hybrid Enrichment for Massively High-Throughput Phylogenomics. Syst Biol. 61(5):727–744. doi:10.1093/SYSBIO/SYS049. [accessed 2021 Aug 18]. https://academic.oup.com/sysbio/article/61/5/727/1737120.

Li B, Dewey CN. 2011. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 12(1):1–16. doi:10.1186/1471-2105-12-323/TABLES/6. [accessed 2022 Feb 3]. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-323.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 34(18):3094–3100. doi:10.1093/BIOINFORMATICS/BTY191. [accessed 2022 Apr 13]. https://academic.oup.com/bioinformatics/article/34/18/3094/4994778.

Li H. 2021. New strategies to improve minimap2 alignment accuracy. Bioinformatics. 37(23):4572–4574. doi:10.1093/BIOINFORMATICS/BTAB705. [accessed 2022 Apr 13]. https://academic.oup.com/bioinformatics/article/37/23/4572/6384570.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 25(14):1754–1760. doi:10.1093/BIOINFORMATICS/BTP324. [accessed 2021 Nov 13]. https://academic.oup.com/bioinformatics/article/25/14/1754/225615.

Li Q, Grossenbacher DL, Angert AL. 2018. The effect of range overlap on ecological niche divergence depends on spatial scale in monkeyflowers. Evolution (N Y). 72(10):2100–2113. doi:10.1111/evo.13567. https://onlinelibrary.wiley.com/doi/10.1111/evo.13567.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 22(13):1658–1659. doi:10.1093/bioinformatics/btl158. [accessed 2021 Feb 15]. https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btl158.

Liu Y, Mao X, Krause J, Fu Q. 2021. Insights into human history from the first decade of ancient human genomics. Science (80-). 373(6562):1479–1484. doi:10.1126/SCIENCE.ABI8202/ASSET/7ADCE72A-2C49-4E0E-801D-D3147B06E4DF/ASSETS/IMAGES/LARGE/SCIENCE.ABI8202-F3.JPG. [accessed 2021 Nov 13]. https://www-science-org.libproxy.clemson.edu/doi/abs/10.1126/science.abi8202.

Lomonte B, Angulo Y, Sasa M, María Gutiérrez J. 2009. The Phospholipase A 2 Homologues of Snake Venoms: Biological Activities and Their Possible Adaptive Roles.

Lomonte B, Rangel J. 2012. Snake venom Lys49 myotoxins: From phospholipases A 2 to non-enzymatic membrane disruptors. Toxicon. 60(4):520–530. doi:10.1016/j.toxicon.2012.02.007.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):1–21. doi:10.1186/S13059-014-0550-8/FIGURES/9. [accessed 2022 Feb 3]. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8.

Mackessy SP. 2008. Venom composition in rattlesnakes: trends and biological significance. In: Hayes WK, Beaman KR, Cardwell MD, Bush SP, editors. The biology of rattlesnakes. Loma Linda, CA: Loma Linda University Press. p. 495–510.

Mackessy SP, editor. 2010. Handbook of Venoms and Toxins of Reptiles. Boca Raton: Taylor & Francis.

Mahadevan J, Pathak AK, Vemula A, Nadella RK, Viswanath B, Jain S, Rao NP, Narayanaswamy JC, Viswanath B, Sivakumar PT, et al. 2021. Analysis of whole exome sequencing in severe mental illness hints at selection of brain development and immune related genes. Sci Reports 2021 111. 11(1):1–10. doi:10.1038/S41598-021-00123-X. [accessed 2021 Nov 13]. https://www-nature-com.libproxy.clemson.edu/articles/s41598-021-00123-x.

Margres MJ, Bigelow AT, Lemmon EM, Lemmon AR, Rokyta DR. 2017. Selection To Increase Expression, Not Sequence Diversity, Precedes Gene Family Origin and Expansion in Rattlesnake Venom. Genetics. 206(3):1569–1580. doi:10.1534/genetics.117.202655. [accessed 2021 Feb 23]. https://doi.org/10.1534/genetics.117.202655.

Margres MJ, Rautsaw RM, Strickland JL, Mason AJ, Schramer TD, Hofmann EP, Stiers E, Ellsworth SA, Nystrom GS, Hogan MP, et al. 2021. The Tiger Rattlesnake genome reveals a complex genotype underlying a simple venom phenotype. Proc Natl Acad Sci. 118(4). doi:10.1073/PNAS.2014634118. [accessed 2021 Feb 5]. https://www.pnas.org/content/118/4/e2014634118.

Margres MJ, Wray KP, Hassinger ATB, Ward MJ, McGivern JJ, Moriarty Lemmon E, Lemmon AR, Rokyta DR. 2017. Quantity, Not Quality: Rapid Adaptation in a Polygenic Trait Proceeded Exclusively through Expression Differentiation. Mol Biol Evol. 34(12):3099–3110. doi:10.1093/molbev/msx231. [accessed 2021 Feb 24]. http://academic.oup.com/mbe/article/34/12/3099/4097632.

Margres MJ, Wray KP, Sanader D, McDonald PJ, Trumbull LM, Patton AH, Rokyta DR. 2021. Varying Intensities of Introgression Obscure Incipient Venom-Associated Speciation in the Timber Rattlesnake (Crotalus horridus). Toxins 2021, Vol 13, Page 782. 13(11):782. doi:10.3390/TOXINS13110782. [accessed 2021 Nov 18]. https://www.mdpi.com/2072-6651/13/11/782/htm.

Martínez-Romero G, Rucavado A, Lazcano D, Gutiérrez JM, Borja M, Lomonte B, Garza-García Y, Zugasti-Cruz A. 2013. Comparison of venom composition and biological activities of the subspecies Crotalus lepidus lepidus, Crotalus lepidus klauberi and Crotalus lepidus morulus from Mexico. Toxicon. 71:84–95. doi:10.1016/j.toxicon.2013.05.006. [accessed 2018 Aug 2]. https://www.sciencedirect.com/science/article/pii/S0041010113001876.

Massey DJ, Calvete JJ, Sánchez EE, Sanz L, Richards K, Curtis R, Boesen K. 2012. Venom variability and envenoming severity outcomes of the Crotalus scutulatus scutulatus (Mojave rattlesnake) from Southern Arizona. J Proteomics. 75(9):2576–2587. doi:10.1016/J.JPROT.2012.02.035.

Mata-Silva V, Desantis DL, Wagler AE, Johnson JD. 2018. Spatial Ecology of Rock Rattlesnakes (Crotalus lepidus) in Far West Texas. Herpetologica. 74(3):245–254. doi:10.1655/Herpetologica-D-16-00030.1.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20(9):1297–1303. doi:10.1101/GR.107524.110. [accessed 2021 Nov 13]. https://genome.cshlp.org/content/20/9/1297.full.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Teeling E, editor. Mol Biol Evol. 37(5):1530–1534. doi:10.1093/molbev/msaa015. [accessed 2021 Feb 15]. https://academic.oup.com/mbe/article/37/5/1530/5721363.

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, et al. 2021. Pfam: The protein families database in 2021. Nucleic Acids Res. 49(D1):D412–D419. doi:10.1093/NAR/GKAA913. [accessed 2022 May 16]. https://academic.oup.com/nar/article/49/D1/D412/5943818.

Nachtigall Pedro G., Kashiwabara AY, Durham AM. 2021. CodAn: predictive models for precise identification of coding regions in eukaryotic transcripts. Brief Bioinform. 22(3):1–11. doi:10.1093/BIB/BBAA045. [accessed 2022 May 19]. https://academic.oup.com/bib/article/22/3/bbaa045/5847603.

Nachtigall Pedro G, Rautsaw RM, Ellsworth SA, Mason AJ, Rokyta DR, Parkinson CL, Junqueira-de-Azevedo ILM. 2021. ToxCodAn: a new toxin annotator and guide to venom gland transcriptomics. Brief Bioinform. doi:10.1093/BIB/BBAB095. [accessed 2021 Jul 11]. https://academic.oup.com/bib/advance-article/doi/10.1093/bib/bbab095/6235957.

Nagelkerke SQ, Tacke CE, Breunis WB, Geissler J, Sins JWR, Appelhof B, Van Den Berg TK, De Boer M, Kuijpers TW. 2015. Nonallelic homologous recombination of the FCGR2/3 locus results in copy number variation and novel chimeric FCGR2 genes with aberrant functional expression. Genes Immun 2015 166. 16(6):422–429. doi:10.1038/gene.2015.25. [accessed 2022 Jun 3]. https://www.nature.com/articles/gene201525.

Neri-Castro E, Lomonte B, Valdés M, Ponce-López R, Bénard-Valle M, Borja M, Strickland JL, Jones JM, Grünwald C, Zamudio F, et al. 2019. Venom characterization of the three species of Ophryacus and proteomic profiling of O. sphenophrys unveils Sphenotoxin, a novel Crotoxin-like heterodimeric β-neurotoxin. J Proteomics. 192:196–207. doi:10.1016/j.jprot.2018.09.002.

Neri-Castro E, Sanz L, Olvera-Rodríguez A, Bénard-Valle M, Alagón A, Calvete JJ. 2020. Venomics and biochemical analysis of the black-tailed horned pitviper, Mixcoatlus melanurus, and characterization of Melanurutoxin, a novel crotoxin homolog. J Proteomics. 225:103865. doi:10.1016/j.jprot.2020.103865.

Ou S, Jiang N. 2018. LTR_retriever: A Highly Accurate and Sensitive Program for Identification of Long Terminal Repeat Retrotransposons. Plant Physiol. 176(2):1410–1422. doi:10.1104/PP.17.01310. [accessed 2022 Apr 13]. https://academic.oup.com/plphys/article/176/2/1410/6117145.

Ou S, Su W, Liao Y, Chougule K, Agda JRA, Hellinga AJ, Lugo CSB, Elliott TA, Ware D, Peterson T, et al. 2019. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. Genome Biol. 20(1):1–18. doi:10.1186/S13059-019-1905-Y/FIGURES/6. [accessed 2022 Apr 13]. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-019-1905-y.

Parks MM, Lawrence CE, Raphael BJ. 2015. Detecting non-allelic homologous recombination from high-throughput sequencing data. Genome Biol. 16(1):1–19. doi:10.1186/S13059-015-0633-1/FIGURES/5. [accessed 2022 Jun 3]. https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0633-1.

Patterson N, Price AL, Reich D. 2006. Population Structure and Eigenanalysis. PLOS Genet. 2(12):e190. doi:10.1371/JOURNAL.PGEN.0020190. [accessed 2022 May 19]. https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0020190.

Perry BW, Gopalan SS, Pasquesi GIM, Schield DR, Westfall AK, Smith CF, Koludarov I, Chippindale PT, Pellegrino MW, Chuong EB, et al. 2022. Snake venom gene expression is coordinated by novel regulatory architecture and the integration of multiple co-opted vertebrate pathways literature on snake venom systems, the mechanistic under-pinning of the regulation of snake VGs and the evolutionary. doi:10.1101/gr.276251.121.

Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 2015 333. 33(3):290–295. doi:10.1038/nbt.3122. [accessed 2022 Apr 13]. https://www.nature.com/articles/nbt.3122.

Petr M, Vernot B, Kelso J. 2019. admixr—R package for reproducible analyses using ADMIXTOOLS. Bioinformatics. 35(17):3194–3195. doi:10.1093/BIOINFORMATICS/BTZ030. [accessed 2022 May 16]. https://academic.oup.com/bioinformatics/article/35/17/3194/5298728.

Platt SG, Hawkes AW, Rainwater TR. 2001. Diet of the canebrake rattlesnake (Crotalus horridus atricaudatus): An additional record and review. Texas J Sci. 53(2):115–120.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006 388. 38(8):904–909. doi:10.1038/ng1847. [accessed 2022 May 19]. https://www.nature.com/articles/ng1847.

Price TD, Grant PR, Lisle Gibbs H, Boag PT. 1984. Recurrent patterns of natural selection in a population of Darwin's finches. Nature. 309(5971):787–789.

Prum RO, Berv JS, Dornburg A, Field DJ, Townsend JP, Lemmon EM, Lemmon AR. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nat 2015 5267574. 526(7574):569–573. doi:10.1038/nature15697. [accessed 2021 Aug 18]. https://www.nature.com/articles/nature15697.

R CoreTeam. 2021. R: A Language and Environment for Statistical Computing.

Rael ED, Johnson JD, Molina O, McCrystal HK. 1992. Distribution of a Mojave Toxin-Like Protein in Rock Rattlesnake (Crotalus lepidus) Venom.

Rautsaw RM, Jiménez-Velázquez G, Hofmann EP, Alencar LR V, Grünwald CI, Martins M, Carrasco P, Doan TM, Parkinson CL. 2022. VenomMaps: Updated species distribution maps and models for New World pitvipers (Viperidae: Crotalinae). Sci Data. 9(1):232. doi:10.1038/s41597-022-01323-4. https://www.nature.com/articles/s41597-022-01323-4.

Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res. 46(D1):D624–D632. doi:10.1093/NAR/GKX1134. [accessed 2022 May 16]. https://academic.oup.com/nar/article/46/D1/D624/4626772.

Rivas E, Neri-Castro E, Bénard-Valle M, Hernánez-Dávila AI, Zamudio F, Alagón A. 2017. General characterization of the venoms from two species of rattlesnakes and an intergrade population (C. lepidus x aquilus) from Aguascalientes and Zacatecas, Mexico. Toxicon. 138:191–195. doi:10.1016/J.TOXICON.2017.09.002.

Robberecht C, Voet T, Esteki MZ, Nowakowska BA, Vermeesch JR. 2013. Nonallelic homologous recombination between retrotransposable elements is a driver of de novo unbalanced translocations. Genome Res. 23(3):411–418. doi:10.1101/GR.145631.112. [accessed 2022 Jun 3]. https://genome.cshlp.org/content/23/3/411.full.

Robertson A. 1967. The nature of quantitative genetic variation. In: Brink RA, editor. Heritage from Mendel. Madison, Wisconsin: University of Wisconsin Press. p. 265–280.

Rokyta DR, Lemmon AR, Margres MJ, Aronow K. 2012. The venom-gland transcriptome of the eastern diamondback rattlesnake (Crotalus adamanteus). BMC Genomics.:1–23.

Rokyta DR, Margres MJ, Ward MJ, Sanchez EE. 2017. The genetics of venom ontogeny in the eastern diamondback rattlesnake (*Crotalus adamanteus*). PeerJ. 5. doi:10.7717/peerj.3249.

Rokyta DR, Wray KP, Margres MJ. 2013. The genesis of an exceptionally lethal venom in the timber rattlesnake (Crotalus horridus) revealed through comparative venom-gland transcriptomics. BMC Genomics. 14(1):394. doi:10.1186/1471-2164-14-394. [accessed 2021 Jan 12]. /pmc/articles/PMC3701607/?report=abstract.

Rokyta DR, Wray KP, McGivern JJ, Margres MJ. 2015. The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the Timber Rattlesnake (Crotalus horridus). Toxicon. 98:34–48. doi:10.1016/j.toxicon.2015.02.015. [accessed 2022 May 3]. https://linkinghub.elsevier.com/retrieve/pii/S0041010115000549.

Rotenberg D, Bamberger ES, Kochva E. 1971. Studies on Ribonucleic Acid Synthesis in the Venom Glands of Vipera palaestinae (Ophidia, Reptilia). Biochem J. 121:609–612.

Ryder OA, Thomas S, Judson JM, Romanov MN, Dandekar S, Papp JC, Sidak-Loftis LC, Walker K, Stalis IH, Mace M, et al. 2021. Facultative Parthenogenesis in California Condors. J Hered. 2021:1–6. doi:10.1093/JHERED/ESAB052. [accessed 2021 Nov 13]. https://academic.oup.com/jhered/advance-article/doi/10.1093/jhered/esab052/6412509.

Salih A, Larkum A, Cox G, Kühl M, Hoegh-Guldberg O. 2000. Fluorescent pigments in corals are photoprotective. Nat 2000 4086814. 408(6814):850–853. doi:10.1038/35048564. [accessed 2021 Nov 13]. https://www.nature.com/articles/35048564.

Saviola AJ, Gandara AJ, Bryson RW, Mackessy SP. 2017. Venom phenotypes of the Rock Rattlesnake (Crotalus lepidus) and the Ridge-nosed Rattlesnake (Crotalus willardi) from México and the United States. Toxicon. 138:119–129. doi:10.1016/J.TOXICON.2017.08.016. [accessed 2019 Mar 25]. https://www.sciencedirect.com/science/article/pii/S0041010117302581?via%3Dihub.

Schlichting CD, Pigliucci M. 1995. Gene regulation, quantitative genetics and the evolution of reaction norms. Evol Ecol. 9:154–168. [accessed 2021 Nov 18]. https://philarchive.org/archive/SCHGRQ.

Shi J, Liang C. 2019. Generic Repeat Finder: A High-Sensitivity Tool for Genome-Wide De Novo Repeat Detection. Plant Physiol. 180(4):1803–1815. doi:10.1104/PP.19.00386. [accessed 2022 Apr 18]. https://academic.oup.com/plphys/article/180/4/1803/6117697.

Shine R. 1991. Why do Larger Snakes Eat Larger Prey Items? Funct Ecol. 5(4):493. doi:10.2307/2389631. https://www.jstor.org/stable/2389631?origin=crossref.

Slon V, Mafessoni F, Vernot B, de Filippo C, Grote S, Viola B, Hajdinjak M, Peyrégne S, Nagel S, Brown S, et al. 2018. The genome of the offspring of a Neanderthal mother and a Denisovan father. Nat 2018 5617721. 561(7721):113–116. doi:10.1038/s41586-018-0455-x. [accessed 2021 Nov 13]. https://www.nature.com/articles/s41586-018-0455-x.

Slotta KH, Fraenkel-Conrat H. 1939. Crotoxin. Nat 1939 1443641. 144(3641):290–291. doi:10.1038/144290c0. [accessed 2022 Jun 14]. https://www.nature.com/articles/144290c0.

Smith A, Hubley R. 2015. RepeatModeler Open-1.0.

Sparks JS, Schelly RC, Smith WL, Davis MP, Tchernov D, Pieribone VA, Gruber DF. 2014. The Covert World of Fish Biofluorescence: A Phylogenetically Widespread and Phenotypically Variable Phenomenon. PLoS One. 9(1):e83259. doi:10.1371/JOURNAL.PONE.0083259. [accessed 2021 Nov 12]. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0083259.

Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics. 24(5):637–644. doi:10.1093/BIOINFORMATICS/BTN013. [accessed 2022 Apr 18]. https://academic.oup.com/bioinformatics/article/24/5/637/202844.

Stanley J, Arendt C. 2020. tidyjson. https://cran.r-project.org/package=tidyjson.

Strickland JL, Borja M, Rautsaw RM, Mason AJ, Castañeda-gaytán G, Rokyta DR, Parkinson CL. The genetic basis for natural variation in rattlesnake venoms. prep.

Strickland JL, Mason AJ, Rokyta DR, Parkinson CL. 2018. Phenotypic variation in mojave rattlesnake (Crotalus scutulatus) venom is driven by four toxin families. Toxins (Basel). 10(4):1–23. doi:10.3390/toxins10040135.

Strickland JL, Smith CF, Mason AJ, Schield DR, Borja M, Castañeda-Gaytán G, Spencer CL, Smith LL, Trápaga A, Bouzid NM, et al. 2018. Evidence for divergent patterns of local selection driving venom variation in Mojave Rattlesnakes (Crotalus scutulatus). Sci Rep. 8(1):17622. doi:10.1038/s41598-018-35810-9. [accessed 2019 Jan 30]. http://www.nature.com/articles/s41598-018-35810-9.

Su W, Gu X, Peterson T. 2019. TIR-Learner, a New Ensemble Method for TIR Transposable Element Annotation, Provides Evidence for Abundant New Transposable Elements in the Maize Genome. Mol Plant. 12(3):447–460. doi:10.1016/J.MOLP.2019.02.008.

Tanner WW, Dixon JR, Harris Jr HS, Harris HS, Harris HS. 1972. A New Subspecies of Crotalus Lepidus from Western Mexico. Gt Basin Nat. 32(1):16–24. doi:10.5962/bhl.part.25765. [accessed 2019 Oct 21]. https://www.biodiversitylibrary.org/part/25765.

Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. 18(12):1979–1990. doi:10.1101/GR.081612.108. [accessed 2022 Apr 18]. https://genome.cshlp.org/content/18/12/1979.full.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22(22):4673. doi:10.1093/NAR/22.22.4673. [accessed 2021 Nov 19]. /pmc/articles/PMC308517/?report=abstract.

Tobler M, Kelley JL, Plath | Martin, R€ |, Riesch U. 2018. Extreme environments and the origins of biodiversity: Adaptation and speciation in sulphide spring fishes. doi:10.1111/mec.14497.

Tsutsui K, Shimada E, Ogawa T, Tsuruwaka Y. 2016. A novel fluorescent protein from the deep-sea anemone Cribrinopsis japonica (Anthozoa: Actiniaria). Sci Reports 2016 61. 6(1):1–9. doi:10.1038/srep23493. [accessed 2021 Nov 13]. https://www.nature.com/articles/srep23493.

Whittington AC, Mason AJ, Rokyta DR. 2018. A single mutation unlocks cascading exaptations in the origin of a potent pitviper neurotoxin. Mol Biol Evol. 35(4):887–898. doi:10.1093/molbev/msx334. [accessed 2020 Sep 21]. https://pubmed.ncbi.nlm.nih.gov/29329419/.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wüster W, Peppin L, Pook CE, Walker DE. 2008. A nesting of vipers: Phylogeny and historical biogeography of the Viperidae (Squamata: Serpentes). Mol Phylogenet Evol. 49(2):445–459. doi:10.1016/J.YMPEV.2008.08.019.

Xiong W, He L, Lai J, Dooner HK, Du C. 2014. HelitronScanner uncovers a large overlooked cache of Helitron transposons in many plant genomes. Proc Natl Acad Sci U S A. 111(28):10263–10268. doi:10.1073/PNAS.1410068111/SUPPL_FILE/PNAS.1410068111.SAPP.PDF. [accessed 2022 Apr 18]. www.pnas.org/cgi/doi/10.1073/pnas.1410068111.

Xu Z, Wang H. 2007. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. 35(suppl_2):W265–W268. doi:10.1093/NAR/GKM286. [accessed 2022 Apr 18]. https://academic.oup.com/nar/article/35/suppl_2/W265/2920813.

Yang ZM, Guo Q, Ma ZR, Chen Y, Wang ZZ, Wang XM, Wang YM, Tsai IH. 2015. Structures and functions of crotoxin-like heterodimers and acidic phospholipases A2 from Gloydius intermedius venom: Insights into the origin of neurotoxic-type rattlesnakes. J Proteomics. 112:210–223.

Yang ZM, Yang YE, Chen Y, Cao J, Zhang C, Liu LL, Wang ZZ, Wang XM, Wang YM, Tsai IH. 2015. Transcriptome and proteome of the highly neurotoxic venom of Gloydius intermedius. Toxicon. 107:175–186. doi:10.1016/j.toxicon.2015.08.010.

Zancolli G, Calvete JJ, Cardwell MD, Greene HW, Hayes WK, Hegarty MJ, Herrmann H-WW, Holycross AT, Lannutti DI, Mulley JF, et al. 2019. When one phenotype is not enough: divergent evolutionary trajectories govern venom variation in a widespread rattlesnake species. Proc R Soc B Biol Sci. 286(1898):20182735. doi:10.1098/rspb.2018.2735. [accessed 2019 Mar 14]. http://www.royalsocietypublishing.org/doi/10.1098/rspb.2018.2735.

Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics. 30(5):614–620. doi:10.1093/BIOINFORMATICS/BTT593. [accessed 2021 Nov 13]. https://academic.oup.com/bioinformatics/article/30/5/614/247231.

Zhang R-G, Li G-Y, Wang X-L, Dainat J, Wang Z-X, Ou S, Ma Y. 2022. TEsorter: An accurate and fast method to classify LTR-retrotransposons in plant genomes. Hortic Res. 9. doi:10.1093/HR/UHAC017. [accessed 2022 Apr 18]. https://academic.oup.com/hr/article/doi/10.1093/hr/uhac017/6532211.

Zimin A V., Puiu D, Luo MC, Zhu T, Koren S, Marçais G, Yorke JA, Dvořák J, Salzberg SL. 2017. Hybrid assembly of the large and highly repetitive genome of Aegilops tauschii, a progenitor of bread wheat, with the MaSuRCA mega-reads algorithm. Genome Res. 27(5):787–792. doi:10.1101/GR.213405.116/-/DC1. [accessed 2022 Apr 18]. https://genome.cshlp.org/content/27/5/787.full.