

Electronic Theses and Dissertations, 2004-2019

2018

Evolution and Distribution of Phenotypic Diversity in the Venom of Mojave Rattlesnakes (*Crotalus scutulatus*)

Jason Strickland
University of Central Florida

 Part of the [Integrative Biology Commons](#), and the [Natural Resources and Conservation Commons](#)
Find similar works at: <https://stars.library.ucf.edu/etd>
University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

STARS Citation

Strickland, Jason, "Evolution and Distribution of Phenotypic Diversity in the Venom of Mojave Rattlesnakes (*Crotalus scutulatus*)" (2018). *Electronic Theses and Dissertations, 2004-2019*. 6054.
<https://stars.library.ucf.edu/etd/6054>

EVOLUTION AND DISTRIBUTION OF PHENOTYPIC DIVERSITY IN THE VENOM OF
MOJAVE RATTLESNAKES (*CROTALUS SCUTULATUS*)

by

JASON STRICKLAND
B.S. Angelo State University, 2009
M.S. Angelo State University, 2011

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the Department of Biology
in the College of Sciences
at the University of Central Florida

Summer Term
2018

Major Professors: Christopher L. Parkinson and Anna E. Savage

© 2018 Jason Strickland

ABSTRACT

Intraspecific phenotype diversity allows for local adaption and the ability for species to respond to changing environmental conditions, enhancing survivability. Phenotypic variation could be stochastic, genetically based, and/or the result of different environmental conditions. Mojave Rattlesnakes, *Crotalus scutulatus*, are known to have high intraspecific venom variation, but the geographic extent of the variation and factors influencing venom evolution are poorly understood. Three primary venom types have been described in this species based on the presence (Type A) or absence (Type B) of a neurotoxic phospholipase A₂ called Mojave toxin and an inverse relationship with the presence of snake venom metalloproteinases (SVMPs). Individuals that contain both Mojave toxin and SVMPs, although rare, are the third, and designated Type A + B. I sought to describe the proteomic and transcriptomic venom diversity of *C. scutulatus* across its range and test whether diversity was correlated with genetic or environmental differences. This study includes the highest geographic sampling of Mojave Rattlesnakes and includes the most venom-gland transcriptomes known for one species. Of the four mitochondrial lineages known, only one was monophyletic for venom type. Environmental variables poorly correlated with the phenotypes. Variability in toxin and toxin family composition of venom transcriptomes was largely due to differences in transcript expression. Four of 19 toxin families identified in *C. scutulatus* account for the majority of differences in toxin number and expression variation. I was able to determine that the toxins primarily responsible for venom types are inherited in a Mendelian fashion and that toxin expression is additive when comparing heterozygotes and homozygotes. Using the genetics to define venom type is more informative and the Type A + B phenotype is not unique, but rather heterozygous for the PLA₂ and/or SVMP alleles. Intraspecific venom variation in *C. scutulatus* highlights the need for fine scale ecological and natural history information to understand how phenotypic diversity is generated and maintained geographically through time.

To Stacy for always being by my side and
to the life-long friends I now have because of this journey.

ACKNOWLEDGMENTS

In my time at UCF, I have been fortunate to interact with many wonderful people personally and professional that made this experience unforgettable. I would first like to thank my advisor, Chris Parkinson. From the moment Stacy and I stepped into your office after driving a U-haul half-way across the country, you welcomed us into your family. Cynthia and Caroline also welcomed us with open arms and Caroline taught me many things including that hermit crabs really will change shells right in front of you with enough options.

My dissertation committee has been immensely helpful and supportive and I would not have been able to make it without them. Anna, thank you for adopting me into your lab and taking on the added responsibility of a stressed out Ph.D. student and always providing valuable advice. Eric, you were stuck with me from the beginning and I will always consider myself a pH lab member and you a great mentor who was willing to discuss my stubborn opinions about species. Darin, thank you for your support, guidance, and patience when I was struggling on how to move forward with my dissertation and opening up your lab to me whenever I needed it.

Hollis, Sharon, and Alejandra, you three showed me why I chose this path and I thank the three of you for spending more hours in the lab than anyone should. I hope each of you learned as much from me as I learned from y'all. My other labmates made the journey worthwhile and fun and I would like to thank Greg Territo, Allyson Fenwick (and Will and Aidan), Haakon Kalkvick, Jason Hickson, Matt Lawrance, Lindsay Arick, Ramin Beheshti, Rachel Acuna, Andrew Mason (and Nia), Rhett Rautsaw, Gina Ferrie, Vicki Villanova, Shelly Gaynor, Alexa Trujillo, Katie Mercier, Matt Holding, Mark Dimeo, Erich Hofmann, Bridget Vincent, Katelyn Lanctot, Carly Grimison, Stephanie Pruneau, Samuel Solano, Millie Baylac, and so many others that I got to learn from in the many lab meetings I attended and discussions we had. I cannot thank the other students I met at UCF enough for their support and camaraderie. The students and the Biology Graduate Student Association are the strength of the department and I am glad I made so many friends at UCF.

I would not have made it through without the help of many people in the UCF family. Patrick Larabee, Steve Jacobs, and Raymond Cload went above and beyond representing UCFIT. They helped me install programs, answered questions, replaced the three hard drives that gave their lives to this dissertation, and fixed my computer more times than I can count. The team that run the UCF STOKES supercomputer including D. Greenspan, N. Lucas, J. Schnaitter, and P. Wiegand were instrumental in helping me complete the analyses for my dissertation. Robert "Bob" Banks and Bryan Kirk at the ADRF animal facility put up with me and my lab mates constantly bringing snakes in and out for our research projects. Bob never hesitated to help us whenever needed and I really appreciate all the advice and assistance he provided. I would especially like to thank Melissa Dagley and the EXCEL/COMPASS/iSTEM team including Sarah Evans, Renee Sackett, Alex Davila, and all of the GTAs I have been fortunate to work with as part of the program. Being a part of that team has been very influential and I will take these experiences with me moving forward.

My dissertation would not have been possible without the people I met along the way. The people that helped me at Florida State University including M. Hogan, M. Margres, J. McGivern, and M. Ward from the Rokyta Lab, M. Seavy and B. Washburn (FSU Department of Biological Science Analytical Lab), and S. Miller (FSU DNA Sequencing Facility) were collegial and always willing to help. I am very happy to now count Miguel Borja, Juan Castañeda-Gaytán, and Gamaliel Castañeda-Gaytán as collaborators and life-long friends. So many people helped me in the field and there are too many to name but I want to specifically thank Jeff Adams and SnakeDays, Diego Ortiz, Dave Weber, and the Arizona Herpetological Association and all the herpers I got to meet, talk to, and find snakes with. A special thank you to Ron Govreau and Kent VanSooy who were taken too soon but left a lasting impression on me and many others with their kindness and willingness to help.

Finally, I would like to thank my wife, Stacy, and the Strickland and Lee families for always supporting me.

TABLE OF CONTENTS

LIST OF FIGURES	xi
LIST OF TABLES	xix
CHAPTER 1: INTRODUCTION	1
Study species	3
Study Aim	5
References	5
CHAPTER 2: PHYLOGEOGRAPHY AND ENVIRONMENTAL VARIABLES POORLY EXPLAIN DIVERSE VENOM PHENOTYPES IN MOJAVE RATTLESNAKES	10
Abstract	10
Introduction	11
Materials and methods	15
Ethics statement	15
Sample collection and DNA extraction	15
Reverse-phased high performance liquid chromatography to determine venom type	16
Mojave toxin assay	16
Venom phylogeography of <i>C. scutulatus</i>	17
Azocasein metalloproteinase assay	19
Kallikrein-like serine protease assay	19
SDS-PAGE protein gel electrophoresis	19
Head morphology analysis	20
Niche modeling of venom type	21

Data availability	22
Results	22
Venom type in <i>C. scutulatus</i>	22
<i>Crotalus scutulatus</i> phylogeography	23
Azocasein metalloproteinase assay	23
Kallikrein-like serine protease assay	24
SDS-PAGE	26
Morphological analysis	27
ENMs between venom type	27
Discussion	30
Conclusion	37
Acknowledgments	37
References	38

CHAPTER 3: PHENOTYPIC VARIATION IN MOJAVE RATTLESNAKE (*CROTALUS*

SCUTULATUS) VENOM IS DRIVEN BY FOUR TOXIN FAMILIES 48

Abstract	48
Introduction	49
Materials and methods	53
Ethics statement	53
Sample collection	53
Venom type determination	54
Venom gland transcriptome sequencing	55
Transcriptome assembly	58
Expression analysis	59
PLA ₂ diversity	60

Data availability	61
Results	61
Venom gland transcriptomes of <i>C. scutulatus</i>	61
Toxin diversity in <i>C. scutulatus</i>	62
Myotoxin <i>a</i> diversity	62
Expression differences in Type A and Type B <i>C. scutulatus</i>	67
PLA ₂ diversity	69
Discussion	72
Conclusions	76
Acknowledgments	77
References	78

CHAPTER 4: ADDITIVE EXPRESSION SUGGESTS MENDELIAN INHERITANCE OF
POLYMORPHIC VENOM PHENOTYPES IN MOJAVE RATTLESNAKES
(*CROTALUS SCUTULATUS*) 86

Abstract	86
Introduction	87
Materials and methods	90
Ethics Statement	90
Sampling and venom type determination	91
Venom gland transcriptome sequencing and assembly	92
Gene expression patterns among venom types	94
Differential expression in Type A + B individuals	94
PLA ₂ and SVMP diversity and inheritance testing	96
Data availability	97
Results	97

Consensus transcriptome of <i>Crotalus scutulatus</i>	97
Venom type expression patterns	98
Differential expression among and between venom types	101
Snake venom metalloproteinase and phospholipase A ₂ s in Type A + B individuals .	105
Inheritance pattern of SVMPs and PLA ₂ s in <i>C. scutulatus</i>	107
Discussion	109
Conclusions	112
Acknowledgments	113
References	114
 CHAPTER 5: CONCLUSION	 124
Significance	126
References	128
 APPENDIX A: CHAPTER THREE LICENSE TERMS AND CONDITIONS	 131
 APPENDIX B: CHAPTER TWO SUPPLEMENTAL TABLE	 134
 APPENDIX C: CHAPTER FOUR SUPPLEMENTAL TABLE	 141

LIST OF FIGURES

Figure 1.1: Distribution map of Mojave Rattlesnakes, <i>Crotalus scutulatus</i>	4
Figure 2.1: Distribution and sampling of Mojave Rattlesnakes collected from throughout their range. Red, purple, and blue represent Type A, Type A + B, and Type B venom, respectively, in the pie charts and the sampling points. Mottling in the distribution are areas of gene flow between lineages. Pie charts represent the proportion of each venom type collected from each lineage. Cladeogram in lower left of the four mitochondrial lineages from Figure 2.3 with the three nuclear clades from Schield <i>et al.</i> [40] numbered: 1 - Sonoran lineage, 2 - Chihuahuan lineage, 3 - Central Mexican Plateau lineage, ? - not sampled/unknown lineage	14
Figure 2.2: Representative reverse-phased high performance liquid chromatography (RP-HPLC) profiles of Type A (top), Type A + B (middle), and Type B (bottom) venom of Mojave Rattlesnakes. The acidic (α) and basic subunit (β) peaks for Mojave toxin are marked and the region where snake venom metalloproteinases elute is marked with a blue bar.	24
Figure 2.3: Bayesian inference phylogeny based on ND4 sequence from 189 <i>C. scutulatus</i> . The dashed line indicates where the Sonoran and Mojave Desert lineage was moved from and no size adjustments occurred. Venom type as discrete characters and metalloproteinase activity on a continuous scale are mapped onto the phylogeny. Individuals without activity values are gray boxes. Dots on nodes represent significant posterior probability values of ≥ 0.95	25

Figure 2.4: Metalloproteinase activity levels for venom types within each lineage. Type A venom had significantly less metalloproteinase activity regardless of lineage. Type B and Type A + B were not significantly from each other in any comparison. 26

Figure 2.5: Kallikrein-like serine protease activity between the Chihuahuan and Sonoran lineages. Type B venom had significantly lower activity than Type A venom but there were no differences in the same venom type between the two populations and Type A + B venom was not significantly different than Type A or Type B venom. 27

Figure 2.6: SDS-PAGE gel images for 59 of 110 samples with toxin families labeled. Venom type of each individual is added to the end of the sample name. When present, the basic subunit of Mojave toxin is at approximately 14kD. See Figure 2.7 for remaining sample images. 28

Figure 2.7: SDS-PAGE gel images for 51 of 110 samples with toxin families labeled. Venom type of each individual is added to the end of the sample name. When present, the basic subunit of Mojave toxin is at approximately 14kD. See Figure 2.6 for remaining sample images. 29

Figure 2.8: Comparison of interfang distance (IF) on the left and fang length (FL) on the right in the Sonoran lineage of Mojave Rattlesnakes. Snout-vent length (SVL) was used to control for different sizes among animals. Smaller values on the y-axis are larger measurements. There was a trend of Type A individuals having longer fangs but it was not significantly different from Type B or Type A + B individuals. Type A individuals did have significantly wider distances between their fangs compared to Type B individuals but not compared to Type A + B individuals. 30

Figure 2.9: Ecological niche models generated in MAXENT using 64 Type A (left) and 31 Type B (right) Mojave Rattlesnakes scaled by probability of presence (pp). Lower maps display model distributions where each venom type is expected to occur based on a threshold point where model sensitivity and specificity are highest ($pp > 0.22$ for Type A, $pp > 0.07$ for Type B). 31

Figure 2.10: Null distributions generated to test niche equivalency using 99 permutations. Results for both Schoener's D and Warren's I reject the null hypotheses that the models for Type A and Type B are identical. 32

Figure 2.11: Null distributions generated using 99 permutations of Type A and Type B individuals to test niche similarity. The top row compares Type A individuals using the model background (bg) of Type B and the bottom row compares Type B individuals using the model background of Type A. Results for both Schoener's D and Warren's I reject the null hypothesis that the two models do not predict the occurrence of each other better or worse than would be expected by chance. Because the D and I values fall in the right tail of the distribution, the niches are more similar than would be expected by chance. 32

Figure 2.12: Model response to variables included in the ecological niche models for Type A and Type B venoms. Asterisks (*) indicate variables that were significantly different between the two models. BIO1 = Annual Mean Temperature, BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp)), BIO4 = Temperature Seasonality (standard deviation *100), BIO6 = Min Temperature of Coldest Month, BIO8 = Mean Temperature of Wettest Quarter, BIO9 = Mean Temperature of Driest Quarter, BIO12 = Annual Precipitation, BIO14 = Precipitation of Driest Month, BIO15 = Precipitation Seasonality (Coefficient of Variation), BIO18 = Precipitation of Warmest Quarter, BIO19 = Precipitation of Coldest Quarter. 33

Figure 3.1: Distribution of Type A and Type B venom in Arizona based on data from Wilkinson et al. (triangles) and the nine transcriptome animals sequenced (stars). The shaded area is the estimated distribution of Type B venom and the dark black line is the outline of the estimated distribution of *Crotalus scutulatus*. CLP: Christopher L. Parkinson field number. 52

Figure 3.2: RP-HPLC profiles for the nine specimens of *C. scutulatus* selected for transcriptome sequencing. Type A individuals are in the left column and Type B individuals are in the right column. α = acidic subunit of Mojave toxin; β = basic subunit of Mojave toxin; Blue bar = region where SVMPS elute. 56

- Figure 3.3: Representation of the toxins in the venom gland transcriptome for nine *C. scutulatus* specimens. Type A are on the left and Type B are on the right, and are ranked by the increasing amount of myotoxins. The bar graphs represent each of the 75 toxins identified. Any toxin that does not have a black bar under it did not meet the criteria for presence in the transcriptome. The pie charts represent the proportion of each toxin family in the venom gland transcriptome. 65
- Figure 3.4: Representation of the average Type A and Type B venom gland transcriptome of *C. scutulatus* from the Sonoran Desert. These were generated by averaging the TPM values of each individual within a venom type. For both Type A and Type B, the majority of the highly expressed transcripts are toxins. The bar graphs with colors represent each of the 75 toxins identified. Any toxin that does not have the black bar under it did not meet the criteria for being present in the transcriptome of any individual of that venom type. The pie charts represent the proportion of each toxin family in the venom gland transcriptome. Type A individuals have very few SVMPs and the Type B individuals are lacking neurotoxic phospholipase A₂ (PLA₂), Mojave toxin. 66
- Figure 3.5: Average expression of toxin families with respect to Type A and Type B *C. scutulatus* after centered log-ratio (clr) transform. Error bars are the standard deviation around the mean of the clr-transformed data. 68

Figure 3.6: Pairwise comparison of the average Type A and Type B venom gland transcriptomes using the centered log-ratio (clr)-transformed TPM data. SVSPs and toxin families that only have one toxin are not labeled. SVSPs did not differ between the two transcriptomes. The red line is the line of best fit through the non toxins and the dashed black lines are the 99% confidence around that line. Any transcript outside the dashed black lines was identified as an outlier. Anything above the upper line is overexpressed in Type B and anything below the lower line is overexpressed in Type A. 69

Figure 4.1: Reverse-phased high performance liquid chromatography (RP-HPLC) profiles of Mojave Rattlesnakes. The acidic (α) and basic subunit (β) peaks for Mojave toxin are marked and the region where snake venom metalloproteinases elute is marked with a blue bar. 99

Figure 4.2: Representation of the toxins in the venom gland transcriptome for six *C. scutulatus* added in this study. The bar graphs represent each of the 84 toxins identified. The pie charts represent the proportion of each toxin family in the venom gland transcriptome. 3FTX: three-finger toxin; BPP: bradykinin potentiating peptide; CRISP: cysteine-rich secretory protein; CTL: C-type lectin; HYAL: hyaluronidase; KUN: Kunitz peptide; LAAO: L-amino-acid oxidase; MYO: myotoxin; NGF: nerve growth factor; NUC: 5' nucleotidase; PDE: phosphodiesterase; PLA₂: Phospholipase A₂; PLB: phospholipase B; SVMP: snake venom metalloproteinase; SVSP: snake venom serine protease; VEGF: vascular endothelial growth factor; MTX: Mojave toxin. 100

- Figure 4.3: Representation of the average Type A + B transcriptome of the four *C. scutulatus* with Type A + B venom. The majority of highly expressed transcripts were the 84 toxins identified. Pie charts represent the proportion of each of the 19 toxin families identified. 101
- Figure 4.4: Hierarchical clustered heatmap of all 84 toxins from the 15 individuals of *C. scutulatus* with their venom gland transcriptome sequenced based on ln-transformed TPM data. Venom type is designated for each individual and dark colors are low expression and lighter colors are higher expression. The putative genotype for each individual is indicated below venom type. . 102
- Figure 4.5: Principle component analysis (PCA) scatterplot of the 15 *C. scutulatus* from the Sonoran Desert based on TPM data for 84 toxins. The squares are the average value for each venom type and the circles are the 95% confidence ellipse for each venom type. The first two axes explain 49.97% of the variability. PCA 1 represents the Type A (left) and Type B (right) dichotomy and PCA 2 represents the presence (bottom) and absence (top) of myotoxins in the venom gland transcriptomes. 103
- Figure 4.6: Graphical representation of differential expression based on the four methods used: DeSEQ2, edgeR, EBseq, and RNentropy. Arrow direction goes from relatively low expression to relatively high expression for each of the possible pairwise comparison between the three venom types in *C. scutulatus*. Numbers indicate number of toxins out of 84 that are differential expressed in each comparison. 104

- Figure 4.7: Pairwise comparison of average Type A and Type A + B venom gland transcriptomes using the centered log-ratio (clr)-transformed TPM data. The red line is the line of best fit through the non toxins and the dashed black lines are the 99% confidence around that line. Anything above the upper line is overexpressed in Type A + B and anything below the lower line is overexpressed in Type A. 105
- Figure 4.8: Pairwise comparison of average Type B and Type A + B venom gland transcriptomes using the centered log-ratio (clr)-transformed TPM data. The red line is the line of best fit through the non toxins and the dashed black lines are the 99% confidence around that line. Anything above the upper line is overexpressed in Type A + B and anything below the lower line is overexpressed in Type B. 106
- Figure 4.9: Principle component analysis (PCA) scatterplot of the 15 *C. scutulatus* from the Sonoran Desert based on TPM data for the four PLA₂s and four SVMPs in Table 4.2. The squares are the average value for each putative genotype and the circles are the 95% confidence ellipse for genotypes with more than two individuals. The first two axes explain 94.78% of the variability. PCA 1 represents MTX presence/SVMP absence (left) and MTX absence/SVMP presence (right) and PCA 2 is the variation of expression of Pla2-gK and Pla2-gB1. 109
- Figure 5.1: Distribution map of Mojave Rattlesnakes, *Crotalus scutulatus*, with venom type labeled with dots and pie charts used to represent the proportion of toxin family in the venom-gland transcriptome of 27 individuals. 127

LIST OF TABLES

Table 2.1:	Logistic regression comparison of the 14 variables used in the Type A and Type B models for <i>C. scutulatus</i> . Bolded variables were significantly different between the two models.	33
Table 3.1:	Sample information for the nine specimens of <i>C. scutulatus</i> sequenced from the U.S. Sonoran Desert. ASNHC = Angelo State Natural History Collection, San Angelo, Texas; ASU = Arizona State University Natural History Museum, Tempe, Arizona; CLP = Christopher L. Parkinson field number.	54
Table 3.2:	Transcripts per million reads (TPM) values for 75 toxins identified in the nine <i>C. scutulatus</i> individuals. TPM values were generated in RSEM [56] with Bowtie2 [53]. CRISP: cysteine-rich secretory protein; BPP: bradykinin potentiating peptide; CTL: C-type lectin; HYAL: hyaluronidase; KUN: Kunitz peptide; LAAO: L-amino-acid oxidase; MYO: myotoxin; NGF: nerve growth factor; NUC: 5' nucleotidase; PDE: phosphodiesterase; PLB: phospholipase B; SVMP: snake venom metalloproteinase; SVSP: snake venom serine protease; VEGF: vascular endothelial growth factor.	63

Table 3.3: Presence and absence data for toxin transcripts that were not found in all individuals. Toxins highlighted in dark blue were found in all Type B individuals but never in Type A individuals. Toxins highlighted in dark red were found in all Type A individuals but never in Type B individuals. Toxins highlighted in light blue or light red were only found in individuals of that venom type but were not found in all individuals. The last row is the number (out of 75) of toxins present in total, which includes 33 toxins present in all individuals not listed in this table. To be present in the transcriptome, toxins had to have at least $5\times$ coverage over 90% of the transcript. MTXA: acidic subunit of Mojave toxin; MTXB: basic subunit of Mojave toxin. 64

Table 3.4: Differential expression analyses for toxins between the nine Type A and Type B *C. scutulatus* as well as the average Type A (AveA) and Type B (AveB) transcriptomes. The UpB and UpA count data were generated by identifying outlier transcripts in the pairwise comparisons of the Type A and Type B individuals (maximum of 20 comparisons). The last four columns are the data from DESeq and DESeq2 identifying differential expression between the two venom types. Toxins highlighted in dark red were found in all Type A individuals but never in Type B individuals. Toxins highlighted in light blue or light red were exclusively found in individuals of that venom type but were not found in all individuals. Toxins highlighted in green were found in all nine individuals. Toxins with NA for P_{adj} only had one individual in one of the treatments with the toxin so it was not possible to calculate in DESeq2. 70

Table 3.5: Presence or absence data for PLA₂s identified by Dowell *et al.* [7] (the first five PLA₂s) and the sixth PLA₂ identified in this study. The first four individuals were the specimen used in Dowell *et al.* [7] and only presence/absence is indicated. The last nine individuals are *C. scutulatus* specimens sequenced in this study with venom type indicated and transcripts per million reads (TPM) values given when that PLA₂ is present. The one TPM value denoted by an * had eight nonsynonymous nucleotide changes in the sequence compared to the other three *C. scutulatus* specimens and matched that of *Crotalus viridis* (Genbank accession AF403134). 71

Table 4.1: Sample information for the six *C. scutulatus* sequenced from the U.S. Sonoran Desert. ASNHC = Angelo State Natural History Collection, San Angelo, TX; ASU = Arizona State University Natural History Museum, Tempe, Arizona; CLP = Christopher L. Parkinson Field Number 91

Table 4.2: Transcripts per million reads (TPM) values rounded to the closest integer for four PLA₂ and four SVMP loci identified as being unique to the genome of Type A and Type B *Crotalus scutulatus* [24, 32]. Values in **bold** met the criteria for presence and values in *italics* did not. Heterozygous individuals for PLA₂s express all four PLA₂s. Heterozygous individuals (CLP1823, CLP2111) at the SVMP loci were putatively identified based on relative expression of the four most highly expressed SVMP loci in the region present in Type B: mp238, mp237, mp240, mp242. The putative genotype is listed with the following code: AA–presence of MTX/absence of Pla2-gK and Pla2-gB1, Aa–expression of all four PLA₂s, aa–absence of MTX/presence of Pla2-gK and Pla2-gB1, BB–high expression of SVMPs, Bb–approximately half of the average expression of SVMPs, bb–no expression of SVMPs. 108

Table B.1: Specimen information and associated data used for analyses in Chapter 2. . 135

Table C.1: Differential toxin expression output for the four methods used. Genes not shown were not differentially expressed in any analysis. For the log-fold changes, anything that is negative was higher in the first venom type being compared. 142

CHAPTER 1: INTRODUCTION

Natural selection can cause divergence between separate lineages when those lineages are exposed to different environmental selection pressures. Natural selection pressures from abiotic factors in the environment and biotic factors such as diet and predation lead to phenotypic traits that should be locally adaptive [1]. If heritable phenotypic variation exists and selective forces are strong on specific traits within populations, phenotypic divergence may occur [2]. By examining diverging phenotypes, it is possible to understand the role which natural selection has played in maintaining phenotypic diversity. This trait evolution framework is particularly useful when the phenotype being studied is highly variable, composed of many simple components, and easily tied to the genome [3]. Comparing the different components within a composite phenotype can lead to identifying adaptive traits and candidate loci that are linked to the genotype yielding a genotype-phenotype map for understanding the molecular basis of adaptation [4, 5]. Generating a genotype-phenotype maps is extremely difficult because phenotypes are generally influenced by multiple genes of unknown origin. One way to bridge the gap between genotype and phenotype is to compare tissue specific transcriptomes [3].

Transcriptome sequencing provides a snapshot in time of what genes are being transcribed and the relative abundance of each transcript in a specific tissue [6]. The transcriptome is the functional backbone of the genome, and through differential expression, can alter the phenotype without changes in the genetic architecture. By only sequencing the genes expressed in a single tissue, it is possible to focus in on specific transcripts of interest that are under high selection or are driving the variability observed in a phenotype [7]. Comparative transcriptomics investigates differential gene expression in conjunction with transcript sequence comparison. Using this approach on individuals with different phenotypes allows for transcriptome-wide scans to locate transcripts with high rates of change in the sequence and/or transcripts that are being over- or under-expressed [3]. Ideally, when using this technique to examine evolutionary questions

in regard to selection, the variation in the phenotype and the transcriptome need to be easily connected. This has been clearly demonstrated in studying venom and venom gland transcriptomes in many different organisms [8–10].

Venom has been of tremendous interest to ecology, evolution, and medicine, particularly in the last 20 years following the advancement of proteomic and sequencing technologies [11]. It is loosely defined as a secretion produced in a gland that is delivered to another organism through a wound. More specifically, venom is a composite phenotype composed of many different proteins that are variable at all biological levels [12, 13]. This complex mixture of toxins (proteins and peptides), salts, and organic compounds that, after introduction, work to disrupt physiological systems often causing pain, death, or debilitation [14, 15]. Venom has convergently evolved in at least seven animal phyla and is found in over 100,000 species.

Whole venom gland transcriptome sequencing has been conducted on many venomous taxa and has resulted in the discovery of new proteins as well as a more thorough understanding of the processes that generate the variation documented in venom [15, 16]. However, the majority of venom gland transcriptomes sequenced have been characterized in snakes. For nearly all studies looking at venom gland transcriptomes, a single individual's mRNA has been used to represent an entire species. This ignores the tremendous amount of variation that has been documented within species [17]. This is especially apparent in venomous snake species where thorough studies of inter- and intraspecific venom variation have been conducted [15].

Within venomous snakes, venom is used for prey capture and/or defense and is thought to be under strong selection [18–20]. There are generally less than 200 unique proteins in venom that fall within 15 - 20 protein families [4]. Most venom studies in snakes have focused on rattlesnakes in the genus *Crotalus*. For rattlesnakes, there is a general dichotomy of snake venoms based on the presence (Type A) or absence (Type B) of Mojave toxin and homologs [21, 22]. Mojave toxin (MTX) is a heterodimeric phospholipase A₂ (PLA₂) which is a presynaptic acting neurotoxin composed of an acidic and basic subunit [23]. Mojave toxin is homologous with toxins

that have been identified in other rattlesnake species including Crotoxin (*C. durissus*), Canebrake toxin (*C. horridus*) and toxins found in several other rattlesnake species [24, 25]. In addition to the differences in neurotoxic activity between venom types, there are other associated differences between Type A and Type B venoms.

Type B venoms have high metalloproteinase activity and relatively low toxicity and Type A venoms have low metalloproteinase activity and high toxicity [26]. Snake venom metalloproteinases (SVMPs) are responsible for the high hemorrhagic activity. The majority of rattlesnakes are Type B, a few are Type A, but there are some interesting and unique cases where the Type A/B dichotomy exists among populations within a species [27]. This has been found in *C. oreganus*, *C. durissus*, *C. horridus*, and *C. scutulatus* [27, 28]. It is best documented in *C. scutulatus* where there is a relatively well defined range for the populations with Type A and Type B venom. The evolutionary history of populations and species exhibiting this dichotomy is unknown but differences in feeding ecology [26] and pedomorphism [28] have been hypothesized.

Study species

Mojave Rattlesnakes, *Crotalus scutulatus*, are found in deserts of the southwestern U.S. and Mexico and are well known by the general public and venom researchers due to their reputation for having toxic and complex venom (Figure 1.1). *Crotalus scutulatus* are a medium sized rattlesnake found predominately in desert grasslands including those dominated by creosote and/or mesquite [29, 30]. This species has very toxic venom, particularly when compared to many other *Crotalus* species, and the severity of snakebites in this species resulted in considerable venom research that began in the 1930s [31].

As research progressed in *C. scutulatus*, geographic variation in venom was identified. It was thought that *C. scutulatus* had Type A venom throughout its range (MTX present/high toxicity/low metalloproteinase activity) which was considered atypical for rattlesnakes [17, 27].

Differences in snakebite pathologies led to the discovery of a population of *C. scutulatus* in Arizona which possess Type B venom, more typically observed in rattlesnakes in that there is no MTX present, low toxicity, and high metalloproteinase activity [32–36]. Glenn and Straight [37] and Wilkinson *et al.* [38] sampled extensively in Arizona and attempted to identify the geographic extent of *C. scutulatus* populations with each venom type. Through this work, they determined that the type B population was confined to central Arizona between Phoenix and Tucson, Arizona. They also identified individuals with a mixed venom called Type A + B in areas where the two populations came into contact. Due to the observed variation in venom protein expression within *C. scutulatus* it makes this species an ideal model to test the role that natural selection has had on venom components.

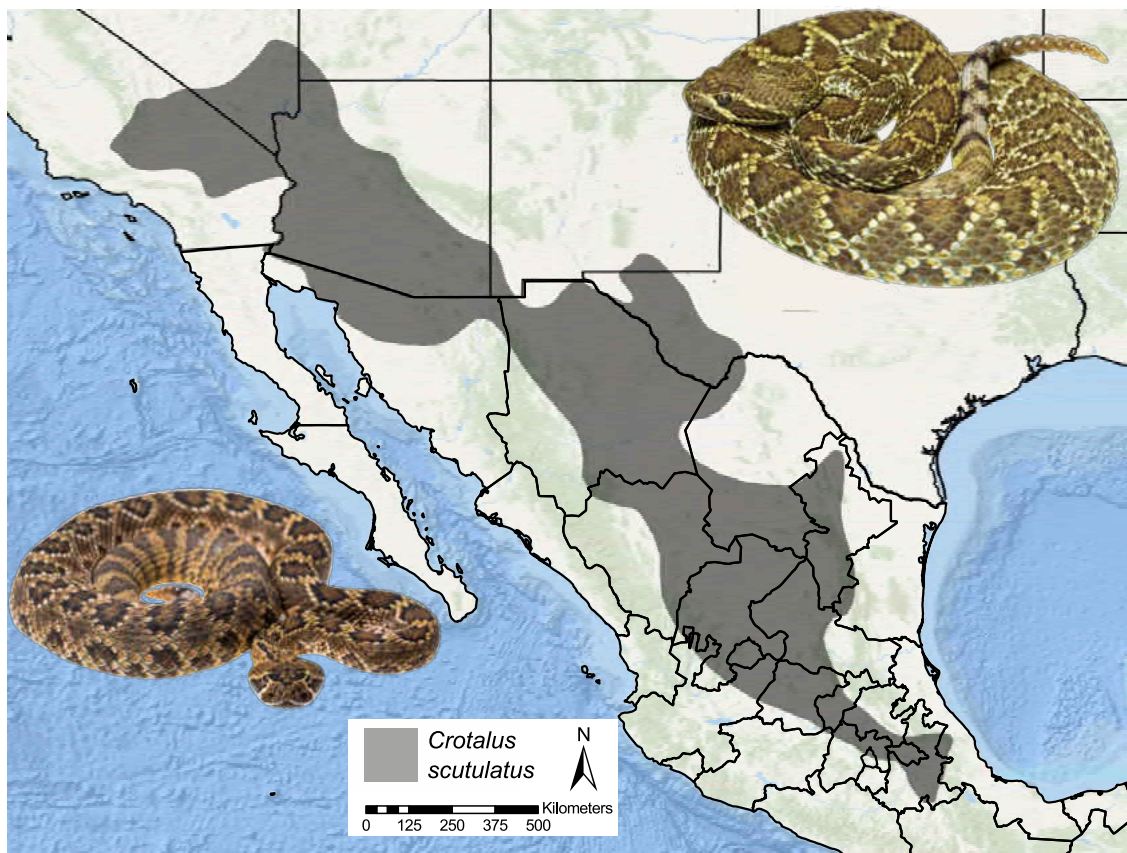


Figure 1.1: Distribution map of Mojave Rattlesnakes, *Crotalus scutulatus*.

Study Aim

To understand the drivers of phenotypic diversity in the venom of Mojave Rattlesnakes, *Crotalus scutulatus*, I aim to meet three goals.

1. Document the distribution of venom phenotypes throughout the range of *C. scutulatus* and determine if it corresponds to ecological differences including climate and functional morphology or if it corresponds to phylogeographic structure.
2. Identify the toxin loci involved in the expression of the venom phenotype dichotomy through comparative transcriptomics of the venom-gland.
3. Determine the inheritance mechanism for the venom dichotomy by testing for additive expression in the Type A + B phenotype.

References

1. Kawecki, T.J.; Ebert, D. Conceptual issues in local adaptation. *Ecology Letters* **2004**, *7*, 1225–1241.
2. Andrés, J.A.; Larson, E.L.; Bogdanowicz, S.M.; Harrison, R.G. Patterns of transcriptome divergence in the male accessory gland of two closely related species of field crickets. *Genetics* **2013**, *193*, 501–13.
3. Harrison, P.W.; Wright, A.E.; Mank, J.E. The evolution of gene expression and the transcriptome-phenotype relationship. *Seminars in Cell & Developmental Biology* **2012**, *23*, 222–229.
4. Gibbs, H.L.; Sanz, L.; Calvete, J.J. Snake population venomomics: proteomics-based analyses of individual variation reveals significant gene regulation effects on venom protein expression in *Sistrurus* rattlesnakes. *Journal of Molecular Evolution* **2009**, *68*, 113–125.

5. Gibbs, H.L.; Chiucchi, J.E. Deconstructing a complex molecular phenotype: population-level variation in individual venom proteins in Eastern Massasauga Rattlesnakes (*Sistrurus c. catenatus*). *Journal of Molecular Evolution* **2011**, *72*, 383–397.
6. Winter, E.E.; Goodstadt, L.; Ponting, C.P. Elevated rates of protein secretion, evolution, and disease among tissue-specific genes. *Genome Research* **2004**, *14*, 54–61.
7. Barreto, F.S.; Moy, G.W.; Burton, R.S. Interpopulation patterns of divergence and selection across the transcriptome of the copepod *Tigriopus californicus*. *Molecular Ecology* **2011**, *20*, 560–72.
8. Rodrigues, R.S.; Boldrini-França, J.; Fonseca, F.P.P.; de la Torre, P.; Henrique-Silva, F.; Sanz, L.; Calvete, J.J.; Rodrigues, V.M. Combined snake venomomics and venom gland transcriptomic analysis of *Bothropoides pauloensis*. *Journal of Proteomics* **2012**, *75*, 2707–2720.
9. Lavergne, V.; Dutertre, S.; Jin, A.h.; Lewis, R.J.; Taft, R.J.; Alewood, P.F. Systematic interrogation of the *Conus marmoreus* venom duct transcriptome with ConoSorter reveals 158 novel conotoxins and 13 new gene superfamilies. *BMC Genomics* **2013**, *14*, 708.
10. Li, J.J.; Bickel, P.J.; Biggin, M.D. System wide analyses have underestimated protein abundances and the importance of transcription in mammals. *PeerJ* **2014**, *2*, e270.
11. Calvete, J.J. Venomomics: digging into the evolution of venomous systems and learning to twist nature to fight pathology. *Journal of Proteomics* **2009**, *72*, 121–6.
12. Mebs, D. Toxicity in animals. Trends in evolution? *Toxicon* **2001**, *39*, 87–96.
13. Casewell, N.R.; Harrison, R.A.; Wüster, W.; Wagstaff, S.C. Comparative venom gland transcriptome surveys of the saw-scaled vipers (Viperidae: *Echis*) reveal substantial intra-family gene diversity and novel venom transcripts. *BMC Genomics* **2009**, *10*, 564.
14. Doley, R.; Kini, R.M. Protein complexes in snake venom. *Cellular and Molecular Life Sciences* **2009**, *66*, 2851–2871.
15. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex cocktails: the

- evolutionary novelty of venoms. *Trends in Ecology & Evolution* **2013**, 28, 219–29.
16. Fox, J.W.; Serrano, S.M.T. Exploring snake venom proteomes: multifaceted analyses for complex toxin mixtures. *Proteomics* **2008**, 8, 909–20.
 17. Chippaux, J.P.; Williams, V.; White, J. Snake venom variability: methods of study. *Toxicon* **1991**, 29, 1279–1303.
 18. Calvete, J.J.; Sanz, L.; Angulo, Y.; Lomonte, B.; Gutiérrez, J.M. Venoms, venomics, antivenomics. *FEBS Letters* **2009**, 583, 1736–1743.
 19. Daltry, J.C.; Wüster, W.; Thorpe, R.S. Diet and snake venom evolution. *Nature* **1996**, 379, 537–540.
 20. Barlow, A.; Pook, C.E.; Harrison, R.A.; Wüster, W. Coevolution of diet and prey-specific venom activity supports the role of selection in snake venom evolution. *Proceedings of the Royal Society B: Biological Sciences* **2009**, 276, 2443–2449.
 21. Bieber, A.L.; Tu, T.; Tu, A.T. Studies of an acidic cardiotoxin isolated from the venom of Mojave Rattlesnake (*Crotalus scutulatus*). *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1975**, 400, 178–188.
 22. Rael, E.D.; Knight, R.; Zepeda, H. Electrophoretic variants of Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) venoms and migration differences of Mojave toxin. *Toxicon* **1984**, 22, 980–984.
 23. Wooldridge, B.; Pineda, G.; Banuelas-Ornelas, J.; Dagda, R.; Gasanov, S.; Rael, E.; Lieb, C. Mojave Rattlesnakes (*Crotalus scutulatus scutulatus*) lacking the acidic subunit DNA sequence lack Mojave toxin in their venom. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **2001**, 130, 169–179.
 24. Weinstein, S.A.; Minton, S.A.; Wilde, C.E. The distribution among ophidian venoms of a toxin isolated from the venom of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*). *Toxicon* **1985**, 23, 825–844.
 25. Rokyta, D.R.; Wray, K.P.; Margres, M.J. The genesis of an exceptionally lethal venom

- in the Timber Rattlesnake (*Crotalus horridus*) revealed through comparative venom-gland transcriptomics. *BMC Genomics* **2013**, *14*, 394.
26. Mackessy, S.P. *Handbook of venoms and toxins of reptiles*; CRC Press/Taylor & Francis, 2010; p. 521.
 27. Mackessy, S.P. Venom composition in rattlesnakes: trends and biological significance. In *The Biology of Rattlesnakes*; Hayes, W.K.; Beaman, K.R.; Cardwell, M.D.; Bush, S.P., Eds.; Loma Linda University Press: Loma Linda, CA, 2008; pp. 495–510.
 28. Calvete, J.J.; Pérez, A.; Lomonte, B.; Sánchez, E.E.; Sanz, L. Snake venomomics of *Crotalus tigris*: the minimalist toxin arsenal of the deadliest nearctic rattlesnake venom. Evolutionary clues for generating a pan-specific antivenom against crotalid type II venoms. *Journal of Proteome Research* **2012**, *11*, 1382–1390.
 29. Campbell, J.A.; Lamar, W.W. *The Venomous Reptiles of the Western Hemisphere*; Comstock Pub. Associates, 2004; p. 870.
 30. Cardwell, M.D. Mohave Rattlesnakes *Crotalus scutulatus* (Kennicott 1861). In *Rattlesnakes of Arizona Species Accounts and Natural History, 1*; Schuett, G.W.; Feldner, M.; Smith, C.; Reiserer, R., Eds.; ECO Herpetological Publishing: Rodeo, NM, 2016; pp. 563–605.
 31. Githens, T.; George, I. Comparative studies on the venoms of certain rattlesnakes. *Bulletin of the Antivenin Institute of America* **1931**, *5*, 31–34.
 32. Glenn, J.; Straight, R. Mojave Rattlesnake *Crotalus scutulatus scutulatus* venom: variation in toxicity with geographical origin. *Toxicon* **1978**, *16*, 81–84.
 33. Glenn, J.J.L.; Straight, R.C.; Wolfe, M.M.C.; Hardy, D.D.L. Geographical variation in *Crotalus scutulatus scutulatus* (Mojave Rattlesnake) venom properties. *Toxicon* **1983**, *21*, 119–130.
 34. Hardy, D.L. Envenomation by the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in southern Arizona, U.S.A. *Toxicon* **1983**, *21*, 111–118.

35. Martinez, M.; Rael, E.D.; Maddux, N.L. Isolation of a hemorrhagic toxin from Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) venom. *Toxicon* **1990**, 28, 685–694.
36. Sánchez, E.E.; Galán, J.A.; Powell, R.L.; Reyes, S.R.; Soto, J.G.; Russell, W.K.; Russell, D.H.; Pérez, J.C. Disintegrin, hemorrhagic, and proteolytic activities of Mohave Rattlesnake, *Crotalus scutulatus scutulatus* venoms lacking Mojave toxin. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **2005**, 141, 124–132.
37. Glenn, J.L.; Straight, R.C. Intergradation of two different venom populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Toxicon* **1989**, 27, 411–8.
38. Wilkinson, J.A.; Glenn, J.L.; Straight, R.C.; Sites, J.W. Distribution and genetic variation in venom A and B populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Herpetologica* **1991**, 47, 54 – 68.

CHAPTER 2: PHYLOGEOGRAPHY AND ENVIRONMENTAL VARIABLES POORLY EXPLAIN DIVERSE VENOM PHENOTYPES IN MOJAVE RATTLESNAKES¹

Abstract

Local adaptation tends to move phenotypic variation to the optimal fitness strategy, but polymorphisms can persist even under strong selection pressures. Polymorphism can be maintained by lack of gene flow between polymorphic populations, differential fitness between environments, or through balancing selection. Balancing selection has been proposed as the mechanism maintaining venom polymorphism in Mojave Rattlesnakes (*Crotalus scutulatus*), but genetic isolation and environmental differences have not been thoroughly tested. In *C. scutulatus*, Type A individuals have venom with low hemorrhagic activity and high toxicity due to the presence of a presynaptic-acting neurotoxin (Mojave toxin), while Type B individuals tend to have high hemorrhagic activity and low toxicity. Venom is under strong selection due to its ecological role in prey acquisition and defense, but the maintenance of the venom dichotomy is unknown. To determine if the venom phenotypes are genetically isolated or if there are environmental differences driving local adaptations we performed a range-wide sampling of *C. scutulatus* and determined venom type for each sample using a combination of Mojave toxin assay and RP-HPLC. We mapped venom type onto the phylogenetic tree of *C. scutulatus* and created ecological niche models for the two venom types. Overall, venom type was not monophyletic and only one population was fixed for venom type. The ecological niche models between the two venom types showed high niche similarity, despite spatially distinct projections. Only one environmental variable, temperature of the coldest month, explained a significant amount of variation in our model and

¹Chapter Two is being prepared as Strickland et al. Phylogeography and environmental variables poorly explain diverse venom phenotypes in Mojave Rattlesnakes.

was also significantly different between the two models. Phylogeographic, morphological, and environmental differences do not fully explain the geographic structure and venom complexity in *C. scutulatus*. The maintenance of these phenotypes is likely due to balancing selection through intense and fine-scale local adaptation between *C. scutulatus* and prey.

Keywords: Balancing Selection, Ecological Niche Modeling, Functional Morphology, Hemorrhagic Activity, Local Adaptation, Mojave Toxin, Phenotype Polymorphism

Introduction

Local adaptation may facilitate speciation if gene flow becomes sufficiently limited between polymorphic populations, but high gene flow can mitigate this, particularly if selection is weak [1, 2]. Generally, when polymorphisms exist, the most fit phenotype is expected to fix over time [3]. However, there are examples where polymorphisms are maintained in populations, even in the face of high gene flow, through balancing selection [1, 4]. The maintenance of phenotypic polymorphisms through balancing selection is considered relatively rare and transient compared to directional, stabilizing, and disruptive selection [5–9]. However, there are well documented cases of balancing selection including the major compatibility complex, self incompatibility in plants, and butterfly mimicry [7, 10].

Dowell *et al.* [11] hypothesized that a venom dichotomy in Mojave Rattlesnake, (*Crotalus scutulatus*), and numerous other rattlesnake species (*Crotalus* and *Sistrurus*), may be under balancing selection. Individuals possessing a heterodimeric neurotoxic Phospholipase A₂ (PLA₂), such as Mojave Toxin (MTX) in *C. scutulatus*, are Type A. Those without MTX are Type B and tend to have high expression of snake venom metalloproteinases (SVMPs) in their venom. Venom is a highly adaptive trait due to its ecological role in feeding and defense so balancing selection could maintain the two polymorphisms if they confer similar or fluctuating fitness over space or time [12–14]. The geographic arrangement of polymorphisms in a population can provide

evidence for the mechanism or mechanisms of balancing selection maintaining the polymorphism [15]. However, because balancing selection is difficult to identify, we first need to rule out other possible explanations at a range-wide scale.

Several selection-driven hypotheses have been proposed to explain the existence of the dichotomy within and between rattlesnake species because venom is thought to be locally adapted to prey which can occur at a very fine geographic scale [16–18]. Many of the hypotheses are centered around potential differences in digestive efficiency of the two venom types [19]. Individuals with Type B venoms are thought to have more efficient digestion in situations with lower temperatures or when temperature fluctuations are more pronounced such as high elevation [20, 21]. However, tests for an increase in digestive efficiency for Type B venoms have been conflicting [19, 22, 23]. Alternatively, neurotoxic venoms are thought to be advantageous when the chance of escape of prey is high [21]. The Type A venom would rapidly incapacitate the prey item through neuromuscular paralysis and ensure a meal but forgo the possible digestive assistance of the SVMPs [12]. Functional morphological changes could also accompany the different venom types, though to date, only basic external features were tested and none were associated with the venom types [24]. Recently, differences regarding the functional morphology of fangs has been proposed based on the ideal injection depth of venom with different compositions [25]. If digestion is important, Type B venom individuals may have longer fangs to inject venom deeper to aid in digestion whereas Type A venom, because it acts presynaptically, would not require deep fang penetration.

Mojave Rattlesnakes are one of ten rattlesnake species with intraspecific variation in the presence of Type A and Type B venoms [26–30] and, although rare, individuals with both MTX and SVMPs present (Type A + B) do occur [31, 32]. Mojave Rattlesnakes are distributed in the deserts of North America, the Central Mexican Plateau, and the volcanic lowlands of south-central Mexico (Figure 3.1). Because of their broad distribution, Mojave Rattlesnakes occupy habitats encompassing a range of elevations and climatic conditions which could impose

differential selection pressures on venom phenotypes. Both Type A and Type B phenotypes as well as the mixed Type A + B phenotype are known from several different locations throughout their distribution with evidence that the phenotypes are geographically structured [24, 30–37]. Overall, Mojave Rattlesnakes have high intraspecific venom variation documented proteomically [26, 36–38] and transcriptomically [39] but the majority of their distribution has not been sampled.

Two currently described subspecies within *C. scutulatus* exist with clear phylogenetic structure among populations [40]. Schield *et al.* [40] identified three lineages (Sonoran Desert, Chihuahuan Deserts, and Central Mexican Plateau) based on nuclear data. The Central Mexican Plateau lineage also contains a fourth mitochondrial lineage comprised of the subspecies *C. scutulatus salvini* which was not distinct from the rest of the individuals in the Central Plateau based on their RADseq dataset thus rendering the other subspecies, *C. scutulatus scutulatus*, paraphyletic [40]. All three venom types have been documented in the three lineages [31, 32, 36, 41] and *C. scutulatus salvini sensu stricto* is thought to be exclusively Type A, but only four venom samples have been analyzed [24, 37]. In regions where the lineages come into contact, significant gene flow is occurring [40]. Gene flow is inferred to be highest between the Chihuahuan Desert and Central Mexican Plateau lineages even though those lineages diverged more anciently than the Chihuahuan Desert and Sonoran Desert populations [40]. This indicates that the historic barrier associated with the rise in elevation at the Mexican Plateau uplift does not completely prevent gene flow.

Mojave Rattlesnakes' distribution spans an ecologically diverse area, genetic lineages are well defined despite gene flow, and spatial diversity of venom phenotypes is known. However, whether spatial patterns of venom diversity can be explained by population structure, environmental variability, or balancing selection remains unknown. Here, we investigate the relationship between phenotype, genotype, and environment to determine if any of these variables can explain the diversity of venom phenotypes observed in Mojave Rattlesnakes. Specifically, we test whether venom phenotypes are monophyletic, whether they are associated with environmental

parameters (particularly elevation and temperature), and whether proteomic and morphological differences exist between populations with the same venom type.

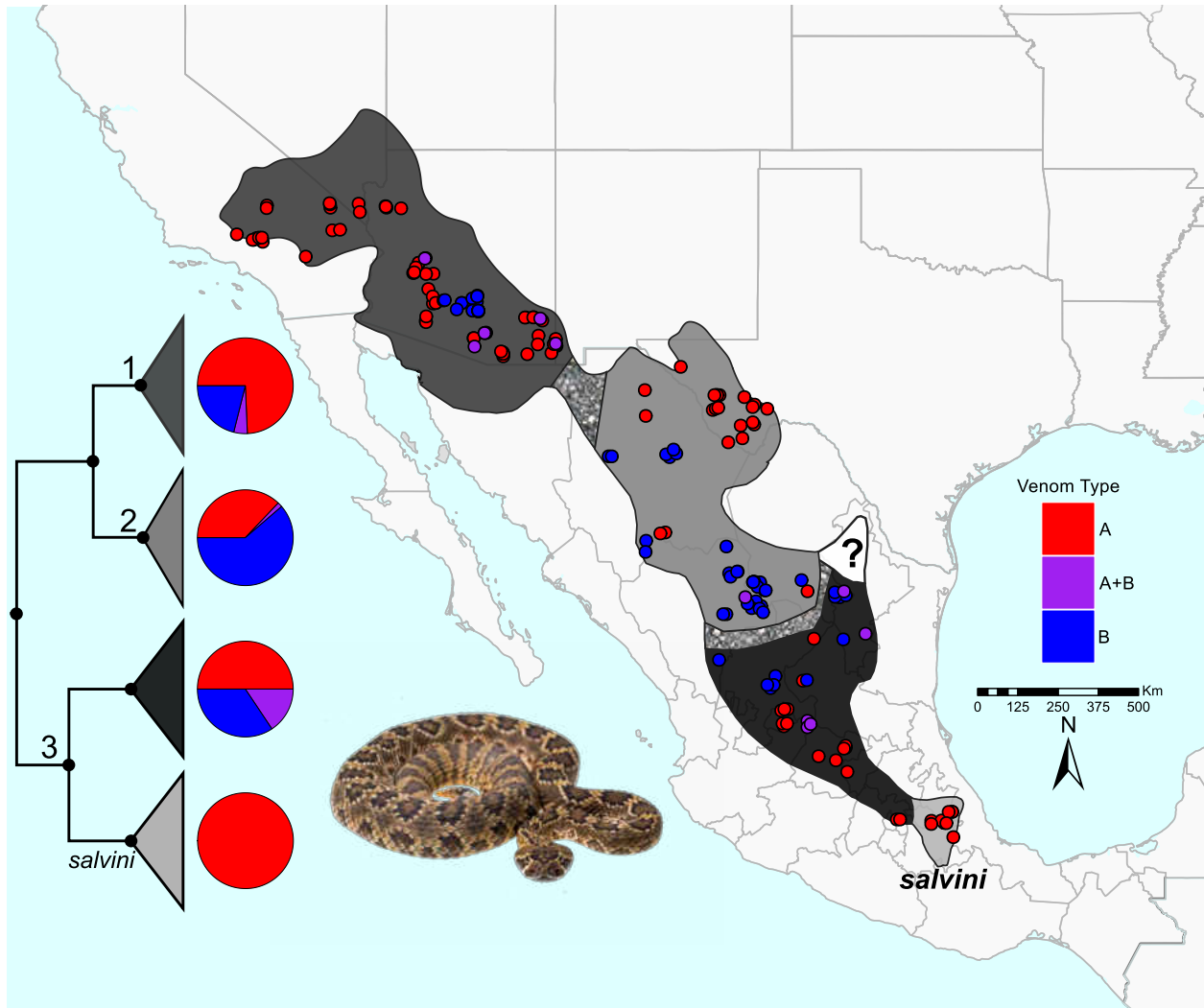


Figure 2.1: Distribution and sampling of Mojave Rattlesnakes collected from throughout their range. Red, purple, and blue represent Type A, Type A + B, and Type B venom, respectively, in the pie charts and the sampling points. Mottling in the distribution are areas of gene flow between lineages. Pie charts represent the proportion of each venom type collected from each lineage. Cladeogram in lower left of the four mitochondrial lineages from Figure 2.3 with the three nuclear clades from Schield *et al.* [40] numbered: 1 - Sonoran lineage, 2 - Chihuahuan lineage, 3 - Central Mexican Plateau lineage, ? - not sampled/unknown lineage

Materials and methods

Ethics statement

Scientific collecting permits in the United States were issued by the State of Arizona Game and Fish Department (SP628489, SP673390, SP673626, SP715023), the California Department of Fish and Wildlife (SC-12985), the New Mexico Department of Game and Fish (3563, 3576) and Texas Parks and Wildlife (SPR-0390-029). In Mexico, collection permits were issued by Secretaria de Medio Ambiente y Recursos Naturales of the Estados Unidos Mexicanos (SEMARNAT: SGPA/DGVS/03562/15, SGPA/DGVS/01090/17 and FAUT-0015). Interactions with animals were approved by the University of Central Florida's (UCF) Institutional Animal Care and Use Committee under protocol 13-17W and followed the American Society of Ichthyologists and Herpetologists ethical guidelines.

Sample collection and DNA extraction

We collected representatives of both subspecies and all lineages of Mojave Rattlesnakes from throughout their distribution. For most samples collected in the field, we obtained venom and tissue. When possible, voucher specimen were created and deposited (Supplemental Table B.1). Tissues were stored in ethanol and RNAlater (ThermoFisher Scientific, Waltham, Massachusetts, USA) and venom was collected and vacuum dried, put in liquid nitrogen, and/or stored at -80°C. We collected a total of 216 individuals. Of these, 114 had tissue and venom, 34 had only venom, and 68 had only tissue (Supplemental Table B.1). Whole genomic DNA was extracted from samples using the Serapure bead extraction protocol of Rohland and Reich [42] following modifications in Faircloth [43].

Reverse-phased high performance liquid chromatography to determine venom type

All venom was either vacuum dried or lyophilization prior to use. We then resuspended venom in Millipore-filtered water and centrifuged to remove insoluble debris. We determined protein concentration on a Qubit 3.0 Fluorometer (ThermoFisher Scientific) using the Qubit Protein Assay (ThermoFisher Scientific) following the manufacturer's protocol. To determine the venom type (A, B, or A + B) of each individual, we used Reverse-phased High Performance Liquid Chromatography (RP-HPLC) based on the protocol in Margres *et al.* [44].

We injected 100 μg of venom onto a Jupiter C18 column (250 x 2 mm; Phenomenex, Torrance, California, USA) using two solvents: 1 = 0.1% trifluoroacetic acid (TFA) in water and 2 = 0.075% TFA in acetonitrile. We used a Beckman System Gold HPLC (Beckman Coulter, Fullerton, California, USA) located in the Florida State University (FSU) Department of Biological Science Analytical Lab. The gradient started with 95% A and 5% B for 5 minutes followed by a 1% per minute linear gradient to 25% B, followed by a 0.25% per minute linear gradient to 55% B, a 2% per minute linear gradient to 75% B, a 14% per minute linear gradient to 5% B and then 5 minutes at the initial conditions all at a 0.2 mL/min flow rate. Run time was 180 minutes for each sample and the effluent was monitored at 220 and 280 nm [45].

We determined venom type for 148 individuals based on the presence of both subunits of MTX (Type A) and the presence of SVMPs (Type B) based on previous RP-HPLC profiles in *C. scutulatus* [38, 46] and under these conditions [13]. When venom was available, this was the primary means of determining venom type.

Mojave toxin assay

To confirm venom type from RP-HPLC and to determine venom type when venom was unavailable, we used PCR assays for both subunits of Mojave Toxin (MTXA and MTXB). We amplified two fragments for each subunit using the primers designed by Zancolli *et al.* [47]. They

were developed to determine MTX presence in Arizona and New Mexico, USA and have also been used successfully in Mexico [32]. PCR amplification was conducted on each DNA sample under the following conditions per 10 μL reactions: 3.5 μL PCR water, 1 μL 10X Sigma Buffer (Sigma-Aldrich, St. Louis, MO, USA), 1 μL of 25 mM MgCl_2 (Sigma-Aldrich), 1.3 μL of 2.5 mM each dNTPs (INFORMATION), 0.5 μL of each primer at 10 μM , 0.2 μL of Taq Polymerase (Sigma-Aldrich), and 2 μL DNA. PCR was conducted on PTC200 Thermal Cycler (Bio-Rad, Hercules, CA, USA): 3.5 minutes at 94°C, 35 cycles of 30 seconds at 94°C, 1 minute at 57°C, and 1 minute at 72°C, and a final extension at 72°C for 5 minutes. PCR product amplification was evaluated on a 2% agarose gel using GelRed dye (Biotium, Fremont, CA, USA) to determine if the subunits were present. If one assay was positive, the individual was considered to have MTX and be Type A because no data exist that suggest an individual can have the gene and not express it [32, 47]. For samples with tissue only, we cannot differentiate between Type A and Type A + B.

Venom phylogeography of C. scutulatus

To determine if venom type correlated with phylogenetic lineage, we PCR amplified and sequenced NADH4 (ND4) for any individual in our dataset not already sequenced in Schield *et al.* [40]. As outgroups, we included one sample each from *C. viridis*, *C. cerberus*, and *C. oreganus* which are the sister species complex to *C. scutulatus* (Supplemental Table B.1). We used the primers ND4 and Leu to sequence the partial ND4 gene as well as the tRNAs His, Ser, and Leu [48]. PCR was conducted under the following conditions per 10 μL reaction: 3.8 μL PCR water, 1 μL 10X Sigma Buffer (Sigma-Aldrich), 1 μL of 25 mM MgCl_2 (Sigma-Aldrich), 0.8 μL of 2.5 mM each dNTPs (Invitrogen, Waltham, Massachusetts, USA), 0.5 μL of each primer at 10 μM , 0.4 μL of Taq Polymerase (Sigma-Aldrich) and 2 μL DNA. PCR was conducted on a PTC200 Thermal Cycler (Bio-Rad): 3.5 minutes at 94°C, 35 cycles of 30 seconds at 94°C, 1 minute at 53°C, and 1 minute at 72°C, and a final extension at 72°C for 5 minutes.

Amplicons were purified by adding 0.5 μL FastAP (ThermoFisher Scientific #EF0651),

0.05 μL Exonuclease 1 (ThermoFisher Scientific #EN0581), and 7.45 μL PCR water to each 30 μL reaction and then placed on the thermal cycler for 30 minutes at 37°C followed by 15 minutes at 85°C. Sequencing was done in both directions at Eurofins Scientific (St. Charles, Missouri, USA) on an ABI 3730 genetic analyzer (Applied Biosystems, Waltham, MA, USA). Sequences were assembled and edited in Geneious v 10.1.2 (Biomatters Ltd., Auckland, New Zealand). Alignments were created with the MAFFT v 7.22 alignment algorithm [49] implemented with default parameters in Geneious. We verified alignments by eye and trimmed low quality nucleotides and also checked to ensure there were no frameshift mutations. Our final alignment was 884 nucleotides for 189 ingroup and three outgroup taxa.

We used PartitionFinder 2.1.1 [50] to determine the best-fit model of evolution for ND4 split by codon position and between protein-coding and tRNA regions. We used the "greedy" search algorithm, Bayesian Information Criterion (BIC), linked branch lengths, and only tested models available in BEAST. The starting tree was generated using PhyML v 3.0 [51]. These analyses were run in the UCF Advanced Research Computing Center (ARCC) on the Stokes High Performance Computer (SHPC). The site models from PartitionFinder2 were HKY+I for the first codon position and the tRNA together, HKY for the second codon position, and HKY+ Γ for the third codon position. All three models estimated nucleotide frequencies (+X).

To determine the mitochondrial lineage of new samples in comparison to Schield *et al.* [40], we used Bayesian inference (BI) in BEAST2 v 2.4.5 [52]. We used BEAUti v 2.4.5 [52] to generate .xml files. We ran the analysis four independent times for 10^8 generations each under a coalescent constant population tree prior and strict molecular clock. We stored 10,000 trees per run. All runs were checked in Tracer v 1.6 [53] to ensure stationarity was reached and all ESS values for the individual and combined runs were ≥ 200 . We then combined .trees files from each run in LogCombiner v 2.4.5 [52] and resampled 10,000 trees. Finally, we annotated the final tree with 10% burn-in in TreeAnnotator v 2.4.5 [52].

Azocasein metalloproteinase assay

To determine if there are differences within venom types among populations, we performed an azocasein metalloproteinase assay on 146 samples in triplicate [54]. We incubated 20 μg of venom with 1 mg of azocasein substrate in buffer composed of 50 mM HEPES and 100 mM NaCl at a pH of 8.0 for 30 minutes at 37°C. We stopped the reaction with 250 μL of 0.5 M trichloroacetic acid, vortexed, and brought it to room temperature. We then centrifuged it at 2000 rpm for 10 minutes. Sample absorbance was read at 342 nm and reported in $\Delta_{342\text{nm}}/\text{min}/\text{mg}$ of venom protein [46]. To determine if there were differences, we used a Kruskal-Wallis tests with venom type and lineage as factors implemented in R v. 3.4.3 (R Development Core Team 2017).

Kallikrein-like serine protease assay

To determine if there are differences within venom types for other toxin classes, we performed a kallikrein-like serine protease assay on 60 samples following Mackessy [55]. We added 0.8 μg of whole venom to 373 μL of the same buffer as above. Samples were incubated for 3 minutes at 37°C and then 50 μL of substrate (Bz-ProPheArg-pNA; Bachem, Torrance, CA, USA), the sample was vortexed and placed back at 37°C for three minutes. The reaction was stopped with 50% acetic acid. Sample absorbance was read at 405 nm and the specific activity was calculated based on a standard curve of p-nitroaniline and reported as nanomoles of product produced per minute per mg of venom [46]. We used a Kruskal-Wallis test with venom type and lineage as factors in R to determine if there were significant differences.

SDS-PAGE protein gel electrophoresis

To estimate overall venom diversity, we performed SDS-PAGE protein gel electrophoresis on 110 samples following Smith and Mackessy [46]. We loaded 20 μg of whole venom into wells of a NuPAGE Novex bis-tris 12% acrylamide mini gel (Life Technologies, Grand Island, NY,

USA) and electrophoresed in MES buffer at 175 volts for 45 minutes. To estimate the molecular weight, we used 7 μ L of Mark 12 standard. We stained gels overnight on a gentle shake with 0.1% Coomassie brilliant blue R-250 in 50% and 20% acetic acid (v/v). Gels were destained for approximately two hours in 30% methanol and 7% glacial acetic acid (v/v) in water until bands were clearly visible. Gels were gently shaken overnight at room temperature in 7% acetic acid (v/v) storage solution and imaged the next day using an HP Scanjet 4570c scanner.

Head morphology analysis

To determine if there are head morphological differences between individuals with Type A and Type B venom, we measured functional morphological characters for 57 individuals from Arizona, USA. We followed Margres *et al.* [25] and measured SVL (snout to vent length), HL (head length), HW (head width), IF (interfang distance), and FL (fang length) for both fangs when not broken and then averaged them. We measured SVL with a tailor's tape from the tip of the snout to the posterior end of the cloaca to the nearest 1 mm. We used IP54 digital calipers (iGaging, San Clemente, CA, USA) to measure HL, HW, IF, and FL to the nearest 0.01mm. Head length was measured from the tip of the snout to the articular-quadrato joint, HW was the widest point behind the eye, IF was the distance between the two fang maxillae, and fang length was from the top of the maxilla to the tip of the fang while folded. Average FL was determined if neither fang was broken. If one fang was broken, then the unbroken fang was used. If both fangs were broken, that individual was not used in analyses involving FL.

All transformations and analyses were done in R v. 3.4.3 (R Development Core Team 2017). Each measurement was natural-log-transformed so they met the assumptions of normality and homoscedasticity. To size correct the data, we used the \ln SVL value and subtracted each other value from it (ex. \ln SVL- \ln HL) to generate new columns for HL, HW, IF, and FL that were standardized based on the length of the snake. Using these data, we compared individuals of the three venom types using Kruskal-Wallis tests and venom type as the factor. When the comparison

of all three venom types was significant, we then did pairwise tests for each venom type.

Niche modeling of venom type

To examine whether differences in venom type could be explained by differential niche occupation, we constructed ecological niche models (ENMs) for the occurrence of A and B venom types across the range of *C. scutulatus*. We used geographic localities for 123 Type A and 68 Type B *C. scutulatus* whose venom types were determined by RP-HPLC and/or MTX assays. ENMs were generated using MAXENT v3.4.1 [56] implemented through the R package dismo [57]. MAXENT uses occurrence records and a user-provided suite of environmental variables to predict the suitability of habitat and likelihood of occurrence across a landscape [58].

To limit the effect of sampling bias in the construction of EMS, we subsampled the total records and retained only those points which were separated by a 0.25 km minimum distance. This reduced the dataset to 64 and 31 representative A and B *C. scutulatus*, respectively. For environmental data we used the 19 climatic variables collecting in the WorldClim dataset v 1.4 [59] as well as elevation, slope, and aspect with a 30 arc second resolution. To avoid biasing the modeling through inclusion of highly correlated inputs [60], we removed 8 variables with a pair-wise pearson's correlation coefficient >0.90 leaving 14 environmental variable characterizing southwestern North America. MAXENT models were run with 20 replicates and average model performance was evaluated by determining Area Under the Curve (AUC) for each model.

To test niche equivalency and niche similarity of Type A and Type B, we used the psuedoreplicate simulation method of Warren *et al.* [61] as implemented in the R package phyloclim [62]. Both tests were run with 99 replicates to build a null distribution against which to test values of Schoener's D and Warren's I inferred from the full data models. To test niche equivalency, the occurrence points of two species (e.g. species A and B) are combined and randomly partitioned into two datasets with sizes equal to those of species A and B. ENMs are generated for each dataset and their similarity values (D and I) are calculated. This process is

repeated to build a null distribution against which to test actual values of D and I inferred from the full data model. This method tests the null hypothesis that the two models are not significantly different.

In contrast, the test of niche similarity compares the ENMs of one species to an ENM of randomly selected subset of background cells (which include both presence and absence locations) of the other species. This is replicated 99 times and is then reversed such that the models for species A is tested against a background model for species B and species B is tested against a background model for species A. Comparison of the true D and I statistics to the pseudoreplicate distributions tests the hypothesis that one species' niche model predicts the occurrences of the other species more than expected by chance. In addition, to more directly test for presence and absence differences as relating to specific environmental variables, species we used logistic regression on the 14 variables used for ENMs. This approach was used to distinguish variables that may be important for *C. scutulatus*' ecology, but may not be correlated with differences between venom types.

Data availability

All novel ND4 sequences generated in this study were deposited in GenBank. Specimen vouchers were deposited in the appropriate museums based on permit requirements. All specimen data including morphology and assay data are listed in Supplemental Table B.1.

Results

Venom type in C. scutulatus

For the 116 samples from which both venom and blood were sampled, the RP-HPLC venom type and the MTX assay were in agreement. None of our sampled individuals that lacked MTX in the RP-HPLC profile had a positive result in the PCR assays (representative examples

in Figure 2.2). For all samples that had both subunits of MTX in their venom profile, they were positive for MTX in at least one PCR assay. All but seven were positive or negative for all four MTX PCR assays. For the 68 samples that only had DNA available, all but five were unanimous for venom type. A total of 12 samples were not positive for all four assays: two were positive for three of four assays, seven were positive for two of four, and three were positive for one assay. In total, 133 samples were Type A, 72 were Type B and 11 were Type A + B (Supplemental Table B.1).

Crotalus scutulatus phylogeography

Our ND4 phylogeny mirrored that of Schield *et al.* [40]. We recovered the three primary clades that also corresponded to their RADseq clades (Figure 2.3). Because of the addition of more sampling in the Central Plateau of Mexico, we identified potential substructure in the region. Based on where each individual was in our phylogeny, we were able to assign each individual to a population and test if there were difference in venom characteristics that corresponded to shared ancestry. For our analyses, we used the three primary mitochondrial lineages to compare populations and venom type. Only the population corresponding to *C. scutulatus salvini* was monophyletic for venom type (Figure 2.3).

Azocasein metalloproteinase assay

Metalloproteinase activity was significantly different between venom types but did not differ within venom types between clades (Figure 2.4). Type A venom metalloproteinase activity ranged from 0 to 0.668 $\Delta_{342nm}/\text{min}/\text{mg}$ (n = 86, Average = 0.052, Median = 0.020), Type A + B from 0.483 to 1.275 $\Delta_{342nm}/\text{min}/\text{mg}$ (n = 10, Average = 0.816, Median = 0.919), and Type B from .208 to 2.043 $\Delta_{342nm}/\text{min}/\text{mg}$ (n = 50, Average = 0.945, Median = 0.919).

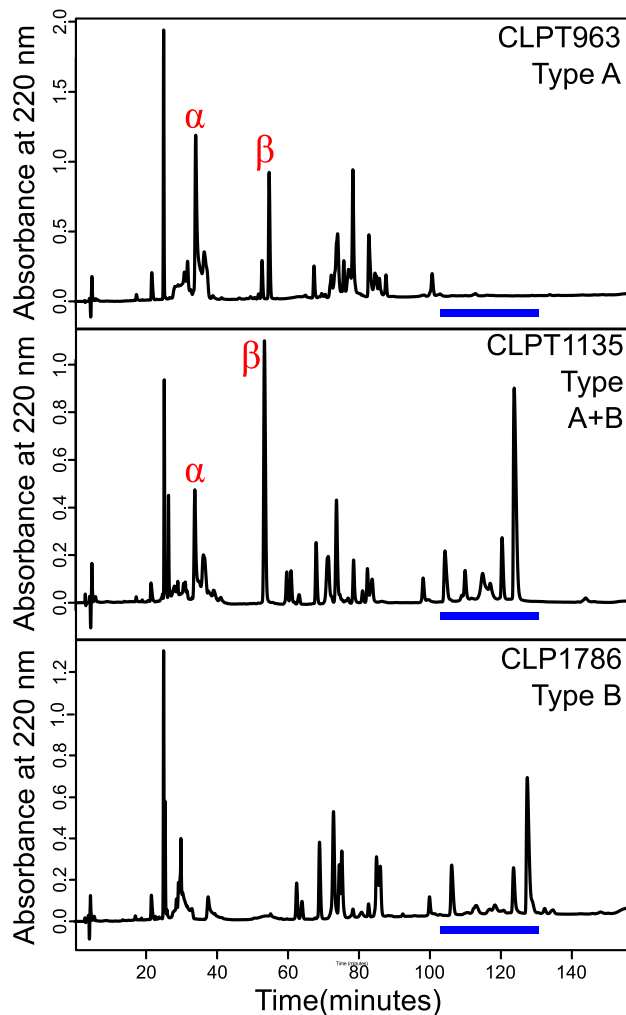


Figure 2.2: Representative reverse-phased high performance liquid chromatography (RP-HPLC) profiles of Type A (top), Type A + B (middle), and Type B (bottom) venom of Mojave Rattlesnakes. The acidic (α) and basic subunit (β) peaks for Mojave toxin are marked and the region where snake venom metalloproteinases elute is marked with a blue bar.

Kallikrein-like serine protease assay

We only had data from the Sonoran and Chihuahuan lineages for Kallikrein-like activity. There were significant differences between venom types but not with regard to population (Figure 2.5). Type A venom had significantly higher activity than Type B individuals. Type A values ranged from 132.77 to 733.20 nmol product/min/mg ($n = 28$, Average = 386.58, Median = 380.10),

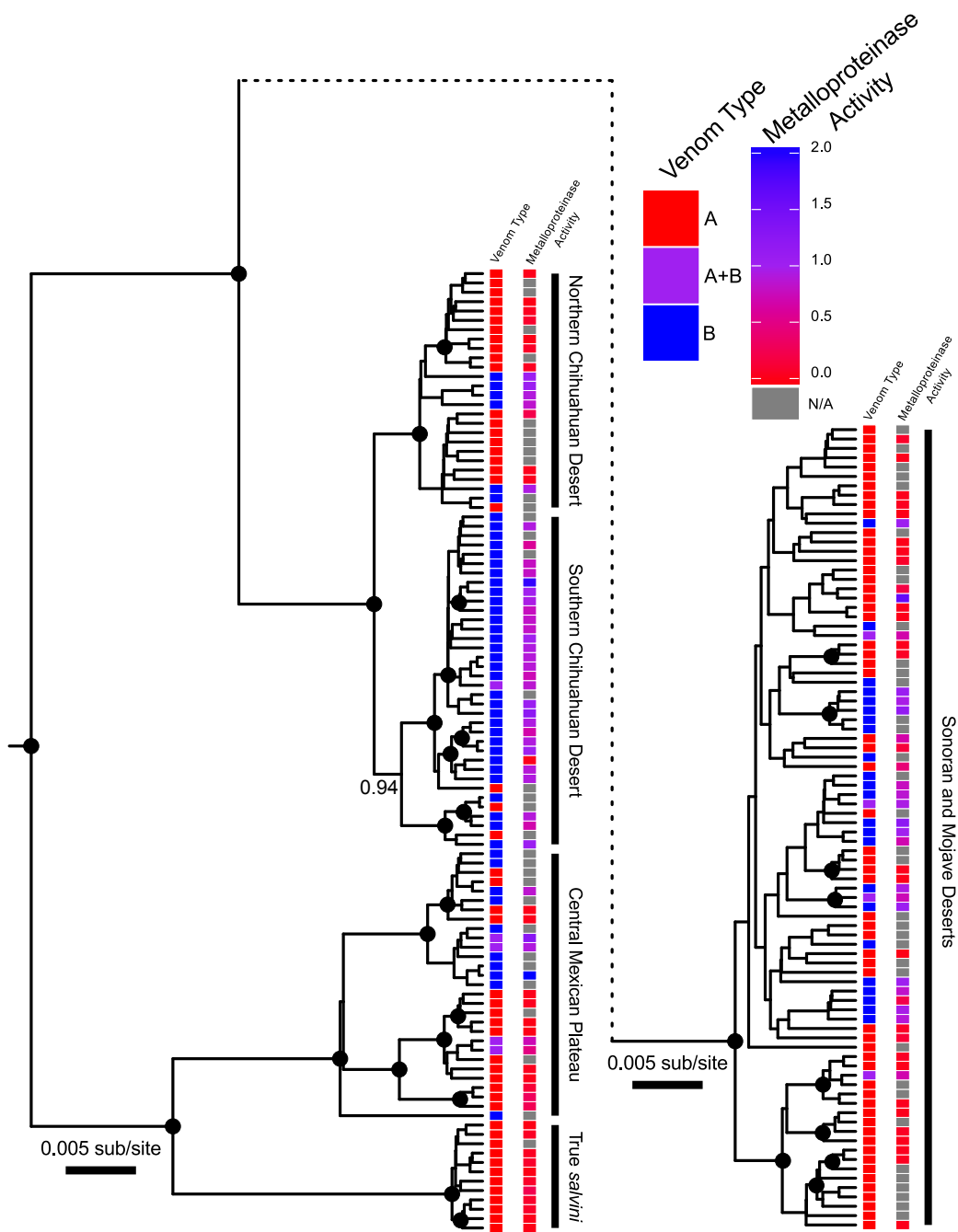


Figure 2.3: Bayesian inference phylogeny based on ND4 sequence from 189 *C. scutulatus*. The dashed line indicates where the Sonoran and Mojave Desert lineage was moved from and no size adjustments occurred. Venom type as discrete characters and metalloproteinase activity on a continuous scale are mapped onto the phylogeny. Individuals without activity values are gray boxes. Dots on nodes represent significant posterior probability values of ≥ 0.95 .

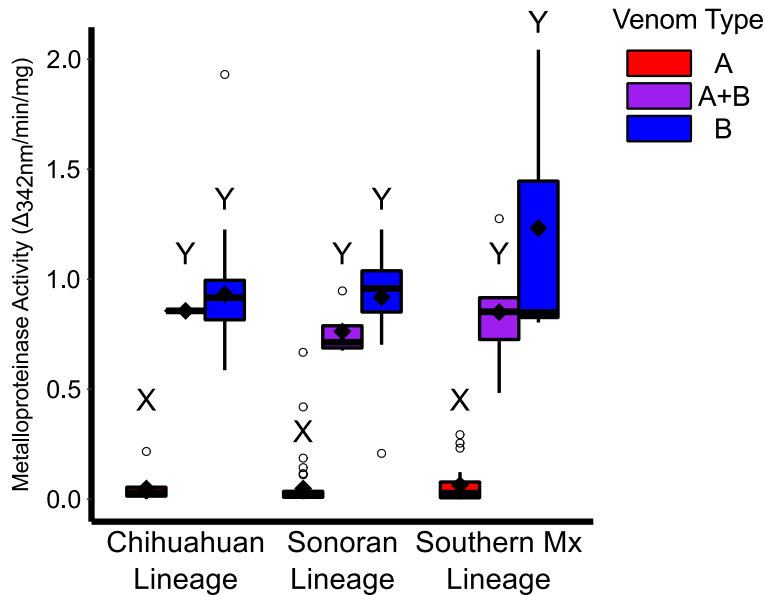


Figure 2.4: Metalloproteinase activity levels for venom types within each lineage. Type A venom had significantly less metalloproteinase activity regardless of lineage. Type B and Type A + B were not significantly from each other in any comparison.

Type A + B from 216.87 to 382.77 nmol product/min/mg (n = 5, Average = 301.54, Median = 298.30), and Type B from 70.69 to 469.24 nmol product/min/mg (n = 27, Average = 277.231, Median = 285.86).

SDS-PAGE

SDS-PAGE confirmed venom type for all samples. The basic subunit of MTX was clear at 14kD in Type A and Type A + B individuals and absent in Type B individuals (Figure 2.6 and 2.7). Additionally, SVMPs were clearly visible in Type B and Type A + B individuals at 55kD and ~22kD and absent in Type A individuals. Other toxin classes, particularly myotoxins were highly variable among individuals regardless of location and lineage. SDS-PAGE illustrates additional diversity in C-type lectins and non MTX phospholipase A₂s as well as the uniformity in snake venom serine proteases and cysteine rich secretory proteins (Figure 2.6 and 2.7).

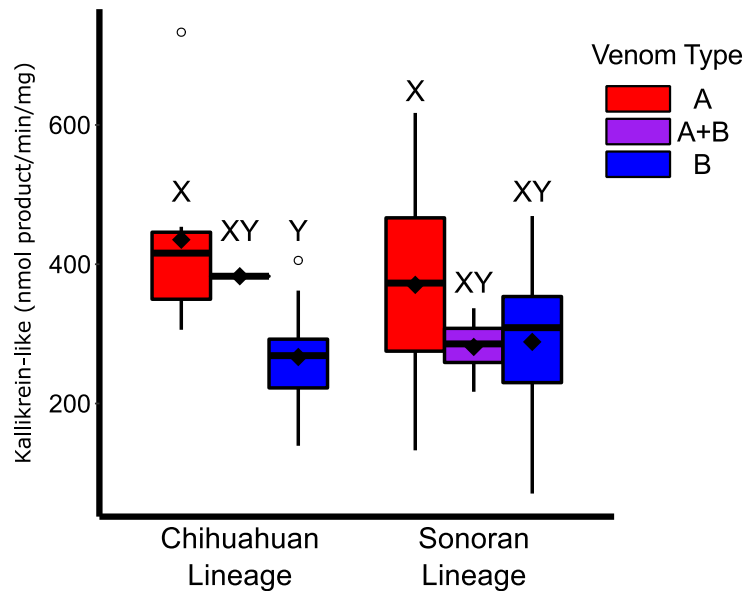


Figure 2.5: Kallikrein-like serine protease activity between the Chihuahuan and Sonoran lineages. Type B venom had significantly lower activity than Type A venom but there were no differences in the same venom type between the two populations and Type A + B venom was not significantly different than Type A or Type B venom.

Morphological analysis

All 57 individuals used in this analysis were from the Sonoran lineage. Thus, they are all genetically similar and provide the best comparison of potential morphological differences associated with venom type. We did not find significant differences between head width or head length between venom types. We did find a significant difference in interfang distance between venom types (Figure 2.8). Type A individuals had a larger distance between the fangs than Type B individuals and Type A + B individuals were intermediate. Additionally, there was a trend for Type A individuals to have longer fangs than Type B and Type A + B individuals (Figure 2.8).

ENMs between venom type

We did not find significant differences between the ecological niche models (ENM) created for Type A and Type B individuals (Figure 2.9). The ENM for Type A and Type B were not

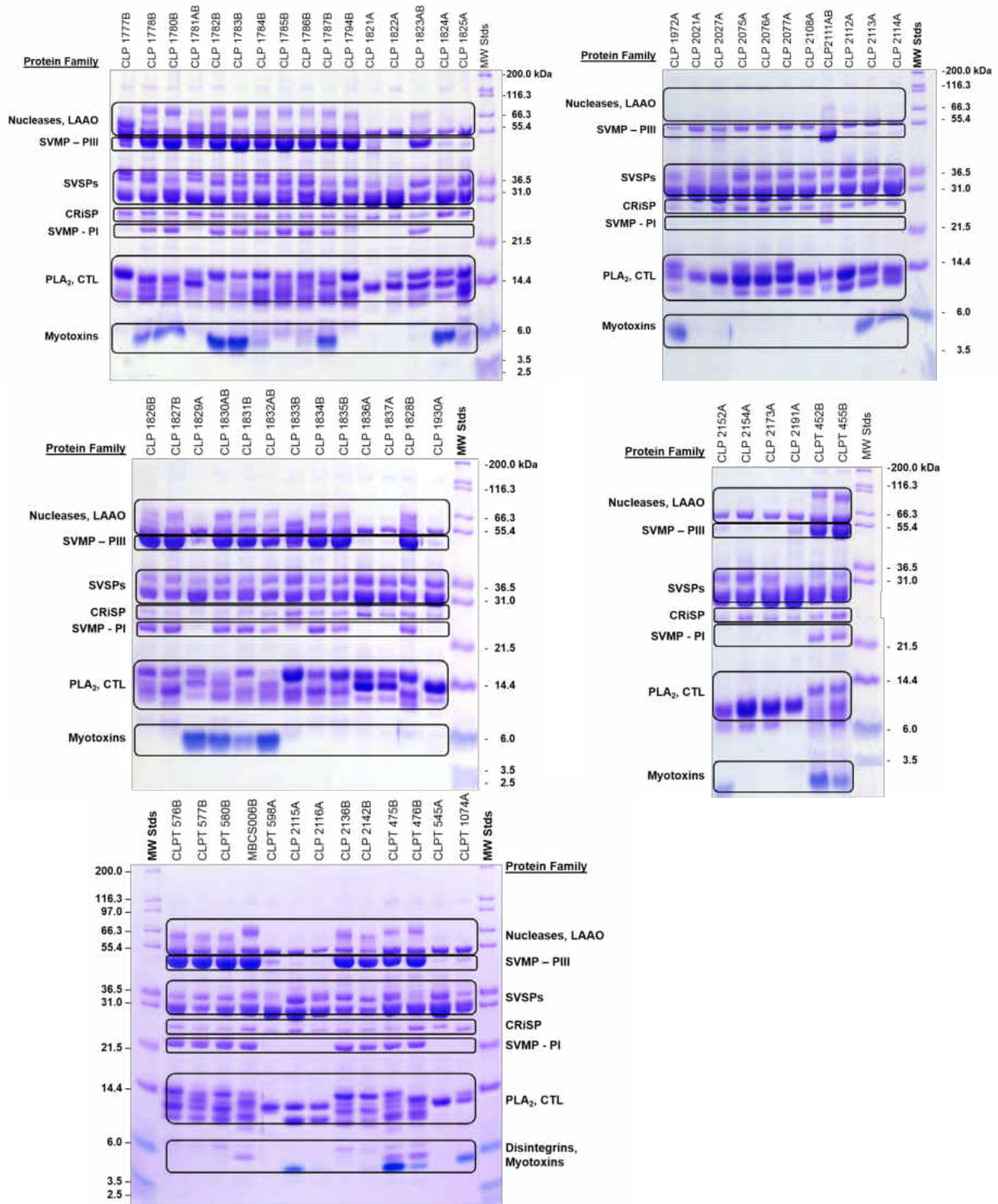


Figure 2.6: SDS-PAGE gel images for 59 of 110 samples with toxin families labeled. Venom type of each individual is added to the end of the sample name. When present, the basic subunit of Mojave toxin is at approximately 14kD. See Figure 2.7 for remaining sample images.

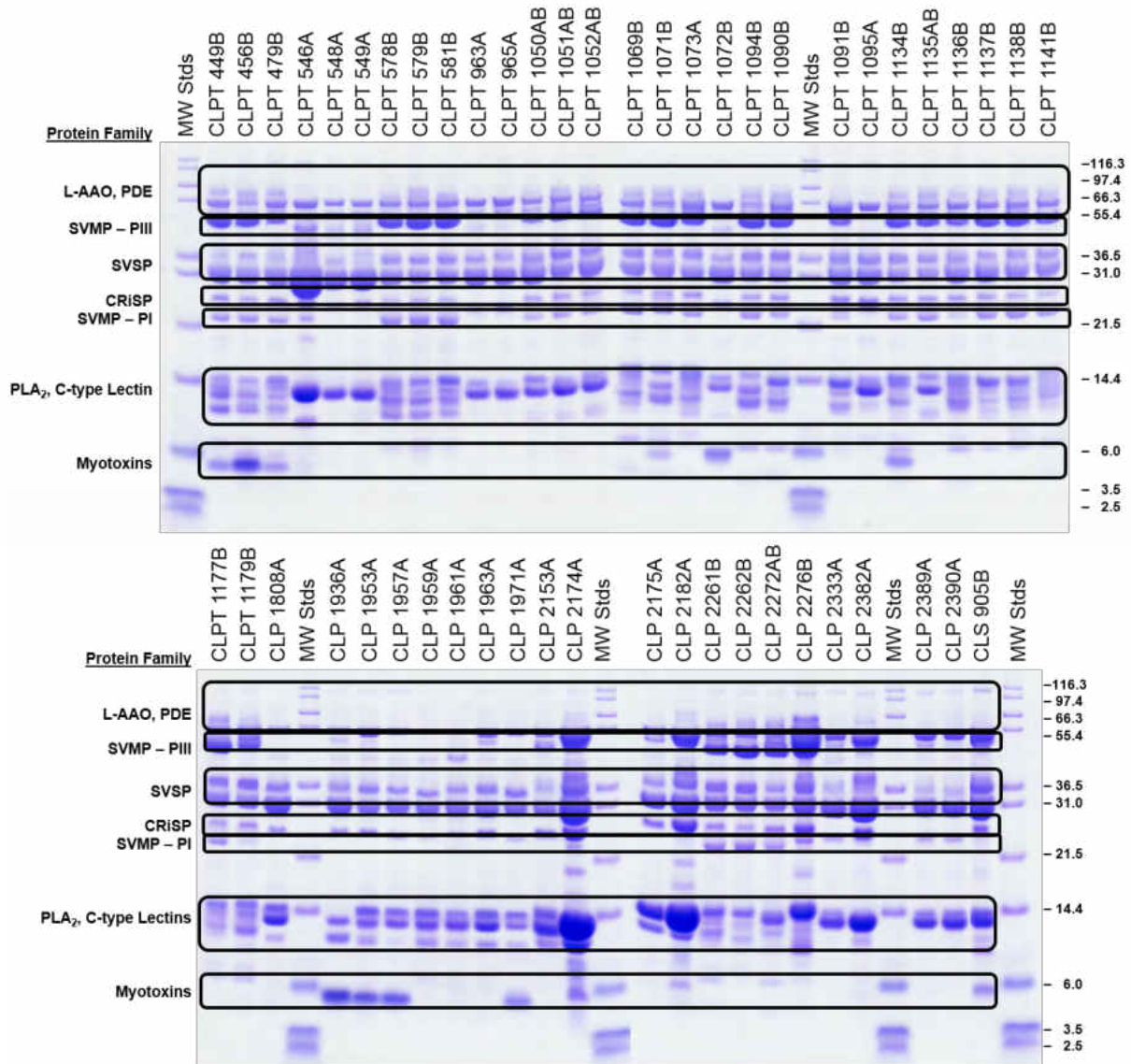


Figure 2.7: SDS-PAGE gel images for 51 of 110 samples with toxin families labeled. Venom type of each individual is added to the end of the sample name. When present, the basic subunit of Mojave toxin is at approximately 14kD. See Figure 2.6 for remaining sample images.

equivalent to each other (Figure 2.10) but they were more similar than would be expected by chance (Figure 2.11). Area under the curve (AUC) from comparison of the model and the Type B (AUC = 0.946) *C. scutulatus* was comparable to AUCs from Type A test data (AUC = 0.956), indicating similarity in predictive power. The BioClim variable BIO6 (minimum temperature of

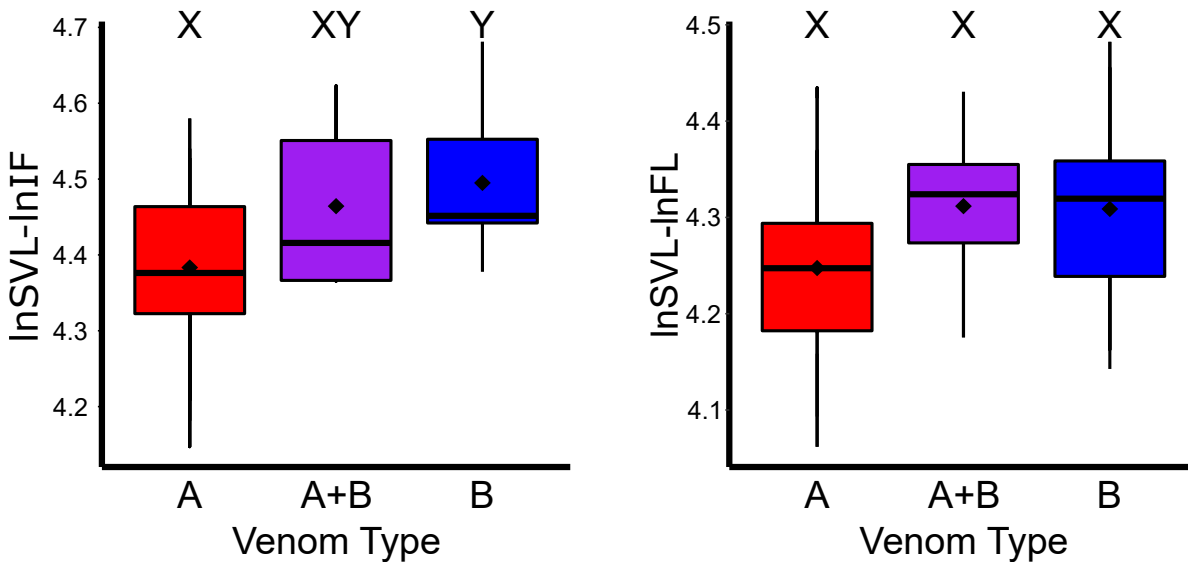


Figure 2.8: Comparison of interfang distance (IF) on the left and fang length (FL) on the right in the Sonoran lineage of Mojave Rattlesnakes. Snout-vent length (SVL) was used to control for different sizes among animals. Smaller values on the y-axis are larger measurements. There was a trend of Type A individuals having longer fangs but it was not significantly different from Type B or Type A + B individuals. Type A individuals did have significantly wider distances between their fangs compared to Type B individuals but not compared to Type A + B individuals.

the coldest month) explained the most variation for each model and was significantly different between the two models (Figure 2.12 and Table 2.1). The only other variable that was significantly different was BIO1 (annual mean temperature) but it explained almost no variation in either model (Figure 2.12 and Table 2.1).

Discussion

The venom phenotype dichotomy was first described in Mojave Rattlesnakes in the 1930's based on different symptoms of snakebite [63]. Subsequent studies of the distribution of venom types led to the conclusion that *C. scutulatus* has neurotoxic venom through the majority of their range [11, 24, 31, 33]. Our data indicate that this generalization is not accurate, and that both Type A and Type B phenotypes occur throughout the distribution with at least three transition

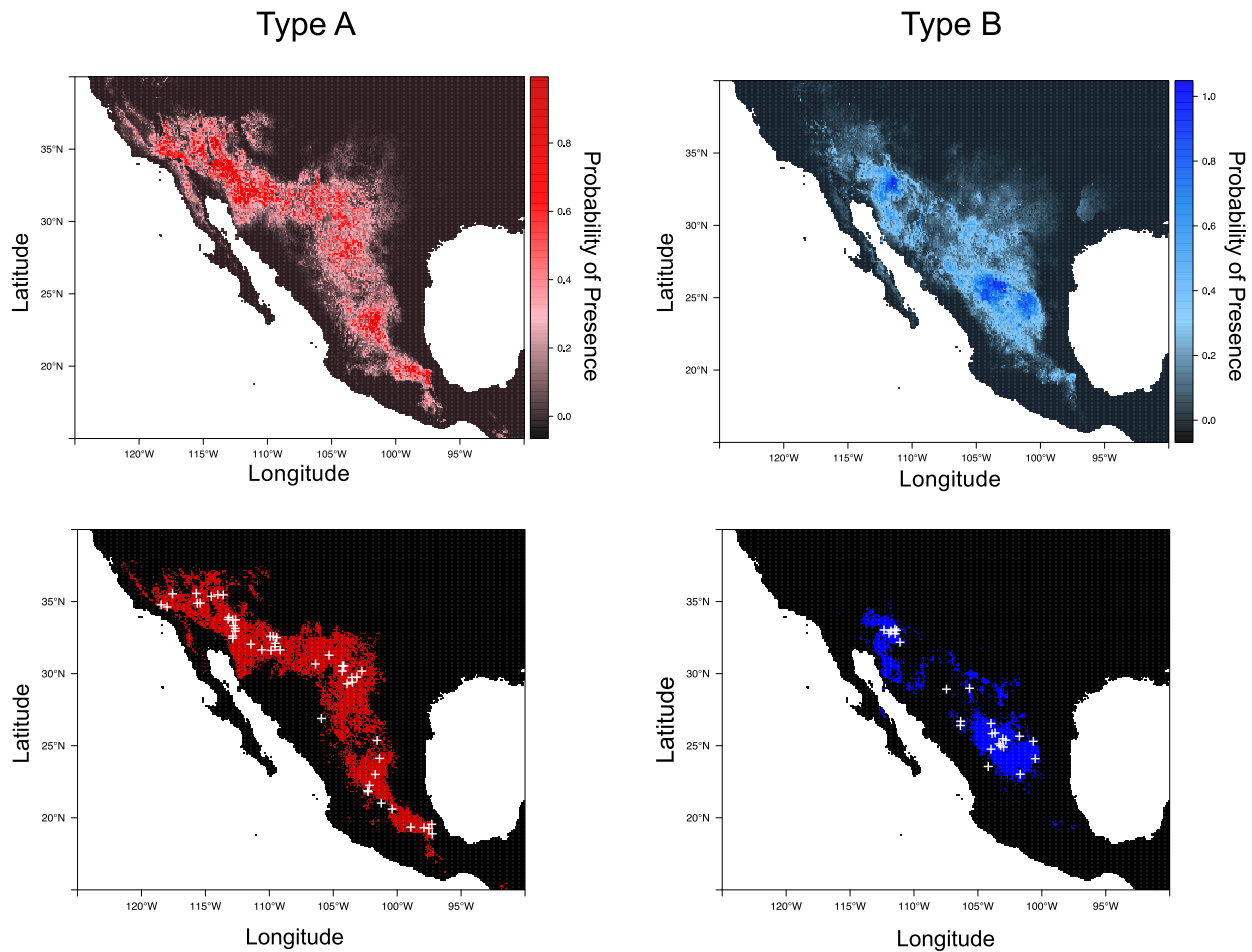


Figure 2.9: Ecological niche models generated in MAXENT using 64 Type A (left) and 31 Type B (right) Mojave Rattlesnakes scaled by probability of presence (pp). Lower maps display model distributions where each venom type is expected to occur based on a threshold point where model sensitivity and specificity are highest (pp > 0.22 for Type A, pp > 0.07 for Type B).

zones. The integrated venom phenotype, Type A + B, does occur in areas where the two primary phenotypes come into contact, but not exclusively. Several of the Type A + B individuals were found in areas with apparently fixed venom types. Each of the major phylogeographic lineages identified by Schield *et al.* [40] possess Type A, Type A + B, and Type B individuals (Figure 3.1 and 2.3).

Schild *et al.* [40] found that the subspecies *C. scutulatus salvini* at the southern end of the distribution was mitochondrially unique but indistinguishable based on RADseq data. *Crotalus*

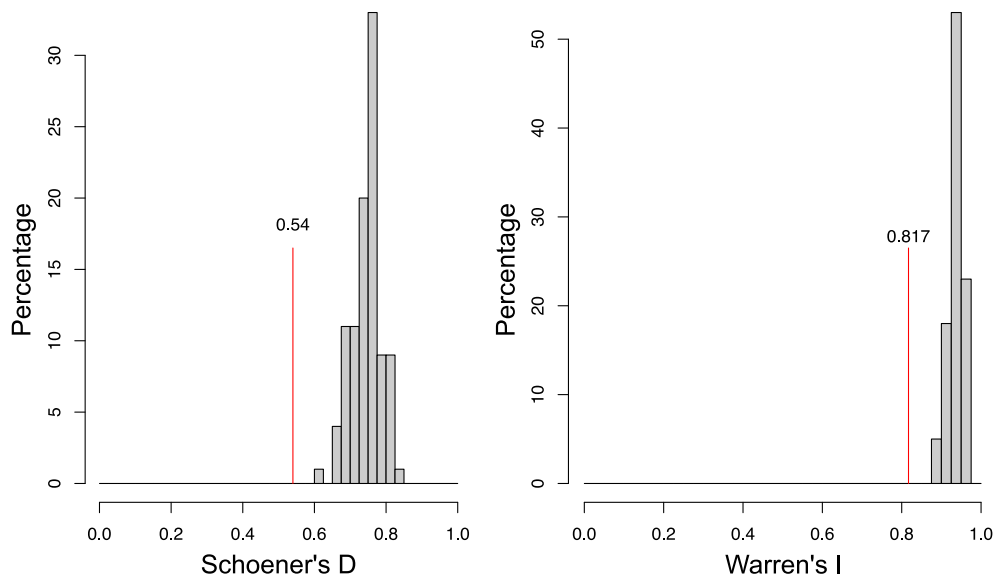


Figure 2.10: Null distributions generated to test niche equivalency using 99 permutations. Results for both Schoener's D and Warren's I reject the null hypotheses that the models for Type A and Type B are identical.

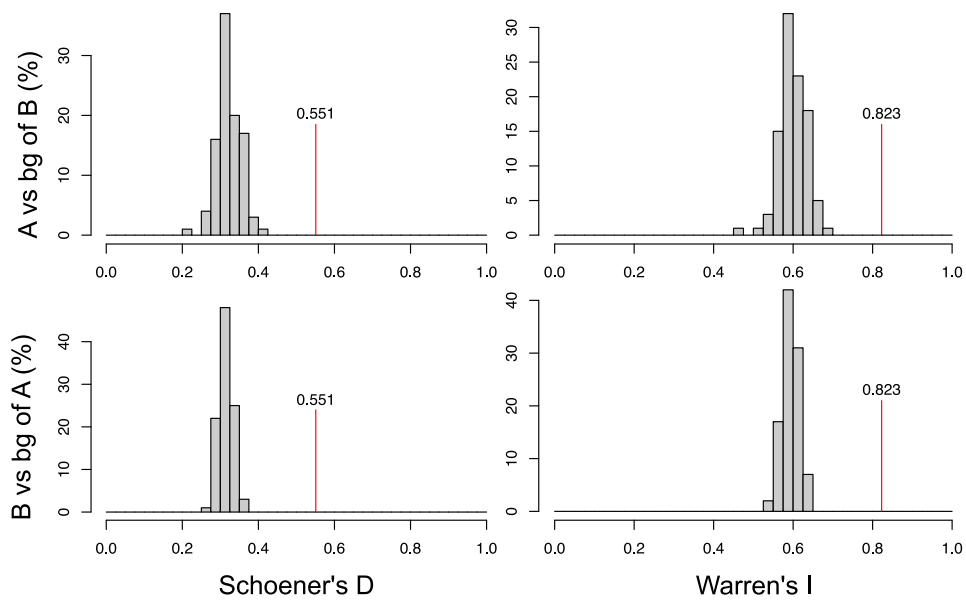


Figure 2.11: Null distributions generated using 99 permutations of Type A and Type B individuals to test niche similarity. The top row compares Type A individuals using the model background (bg) of Type B and the bottom row compares Type B individuals using the model background of Type A. Results for both Schoener's D and Warren's I reject the null hypothesis that the two models do not predict the occurrence of each other better or worse than would be expected by chance. Because the D and I values fall in the right tail of the distribution, the niches are more similar than would be expected by chance.

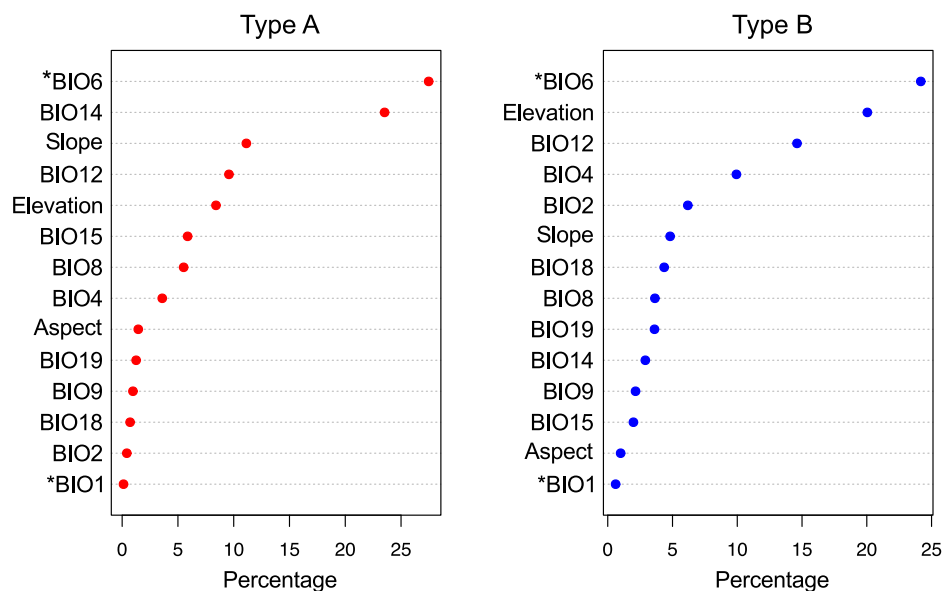


Figure 2.12: Model response to variables included in the ecological niche models for Type A and Type B venoms. Asterisks (*) indicate variables that were significantly different between the two models. BIO1 = Annual Mean Temperature, BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp)), BIO4 = Temperature Seasonality (standard deviation *100), BIO6 = Min Temperature of Coldest Month, BIO8 = Mean Temperature of Wettest Quarter, BIO9 = Mean Temperature of Driest Quarter, BIO12 = Annual Precipitation, BIO14 = Precipitation of Driest Month, BIO15 = Precipitation Seasonality (Coefficient of Variation), BIO18 = Precipitation of Warmest Quarter, BIO19 = Precipitation of Coldest Quarter.

Table 2.1: Logistic regression comparison of the 14 variables used in the Type A and Type B models for *C. scutulatus*. Bolded variables were significantly different between the two models.

Independent Variable	Df	F-value	P-value
BIO1	93	10.38	0.001
BIO2	93	3.59	0.058
BIO4	93	2.23	0.136
BIO6	93	9.78	0.002
BIO8	93	3.08	0.079
BIO9	93	0.35	0.556
BIO12	93	0.01	0.925
BIO14	93	1.81	0.179
BIO15	93	0.32	0.571
BIO18	93	2.56	0.11
BIO19	93	0.15	0.698
Elevation	93	0.07	0.797
Aspect	93	0.115	0.915
Slope	93	0.51	0.475

scutulatus salvini is the only potential population that was monomorphic for venom type. All 13 individuals collected were Type A and had very little metalloproteinase activity. Given that this population is mitochondrially and venomically distinct, it is possible that a founder event could have occurred and this population invaded the Neovolcanic Axis Mountain basin south of Mexico City. The remaining lineages, which are currently classified as *C. scutulatus scutulatus*, have all three venom types within them (Figure 2.3). The region where the Chihuahuan lineage and the southern Mexican lineage come into contact and where Schield *et al.* [40] found the highest gene flow among lineages has a high concentration of Type A + B individuals. However, in Arizona where *C. scutulatus* is genetically panmictic, the Type B population is geographically isolated, there is also a high concentration of Type A + B individuals. The lack of genetic correlation could be due to fine scale local adaptation for the phenotypes.

The biochemical properties of the venom did vary between venom types but values were not significantly different between lineages. As expected, metalloproteinase activity was higher in Type B individuals, but Type B animals from all three lineages had the same metalloproteinase activity across venom type (Figure 2.4). Average metalloproteinase activity was lower in Type A + B individuals than in Type B individuals for all three populations, but it was not significantly different (Figure 2.4). Kallikrein-like serine protease activity was highly variable with Type B individuals having significantly lower activity in the Chihuahuan lineage (Figure 2.5). The trend also exists in the Sonoran lineage (Figure 2.5). Other than SVMPs and PLA₂s, the SDS PAGE profiles indicate some variability among individuals in other protein classes, particularly myotoxins (Figure 2.6 and 2.7). Myotoxins were found in 56 of 110 individuals tested and were generally linked with Type A individuals in the Sonoran lineage and Type B individuals in the Chihuahuan lineage. This is in line with recent work documenting high myotoxin variability proteomically [38] and transcriptomically [39]. Myotoxins increase the lethality of venom in Type B individuals but it is unclear what they add to individuals with Type A venom [38].

Morphological differences may accompany the venom dichotomy and be the second

example of phenotypic integration of traits involved in envenomation [25]. Within the Sonoran lineage, there was a significant difference in the interfang distance in Type A and Type B individuals (Figure 2.8). Type A individuals had a larger distance between the fangs. There was also a trend for fang length to be longer in Type A individuals (Figure 2.8). These morphological differences between Type A and Type B venom need to be tested in each lineage to determine if it is a broader pattern. If this pattern occurs across lineages it would run counter to what was predicted. Instead of Type B venom having longer fangs for deeper penetration of venom, it may be that the increased range of Type A fangs increases the likelihood of making contact and even a small amount of venom could incapacitate the prey item. It is likely that kangaroo rats (*Dipodomys*) are the primary diet component as they share similar distributions and field observations document their interactions frequently [64]. Kangaroo rats are agile and have several defensive behaviors that seem to be a response to snake predation so it is possible the quick action of the neurotoxic venom in *C. scutulatus* is a response to rapidly incapacitating kangaroo rats [65–68]. This could represent an arms race between Mojave Rattlesnakes and kangaroo rats as seen in other systems [69, 70].

The ecological niche models were not identical between the two venom types but were more similar than would be expected by chance. Only two of the fourteen model parameters were significantly different between the two ENMs and they were the highest and lowest variables in explaining variation. The most important variable, BIO 6, which is mean temperature of the coldest month was significantly higher in the Type B model. Several hypotheses for the maintenance of the two phenotypes involve temperature and suggest that in cooler areas, Type B venom should be more common [21]. However, it is important to recognize that our current findings suggest the strongest impact of temperature at the coldest part of the year. Thus, while neither venom type may have a selective advantage during warmer parts of the year, the Type B venom may allow for individuals to be active longer at the beginning and ends of the active season. There is some evidence that Type B venoms speed up digestion at cooler temperatures [19, 71] but these data

do not support greater digestive efficacy [23, 71]. If our hypothesis of longer active season is correct, we would expect differing activity patterns for *C. scutulatus* with the two venom types. Alternatively, another ecological parameter not tested in this study may differ such as proportion of different prey items in the diets of individuals with the different venom types.

We found little to no evidence for phylogenetic or environmental variables being responsible for the complex venom phenotypes in *C. scutulatus*. Venom phenotypes in *C. scutulatus* are geographically structured, occur in ecologically similar environments, and individuals are genetically similar at neutral loci. All three venom phenotypes are found within the three lineages identified by Schield *et al.* [40] as part of panmictic populations (Figure 2.3). The maintenance of these two phenotypes within the population in spite of these factors represents an exemplary case of adaptation through balancing selection. Of the six types of balancing selection, we can eliminate two. Heterozygote advantage is unlikely because the Type A + B phenotype is rare and was found in less than 10% of individuals sampled. Sexual antagonism is also not possible because there are no differences between males and females in venom phenotypes. Due to the lack of ecological data, we were unable to distinguish between frequency-dependent selection, density dependent selection, and local adaptation over time and space on the venom type dichotomy in *C. scutulatus*.

Linkage disequilibrium based on the genetic architecture of polymorphisms or similar selective regimes can allow recombination through most of the genome but not the region(s) controlling the trait [72]. Loci involved form a 'supergene' complex that can be maintained in the face of gene exchange [10, 73, 74]. Maladapted individuals are quickly removed from the population and intermediate phenotypes would be rare due to low fitness [2]. The venom phenotype dichotomy in *C. scutulatus* has characteristics of a supergene complex [72]. The integrated phenotype (Type A + B) was rare (11 of 216 individuals - 5.1%) and the two subunits of MTX are close to each other in the genome [75] and they are always found expressed together [32, 47]. The 15-17 SVMP loci are also located close to each other and Type A individuals have a deletion

in the genome of all but four of them [11]. The SVMP and MTX regions are likely not physically linked but it is possible for different regions in the genome to be in linkage disequilibrium and have the same response to balancing selection [6]. When comparing the dichotomy across rattlesnake species, the lack of phylogenetic signal also supports balancing selection [11] and may represent a trans-species polymorphism [76, 77].

Conclusion

We found that the geographic pattern of the three venom phenotypes in *C. scutulatus* is much more complex than previously hypothesized. We were able to rule out venom type being fixed in genetic populations because only one population is potentially monomorphic. Additionally, only one environmental variable, temperature of the coldest month, was both statistically different between Type A and Type B venom and explained a significant amount of variability in our environmental niche models. Overall, the models were not equal but were significantly similar and ruled out elevation as being different between venom types. The lack of genetic and environmental correlation with venom type coupled with the morphological difference in interfang distance suggests that there may be intense local selection between Mojave Rattlesnakes and their prey through venom resistance or behavioral modifications. The presence of both phenotypes may also be type of historical contingency combined with selection. By using range-wide sampling of *C. scutulatus*, we were able to rule out broad genetic patterns and clear environmental differences as explaining the high level of venom variability found in this species making it possible that their persistence is being maintained through balancing selection.

Acknowledgments

Support and assistance in the field was provided by J. Aceves, J. Adams, I. Ahumada, L. Badillo, M. Bernard, B. Bloom, T. Burkhardt, J. Castañeda-Gaytán, B. Chambers, H. Dahn, D.

Deem, T. Dimler, B. Eaton, R. Engeldorf, E. Fanti, M. Feldner, T. Fisher, E. Flores, H. Franz, F. García, R. Govreau, C. Grünwald, A. Gutierrez, F. Hernandez, M. Holding, J. Houck, W. Howell, C. Howell, T. Jones, J. Jones, T. LaDuc, P. Lindsey, M. Linsalata, J. Lock, C. Mallery, M. Margres, C. May, S. May, R. Mayerhofer, E. McCormick, C. McMartin, J. McNally, C. Moreno, D. Ortiz, A. Owens, T. Petty, M. Price, A. Quillen, G. Quintero, R. Ramirez, R. Rautsaw, E. Rivas, M. Rodriguez, B. Rodriguez, A. Salas, G. Salmon, J. Sigala, J. Slone, R. Solis, D. Speckin, S. Stevens, R. Swanson, K. Swanson, E. Swanson, G. Territo, B. Townsend, C. Trumbower, S. Valenzuela, K. VanSooy, I. Villalobos, C. Vratil, D. Weber, K. Wray, W. Wüster, G. Zancolli, and G. Zavala. We thank M. Hogan, M. Margres, J. McGivern, and M. Ward from the Rokyta Lab and M. Seavy (FSU Department of Biological Science Analytical Lab) for assistance with RP-HPLC. Comments and suggestions were given by E. Hofmann, R. Rautsaw, and A. Savage which greatly improved the quality of this manuscript. Financial support was generously provided by the National Science Foundation to C.L.P (DUE 1161228 which funded J.L.S. and A.J.M. and DEB 1638879), to D.R.R. (DEB 1145987 and DEB 1638902), to T.A.C. and S.P.M (DEB 1655571), and to D.R.S and T.A.C. (DDIG Grant DEB-1501886). Additional support was provided by UC Mexus to C.L.S. (CN-11-548), University of Texas at Arlington through faculty startup to T.A.C., Consejo Nacional de Ciencia y Tecnologia to M.B. (CONACYT 247437), Phi Sigma Beta Phi Chapter (D.R.S), Prairie Biotic Research Inc. (J.L.S.), Sigma Xi Grants-in-aid-of-research (J.L.S.), SnakeDays Research Grant (J.L.S.), the Southwestern Association of Naturalists McCarley Research Grant (J.L.S., A.J.M and M.B.), and the Theodore Roosevelt Memorial Fund through the American Museum of Natural History (A.J.M and J.L.S.).

References

1. Tigano, A.; Friesen, V.L. Genomics of local adaptation with gene flow. *Molecular Ecology* **2016**, *25*, 2144–2164.

2. Pinho, C.; Hey, J. Divergence with gene flow: models and data. *Annual Review of Ecology, Evolution, and Systematics* **2010**, *41*, 215–30.
3. Cox, C.L.; Davis Rabosky, A.R. Spatial and temporal drivers of phenotypic diversity in polymorphic snakes. *The American Naturalist* **2013**, *182*, E40–E57.
4. Moody, K.N.; Hunter, S.N.; Childress, M.J.; Blob, R.W.; Schoenfuss, H.L.; Blum, M.J.; Ptacek, M.B. Local adaptation despite high gene flow in the waterfall-climbing Hawaiian Goby, *Sicyopterus stimpsoni*. *Molecular Ecology* **2015**, *24*, 545–563.
5. Hittinger, C.T.; Gonçalves, P.; Sampaio, J.P.; Dover, J.; Johnston, M.; Rokas, A. Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature* **2010**, *464*, 54–58.
6. Llaurens, V.; Whibley, A.; Joron, M. Genetic architecture and balancing selection: the life and death of differentiated variants. *Molecular Ecology* **2017**, *26*, 2430–2448.
7. Fijarczyk, A.; Babik, W. Detecting balancing selection in genomes: limits and prospects. *Molecular Ecology* **2015**, *24*, 3529–3545.
8. Charlesworth, D. Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genetics* **2006**, *2*, e64.
9. Greene, J.S.; Brown, M.; Dobosiewicz, M.; Ishida, I.G.; Macosko, E.Z.; Zhang, X.; Butcher, R.A.; Cline, D.J.; McGrath, P.T.; Bargmann, C.I. Balancing selection shapes density-dependent foraging behaviour. *Nature* **2016**, *539*, 254–258.
10. Joron, M.; Frezal, L.; Jones, R.T.; Chamberlain, N.L.; Lee, S.F.; Haag, C.R.; Whibley, A.; Becuwe, M.; Baxter, S.W.; Ferguson, L.; Wilkinson, P.A.; Salazar, C.; Davidson, C.; Clark, R.; Quail, M.A.; Beasley, H.; Glithero, R.; Lloyd, C.; Sims, S.; Jones, M.C.; Rogers, J.; Jiggins, C.D.; Ffrench-Constant, R.H. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **2011**, *477*.
11. Dowell, N.L.; Giorgianni, M.W.; Griffin, S.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. Extremely divergent haplotypes in two toxin gene complexes encode

- alternative venom types within rattlesnake species. *Current Biology* **2018**, *28*, 1016–1026.
12. Mackessy, S.P. Venom composition in rattlesnakes: trends and biological significance. In *The Biology of Rattlesnakes*; Hayes, W.K.; Beaman, K.R.; Cardwell, M.D.; Bush, S.P., Eds.; Loma Linda University Press: Loma Linda, CA, 2008; pp. 495–510.
 13. Rokyta, D.R.; Wray, K.P.; McGivern, J.J.; Margres, M.J. The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the Timber Rattlesnake (*Crotalus horridus*). *Toxicon* **2015**, *98*, 34–48.
 14. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex cocktails: the evolutionary novelty of venoms. *Trends in ecology & evolution* **2013**, *28*, 219–29.
 15. Holmes, I.A.; Grundler, M.R.; Davis Rabosky, A.R. Predator perspective drives geographic variation in frequency-dependent polymorphism. *The American Naturalist* **2017**, *190*, E78–E93.
 16. Daltry, J.C.; Wüster, W.; Thorpe, R.S. Diet and snake venom evolution. *Nature* **1996**, *379*, 537–540.
 17. Smiley-Walters, S.A.; Farrell, T.M.; Gibbs, H.L. Evaluating local adaptation of a complex phenotype: reciprocal tests of Pigmy Rattlesnake venoms on treefrog prey. *Oecologia* **2017**, *184*, 739–748.
 18. Smiley-Walters, S.A.; Farrell, T.M.; Gibbs, H.L. The importance of species: Pigmy Rattlesnake venom toxicity differs between native prey and related non-native species. *Toxicon* **2018**, *144*, 42–47.
 19. Thomas, R.G.; Pough, F.H. The effect of rattlesnake venom on digestion of prey. *Toxicon* **1979**, *17*, 221–228.
 20. Mackessy, S.P. Venom ontogeny in the Pacific Rattlesnakes *Crotalus viridis helleri* and *C. v. oregonus*. *Copeia* **1988**, *1988*, 92–101.
 21. Mackessy, S.P. Evolutionary trends in venom composition in the Western Rattlesnakes (*Crotalus viridis sensu lato*): toxicity vs. tenderizers. *Toxicon* **2010**, *55*, 1463–74.

22. McCue, M.D. Prey envenomation does not improve digestive performance in Western Diamondback Rattlesnakes (*Crotalus atrox*). *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* **2007**, *307*, 568–577.
23. Chu, C.W.; Tsai, T.S.; Tsai, I.H.; Lin, Y.S.; Tu, M.C. Prey envenomation does not improve digestive performance in Taiwanese pit vipers (*Trimeresurus gracilis* and *T. stejnegeri*). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* **2009**, *152*, 579–585.
24. Glenn, J.J.L.; Straight, R.C.; Wolfe, M.M.C.; Hardy, D.D.L. Geographical variation in *Crotalus scutulatus scutulatus* (Mojave Rattlesnake) venom properties. *Toxicon* **1983**, *21*, 119–130.
25. Margres, M.J.; Wray, K.P.; Seavy, M.; McGivern, J.J.; Sanader, D.; Rokyta, D.R. Phenotypic integration in the feeding system of the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*). *Molecular Ecology* **2015**, *24*, 3405–3420.
26. Glenn, J.; Straight, R. Mojave Rattlesnake *Crotalus scutulatus scutulatus* venom: variation in toxicity with geographical origin. *Toxicon* **1978**, *16*, 81–84.
27. Weinstein, S.A.; Minton, S.A.; Wilde, C.E. The distribution among ophidian venoms of a toxin isolated from the venom of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*). *Toxicon* **1985**, *23*, 825–844.
28. Wooldridge, B.; Pineda, G.; Banuelas-Ornelas, J.; Dagda, R.; Gasanov, S.; Rael, E.; Lieb, C. Mojave Rattlesnakes (*Crotalus scutulatus scutulatus*) lacking the acidic subunit DNA sequence lack Mojave toxin in their venom. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **2001**, *130*, 169–179.
29. Sánchez, E.E.; Galán, J.A.; Powell, R.L.; Reyes, S.R.; Soto, J.G.; Russell, W.K.; Russell, D.H.; Pérez, J.C. Disintegrin, hemorrhagic, and proteolytic activities of Mohave Rattlesnake, *Crotalus scutulatus scutulatus* venoms lacking Mojave toxin. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **2005**, *141*, 124–132.

30. Martinez, M.; Rael, E.D.; Maddux, N.L. Isolation of a hemorrhagic toxin from Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) venom. *Toxicon* **1990**, *28*, 685–694.
31. Wilkinson, J.A.; Glenn, J.L.; Straight, R.C.; Sites, J.W. Distribution and genetic variation in venom A and B populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Herpetologica* **1991**, *47*, 54 – 68.
32. Borja, M.; Neri-Castro, E.; Castañeda-Gaytán, G.; Strickland, J.L.; Parkinson, C.L.; Castañeda-Gaytán, J.; Ponce-López, R.; Lomonte, B.; Olvera-Rodríguez, A.; Alagón, A.; Pérez-Morales, R.; Olvera-Rodríguez, A.; Alagón, A.; Pérez-Morales, R. Biological and proteolytic variation in the venom of *Crotalus scutulatus scutulatus* from Mexico. *Toxins* **2017**, *10*, 35.
33. Glenn, J.L.; Straight, R.C. Intergradation of two different venom populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Toxicon* **1989**, *27*, 411–8.
34. Minton, S.A. Observations on the amphibians and reptiles of the Big Bend region of Texas. *The Southwestern Naturalist* **1958**, *3*, 28–54.
35. Bieber, A.L.; Tu, T.; Tu, A.T. Studies of an acidic cardiotoxin isolated from the venom of Mojave Rattlesnake (*Crotalus scutulatus*). *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1975**, *400*, 178–188.
36. Borja, M.; Castañeda, G.; Espinosa, J.; Neri, E.; Carbajal, A.; Clement, H.; García, O.; Alagon, A. Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) with Type B venom from Mexico. *Copeia* **2014**, *2014*, 7–13.
37. Dobson, J.; Yang, D.C.; op den Brouw, B.; Cochran, C.; Huynh, T.; Kurrupu, S.; Sánchez, E.E.; Massey, D.J.; Baumann, K.; Jackson, T.N.N.; Nouwens, A.; Josh, P.; Neri-Castro, E.; Alagón, A.; Hodgson, W.C.; Fry, B.G. Rattling the border wall: pathophysiological implications of functional and proteomic venom variation between Mexican and US subspecies of the desert rattlesnake *Crotalus scutulatus*. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **2017**.

38. Massey, D.J.; Calvete, J.J.; Sánchez, E.E.; Sanz, L.; Richards, K.; Curtis, R.; Boesen, K. Venom variability and envenoming severity outcomes of the *Crotalus scutulatus scutulatus* (Mojave Rattlesnake) from southern Arizona. *Journal of Proteomics* **2012**, *75*, 2576–2587.
39. Strickland, J.; Mason, A.; Rokyta, D.; Parkinson, C. Phenotypic variation in Mojave Rattlesnake (*Crotalus scutulatus*) venom is driven by four toxin families. *Toxins* **2018**, *10*, 135.
40. Schield, D.R.; Card, D.C.; Adams, R.H.; Corbin, A.; Jezkova, T.; Hales, N.; Meik, J.M.; Spencer, C.L.; Smith, L.; Campillo-Garcia, G.; Bouzid, N.; Strickland, J.L.; Parkinson, C.L.; Flores-Villela, O.; Mackessy, S.P.; Castoe, T.A. Cryptic genetic diversity, population structure, and gene flow in the Mojave Rattlesnake (*Crotalus scutulatus*). *Molecular Phylogenetics and Evolution* **2018**.
41. Dagda, R.K.; Gasanov, S.; De La Oiii, Y.; Rael, E.D.; Lieb, C.S. Genetic basis for variation of metalloproteinase-associated biochemical activity in venom of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*). *Biochemistry Research International* **2013**, *2013*, 251474.
42. Rohland, N.; Reich, D. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research* **2012**, *22*, 939–46.
43. Faircloth, B. Protocol: preparation of an AMPure XP substitute (AKA Serapure) **2014**. p. doi: 10.6079/J9MW2F26.
44. Margres, M.J.; McGivern, J.J.; Wray, K.P.; Seavy, M.; Calvin, K.; Rokyta, D.R. Linking the transcriptome and proteome to characterize the venom of the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*). *Journal of Proteomics* **2014**, *96*, 145–158.
45. Rokyta, D.R.; Wray, K.P.; Lemmon, A.R.; Lemmon, E.M.; Caudle, S.B. A high-throughput venom-gland transcriptome for the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*) and evidence for pervasive positive selection across toxin classes. *Toxicon* **2011**, *57*, 657–671.
46. Smith, C.F.; Mackessy, S.P. The effects of hybridization on divergent venom phenotypes:

- characterization of venom from *Crotalus scutulatus scutulatus* x *Crotalus oreganus helleri* hybrids. *Toxicon* **2016**, *120*, 110–123.
47. Zancolli, G.; Baker, T.; Barlow, A.; Bradley, R.; Calvete, J.; Carter, K.; de Jager, K.; Owens, J.; Price, J.; Sanz, L.; Scholes-Higham, A.; Shier, L.; Wood, L.; Wüster, C.; Wüster, W. Is hybridization a source of adaptive venom variation in rattlesnakes? A test, using a *Crotalus scutulatus* x *viridis* hybrid zone in southwestern New Mexico. *Toxins* **2016**, *8*, 188.
 48. Arevalo, E.; Davis, S.K.; Sites, J.W. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* Complex (Phrynosomatidae) in Central Mexico. *Systematic Biology* **1994**, *43*, 387–418.
 49. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **2013**, *30*, 772–780.
 50. Lanfear, R.; Frandsen, P.B.; Wright, A.M.; Senfeld, T.; Calcott, B. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* **2016**, *34*, msw260.
 51. Guindon, S.; Dufayard, J.F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **2010**, *59*, 307–321.
 52. Bouckaert, R.; Heled, J.; Kühnert, D.; Vaughan, T.; Wu, C.H.; Xie, D.; Suchard, M.A.; Rambaut, A.; Drummond, A.J. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **2014**, *10*, e1003537.
 53. Rambaut, A.; Suchard, M.A.; Xie, D.; Drummond, A.J. Tracer v1.6., 2010.
 54. Aird, S.D.; Kruggel, W.G.; Kaiser, I.I. Amino acid sequence of the basic subunit of Mojave toxin from the venom of the Mojave Rattlesnake (*Crotalus s. scutulatus*). *Toxicon* **1990**, *28*, 669–673.

55. Mackessy, S.P. Fibrinogenolytic proteases from the venoms of juvenile and adult Northern Pacific Rattlesnakes (*Crotalus viridis oreganus*). *Comparative Biochemistry and Physiology – Part B: Comparative Biochemistry* **1993**, *106*, 181–189.
56. Phillips, S.J.; Dudik, M.; Schapire, R.E. Maxent software for modeling species niches and distributions, 2018.
57. Hijmans, R.J.; Phillips, S.; Leathwick, J.; Elith, J. dismo: Species distribution modeling. R package version 1.1-4, 2017.
58. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **2006**, *190*, 231–259.
59. Fick, S.E.; Hijmans, R.J. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **2017**, *37*, 4302–4315.
60. Dormann, C.F.; Elith, J.; Bacher, S.; Buchmann, C.; Carl, G.; Carré, G.; Marquéz, J.R.G.; Gruber, B.; Lafourcade, B.; Leitão, P.J.; Münkemüller, T.; McClean, C.; Osborne, P.E.; Reineking, B.; Schröder, B.; Skidmore, A.K.; Zurell, D.; Lautenbach, S. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **2013**, *36*, 27–46.
61. Warren, D.L.; Glor, R.E.; Turelli, M. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **2008**, *62*, 2868–83.
62. Heibl, C.; Calenge, C. phyloclim: Integrating phylogenetics and climatic niche modeling, 2015.
63. Githens, T.; George, I. Comparative studies on the venoms of certain rattlesnakes. *Bulletin of the Antivenin Institute of America* **1931**, *5*, 31–34.
64. Cardwell, M.D. Mohave Rattlesnakes *Crotalus scutulatus* (Kennicott 1861). In *Rattlesnakes of Arizona Species Accounts and Natural History, 1*; Schuett, G.W.; Feldner, M.; Smith, C.; Reiserer, R., Eds.; ECO Herpetological Publishing: Rodeo, NM, 2016; pp. 563–605.

65. Higham, T.E.; Clark, R.W.; Collins, C.E.; Whitford, M.D.; Freymiller, G.A. Rattlesnakes are extremely fast and variable when striking at kangaroo rats in nature: three-dimensional high-speed kinematics at night **2016**.
66. Clark, R.W.; Tangco, S.; Barbour, M.A. Field video recordings reveal factors influencing predatory strike success of free-ranging rattlesnakes (*Crotalus spp.*). *Animal Behaviour* **2012**, *84*, 183–190.
67. Schraft, H.A.; Clark, R.W. Kangaroo rats change temperature when investigating rattlesnake predators. *Physiology & Behavior* **2017**, *173*, 174–178.
68. Freymiller, G.A.; Whitford, M.D.; Higham, T.E.; Clark, R.W. Recent interactions with snakes enhance escape performance of desert kangaroo rats (Rodentia: Heteromyidae) during simulated attacks. *Biological Journal of the Linnean Society* **2017**, *122*, 651–660.
69. Pomento, A.M.; Perry, B.W.; Denton, R.D.; Gibbs, H.L.; Holding, M.L. No safety in the trees: local and species-level adaptation of an arboreal squirrel to the venom of sympatric rattlesnakes. *Toxicon* **2016**, *118*, 149–155.
70. Holding, M.L.; Drabeck, D.H.; Jansa, S.A.; Gibbs, H.L. Venom resistance as a model for understanding the molecular basis of complex coevolutionary adaptations. *Integrative and Comparative Biology* **2016**, *56*, 1032–1043.
71. LaBonte, J.; Welch, K.; Suarez, R. Digestive performance in neonatal Southern Pacific Rattlesnakes (*Crotalus oreganus helleri*). *Canadian Journal of Zoology* **2011**, *89*, 705–713.
72. Thompson, M.; Jiggins, C. Supergenes and their role in evolution. *Heredity* **2014**, *11320*, 1–8.
73. Li, J.; Cocker, J.M.; Wright, J.; Webster, M.A.; McMullan, M.; Dyer, S.; Swarbreck, D.; Caccamo, M.; van Oosterhout, C.; Gilmartin, P.M. Genetic architecture and evolution of the S locus supergene in *Primula vulgaris*. *Nature Plants* **2016**, *2*, 16188.
74. Tuttle, E.M.; Bergland, A.O.; Korody, M.L.; Warren, W.C.; Gonser, R.A.; Balakrishnan Correspondence, C.N.; Brewer, M.S.; Newhouse, D.J.; Minx, P.; Stager, M.; Betuel,

- A.; Cheviron, Z.A.; Balakrishnan, C.N. Divergence and functional degradation of a sex chromosome-like supergene. *Current Biology* **2016**, *26*, 344–350.
75. Dowell, N.L.; Giorgianni, M.W.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. The deep origin and recent loss of venom toxin genes in rattlesnakes. *Current Biology* **2016**, *26*, 2434–2445.
76. Klein, J. Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. *Human Immunology* **1987**, *19*, 155–162.
77. Klein, J.; Sato, A.; Nagl, S.; O’huigín, C. Molecular trans-species polymorphism. *Annual Review of Ecology and Systematics* **1998**, *29*, 1–21.

CHAPTER 3: PHENOTYPIC VARIATION IN MOJAVE RATTLESNAKE (*CROTALUS SCUTULATUS*) VENOM IS DRIVEN BY FOUR TOXIN FAMILIES²

Abstract

Phenotypic diversity generated through altered gene expression is a primary mechanism facilitating evolutionary response in natural systems. By linking the phenotype to genotype through transcriptomics, it is possible to determine what changes are occurring at the molecular level. High phenotypic diversity has been documented in rattlesnake venom, which is under strong selection due to its role in prey acquisition and defense. Rattlesnake venom can be characterized by the presence (Type A) or absence (Type B) of a type of neurotoxic phospholipase A₂ (PLA₂), such as Mojave toxin, that increases venom toxicity. Mojave Rattlesnakes (*Crotalus scutulatus*), represent this diversity as both venom types are found within this species and within a single panmictic population in the Sonoran Desert. We used comparative venom gland transcriptomics of nine specimens of *C. scutulatus* from this region to test whether expression differences explain diversity within and between venom types. Type A individuals expressed significantly fewer toxins than Type B individuals owing to the diversity of C-type lectins (CTLs) and snake venom metalloproteinases (SVMPs) found in Type B animals. As expected, both subunits of Mojave toxin were exclusively found in Type A individuals but we found high diversity in four additional PLA₂s that was not associated with a venom type. Myotoxin *a* expression and toxin number variation was not associated with venom type, and myotoxin *a* had the highest range of expression of any toxin class. Our study represents the most comprehensive transcriptomic profile of the venom

²Chapter Three was previously published. The reference for this paper is as follows:
Strickland JL, Mason AJ, Rokyta DR, Parkinson CL. 2018. Phenotypic variation in Mojave Rattlesnake (*Crotalus scutulatus*) venom is driven by four toxin families. *Toxins* 10:135. doi:10.3390/toxins10040135

type dichotomy in rattlesnakes and *C. scutulatus*. Even intra-specifically, Mojave Rattlesnakes showcase the diversity of snake venoms and illustrate that variation within venom types blurs the distinction of the venom dichotomy.

Keywords: C-type lectins; Hemorrhagic; Mojave toxin; Myotoxins *a*; Neurotoxic; Phospholipase A₂s; RNA-seq; Snake Venom Metalloproteinases

Introduction

Differential gene expression is a primary component with respect to the genotype and phenotype that facilitates rapid evolutionary response in the face of changing environmental pressures by generating phenotypic diversity [1]. Comparative transcriptomics has emerged as the tool to understand these responses by linking the phenotype to the genotype through mRNA sequencing. However, the molecular mechanisms underlying phenotypic divergence are difficult to determine because there is usually no one-to-one link between the genotype and phenotype due to pleiotropic and epistatic effects [2, 3]. Venom is an exception to this because it is a complex trait that is highly tractable from the gene being expressed to the final protein product [4]. Venom is under strong selection as it aids in prey acquisition and/or serves as a predator deterrent [5]. Changes within the venom phenotype occur through regulatory shifts in protein expression [6], through loss of specific genes [7], duplication [8], and point mutations [9]. Transcriptomics cannot detect all possible mechanisms resulting in phenotypic diversity, particularly with regard to genes in the genome that are not expressed or have multiple copies [10]. However, for those genes that are expressed, transcriptomics offer an effective means of linking sequence-based and regulatory variation to changes in a composite phenotype [11]. With the increase in availability of high-throughput proteomic, transcriptomic, and genomic resources, rattlesnakes and their venom have become a model system to understand these mechanisms as they exhibit high phenotypic diversity [12–15].

Rattlesnake venoms can be broadly characterized by the presence or absence of heterodimeric phospholipases A₂ (PLA₂s). Type A venoms contain this PLA₂ which is a β -neurotoxin responsible for highly toxic venom in individuals that express it. These venoms also have little hemorrhagic activity due to low expression of snake venom metalloproteinases (SVMPs). Type B venoms lack the neurotoxic PLA₂ and are also characterized by high hemorrhagic activity due to high expression of SVMPs [12, 13, 15–17]. When present, the PLA₂ acts presynaptically to disrupt the nervous system and both the acidic and basic subunits of the heterodimer must be expressed in the venom for the neurotoxic effect to occur. The origin of this toxin and its effect is due to a single nucleotide substitution that allowed the interaction of the two subunits to be energetically favorable [9], and no evidence exists of this PLA₂ being in the genome and not being expressed proteomically [18–20]. Examples include Mojave toxin (MTX) in *Crotalus scutulatus* and its close relatives, Sistruxin in *Sistrurus catenatus*, Crotoxin in the *Crotalus durissus* complex, and Canebrake toxin in *Crotalus horridus* [21–25]. Of the 48 species of rattlesnakes (*Crotalus* and *Sistrurus*) currently recognized [26], 38 are considered as Type B, one is Type A, and 9 are documented as polymorphic [25].

The lack of a phylogenetic pattern in venom phenotypes exhibited by rattlesnakes has hampered our understanding of the evolution of this dichotomy. The most recent hypothesis, based on genomic sequencing and ancestral state reconstruction, is that the ancestral state was neurotoxic (Type A) and lineages no longer possessing the neurotoxic PLA₂ (hereafter Mojave toxin/MTX) have lost it [7]. Dowell *et al.* [7] examined three rattlesnake species (*Crotalus adamanteus*, *C. atrox*, and *C. scutulatus*) and found that these three species have different sets of PLA₂s based on venom type, which is supported by current transcriptomic evidence. Rokyta *et al.* [27] compared the venom gland transcriptomes of one individual from a Type B species (*C. adamanteus*) to one Type A species (*C. horridus*) and found different sets of PLA₂s. This was further supported by Rokyta *et al.* [28] when they compared the venom gland transcriptomes of one Type B individual and one Type A individual of *C. horridus* to examine intraspecific variation.

They found different sets of PLA₂s expressed in the transcriptome between Type A and Type B *C. horridus*. Additionally, both studies found evidence to support the dichotomy. Type A individuals had simpler venom in that fewer toxins were expressed and the primary difference was the trade-off between MTX and SVMPs, but they also found differences in C-type lectins (CTLs) and myotoxin *a* (hereafter MYO or myotoxins) [27, 28]. Like *C. horridus*, *C. scutulatus* (Dowell *et al.* [7]’s neurotoxic/Type A representative), have well documented intraspecific venom variability corresponding to the Type A and Type B venom phenotypes throughout their distribution [20, 21, 29–33].

Mojave rattlesnakes (*C. scutulatus*) are known to be present from the deserts of the southwestern United States and as far south as the state of Puebla in Mexico (Figure 3.1). *Crotalus scutulatus* is comprised of three phylogeographic lineages [34]. The basal lineage includes the subspecies *Crotalus scutulatus salvini*, which has Type A venom and is located at the southern end of the distribution [12, 33]. The remaining two lineages are *Crotalus scutulatus scutulatus* distributed in the Chihuahuan Desert and Sonoran/Mojave Deserts (hereafter, Sonoran), respectively [34]. Both Type A and Type B venoms are found in the latter two lineages, and although rare, snakes possessing a Type A + B phenotype (highly expressing both MTX and SVMPs) have also been documented [18, 20, 31, 35, 36].

The venom phenotype complexity in *C. scutulatus* is best documented within the Sonoran lineage [37–39]. Proteomic differences in the venom seem to be geographically fixed, but intergradation occurs where Type A and Type B ranges come into contact [12, 31, 35, 39, 40]. In addition to the dichotomy between PLA₂s and SVMPs, variability in other toxins has been documented [36, 39]. Particularly, Massey *et al.* [39] described variation in myotoxins in one small area of the distribution where some individuals had ~25% of their proteome made up of myotoxins. This finding prompted them to suggest further dividing venom types within *C. scutulatus* into Types A–F to account for variation between PLA₂s, SVMPs, and myotoxins, and hypothesized that myotoxins occur when the Type A and Type B phenotypes come into

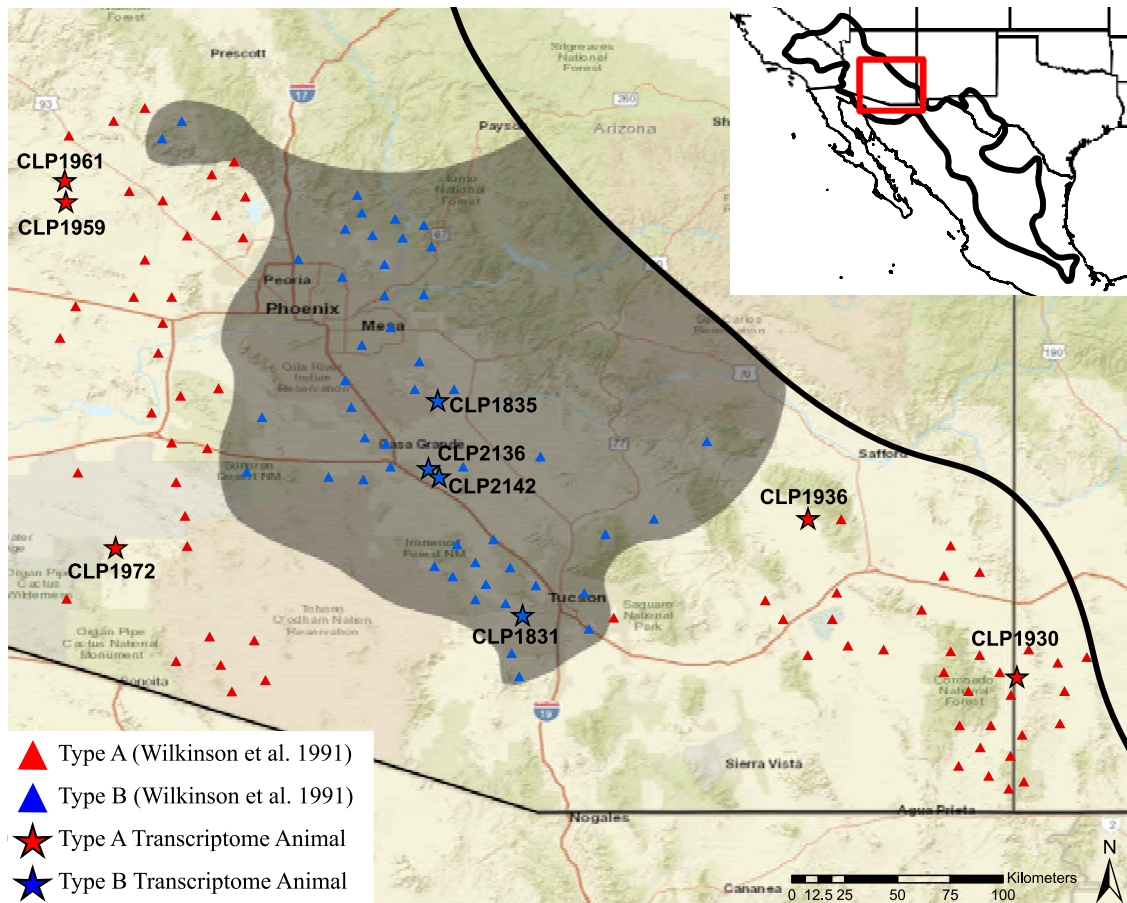


Figure 3.1: Distribution of Type A and Type B venom in Arizona based on data from Wilkinson et al. (triangles) and the nine transcriptome animals sequenced (stars). The shaded area is the estimated distribution of Type B venom and the dark black line is the outline of the estimated distribution of *Crotalus scutulatus*. CLP: Christopher L. Parkinson field number.

contact [39]. No phylogeographic structure exists and high gene flow occurs between individuals regardless of venom type within this region based on allozymes [31], mitochondrial (ND4), and nuclear (RADSeq) data [34], so genetic recombination at the contact zone is possible.

To examine the role of differential expression in the evolution of venom phenotypes, we used Sonoran Desert Mojave Rattlesnakes because they are a microcosm for the diversity found in rattlesnakes. Within the Sonoran lineage, high venom diversity exists and six phenotypes are described [36, 39] with Type A and Type B venom phenotypes being geographically fixed despite high gene flow [31, 34]. Moreover, the ancestral lineage within Mojave Rattlesnakes, *C. scutulatus*

salvini, is neurotoxic [33, 34], as is the hypothesized ancestral rattlesnake [7]. Thus, mechanisms underlying the presence or absence of Mojave toxin in other rattlesnake species may become more apparent by focusing on *C. scutulatus*. Through comparative venom-gland transcriptomics, we link the patterns found genomically and the diversity identified proteomically to test: (1) whether myotoxin expression is localized to contact areas between individuals with Type A and Type B venom; (2) if expression patterns within and between venom phenotypes are consistent among species and individuals; and (3) whether individuals with Type A and Type B venom will express distinct sets of PLA₂s as hypothesized by Dowell *et al.* [7]. In this pursuit, we present the most extensive transcriptomic sampling to date of *Crotalus scutulatus* and the A/B venom dichotomy.

Materials and methods

Ethics statement

Scientific collecting permits were issued by the New Mexico Department of Game and Fish (3563, 3576) and the State of Arizona Game and Fish Department (SP628489, SP673390, SP673626, SP715023). All interactions with animals were approved by the University of Central Floridas Institutional Animal Care and Use Committee under protocol 13–17 W and followed the American Society of Ichthyologists and Herpetologists ethical guidelines.

Sample collection

In the summers of 2013–2015, we collected *Crotalus scutulatus* from the Sonoran Desert in Arizona and New Mexico. In sampling, we targeted areas described by Wilkinson *et al.* [31] to have individuals with either Type A or Type B venoms (Figure 3.1). We collected a total of 42 individuals and processed them in preparation for venom gland transcriptome sequencing. First, we collected venom from each animal by tubing the individual in polycarbonate tubes (Get Hooked L.L.C., Sanford, FL, USA) and allowed them to move through the tube until just their

head was protruding. Then, they were able to voluntarily bite a sterile collection cup covered in parafilm. Venom was collected, vacuum dried, and stored at -80°C for future use. Four days after venom was collected and transcription was maximized [41], we sacrificed the animal using an intracoelomic injection of sodium pentobarbital (100 mg/kg). We then removed the venom glands and stored them separately in either RNAlater (Thermo Fisher Scientific, Waltham, MA, USA) at 4°C overnight or in liquid nitrogen before moving to -80°C for long-term storage. Each specimen was fixed in 10% buffered formalin for five days and then transferred to 70% ethanol and deposited in a natural history museum (Table 3.1).

Table 3.1: Sample information for the nine specimens of *C. scutulatus* sequenced from the U.S. Sonoran Desert. ASNHC = Angelo State Natural History Collection, San Angelo, Texas; ASU = Arizona State University Natural History Museum, Tempe, Arizona; CLP = Christopher L. Parkinson field number.

Specimen ID	Museum ID	Venom Type	Sex	SVL (mm)	Mass (g)	State	County	Sequencing Platform	Read Pairs	Merged Reads	BioSample ¹ Accession
CLP1930	ASNHC14997	A	F	724	195	NM	Hidalgo	MiSeq	15,649,085	13,083,925	SAMN08596271
CLP1936	ASU36035	A	F	441	48	AZ	Graham	MiSeq	20,835,668	18,182,033	SAMN08596272
CLP1959	ASU36061	A	M	730	204	AZ	Yavapai	MiSeq	20,577,779	17,091,930	SAMN08596273
CLP1961	ASU36062	A	M	564	90	AZ	Yavapai	HiSeq	11,929,639	10,061,262	SAMN08596274
CLP1972	ASU36092	A	M	635	126	AZ	Pima	MiSeq	13,168,704	11,521,499	SAMN08596275
CLP1831	ASU36089	B	F	795	344	AZ	Pima	HiSeq	15,448,552	13,526,047	SAMN08596267
CLP1835	ASU36102	B	F	685	146	AZ	Pinal	HiSeq	16,271,477	14,210,557	SAMN08596269
CLP2136	ASU36103	B	M	1030	627	AZ	Pinal	MiSeq	7,771,613	6,864,270	SAMN08596277
CLP2142	ASU36104	B	M	775	262	AZ	Pinal	MiSeq	10,039,268	8,893,097	SAMN08596278

¹ National Center for Biotechnology Information under BioProject PRJNA88989.

Venom type determination

We determined the venom type of each individual using reverse-phased high performance liquid chromatography (RP-HPLC) similar to Margres *et al.* [42]. We resuspended the dry venom in water and removed insoluble material via centrifugation. We determined the concentration of the venom utilizing the Qubit Protein Assay (Thermo Fisher Scientific) following the manufacturer's protocol on a Qubit 3.0 Fluorometer (Thermo Fisher Scientific). We then injected 100 μg of venom onto a Jupiter C18 column (250 \times 2 mm; Phenomenex, Torrance, CA, USA) using two solvents: A = 0.1% trifluoroacetic acid (TFA) in water and B = 0.075% TFA in acetonitrile. RP-HPLC

was conducted on a Beckman System Gold HPLC (Beckman Coulter, Fullerton, CA, USA). The gradient began with 95% A and 5% B for 5 min followed by a 1% per minute linear gradient to 25% B. This was followed by a 0.25% per minute linear gradient to 55% B, a 2% per minute linear gradient to 75% B, a 14% per minute linear gradient to 5% B, and then 5 min at the initial conditions all at a 0.2 mL/min flow rate. Total run time was 180 min for each sample and the effluent was monitored at 220 and 280 nm [43].

We were able to distinguish Type A and Type B venoms from each other based on the presence or absence of MTX and SVMPs based on previous RP-HPLC profiles in *C. scutulatus* [39] and under these conditions [28]. We selected a total of nine individuals, five with Type A venom and four with Type B venom (Figure 3.2), for venom gland transcriptome sequencing. The Type A individuals were selected to maximize the geographic breadth around the Type B zone in central Arizona (Figure 3.1).

Venom gland transcriptome sequencing

For each transcriptome animal, total RNA was extracted from the left and right venom glands independently using a TRIzol-based RNA extraction as previously described [43]. We first diced approximately 100 mg of venom gland tissue into small pieces and added it to 500 μ L of TRIzol (Invitrogen, Carlsbad, CA, USA). We aspirated the tissue and TRIzol through a 20-gauge needle at least 10 \times until the tissue was homogenized. We added 500 μ L of fresh TRIzol and added the entire solution to Phase Lock Gel Heavy tubes (5Prime 2302830, Quantabio, Beverly, MA, USA). We added 200 μ L of 20% chloroform, shook the solution, and centrifuged for 20 min after three minutes of incubation to separate the RNA. RNA was isolated with isopropyl alcohol precipitation overnight at -20 $^{\circ}$ C and washed with 75% ethanol. We estimated the concentration of total RNA using the Qubit RNA BR Assay (Thermo Fisher Scientific) and evaluated the quality and final concentration using an Agilent Bioanalyzer 2100 with an RNA 6000 Pico Kit (Agilent Technologies, Santa Clara, CA, USA) following the manufacturer's instructions. We

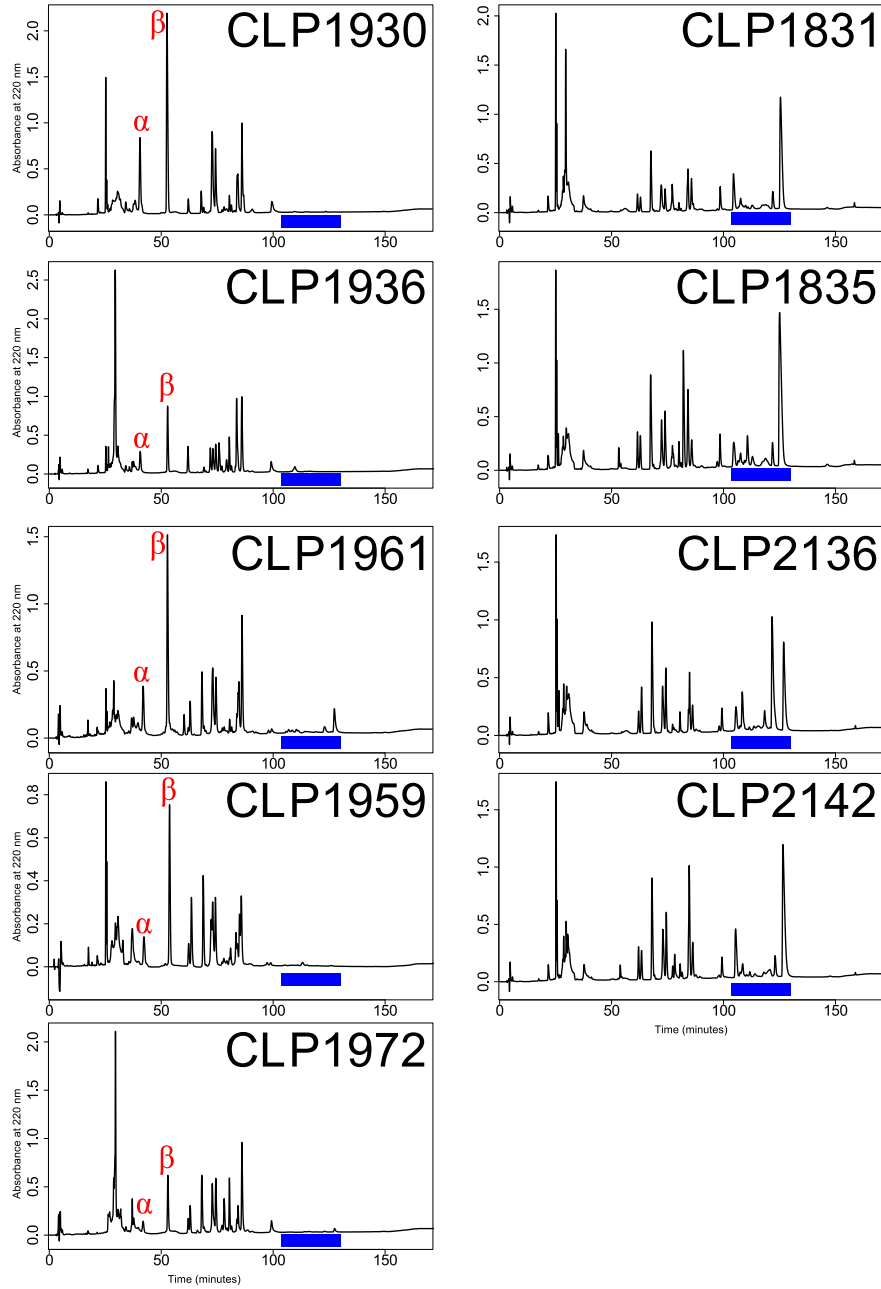


Figure 3.2: RP-HPLC profiles for the nine specimens of *C. scutulatus* selected for transcriptome sequencing. Type A individuals are in the left column and Type B individuals are in the right column. α = acidic subunit of Mojave toxin; β = basic subunit of Mojave toxin; Blue bar = region where SVMPs elute.

then combined RNA from both glands in equal concentration for library preparation as there is no difference in the transcriptome between the two glands [44].

We used 1 μ g of total RNA for mRNA isolation and cDNA library preparation [45, 46]. We used the New England Biolabs (NEB, Ipswich, MA, USA) NEBNext Poly(A) mRNA magnetic isolation kit (E7490S), the NEBNext Ultra RNA Library Prep Kit for Illumina (E7530), and the NEBNext Multiplex Oligos for Illumina (E7335 -Index primer set 1) following the manufacturer's protocols with a target mean insert size of 370 bp, a fragmentation time of 15.5 min, and 14 PCR cycles in the final enrichment step to yield the appropriate cDNA concentration. To do this, we isolated RNA with poly-A tails from the total RNA with oligo-dT beads and then immediately moved to first and second strand cDNA synthesis [47]. We used Agencourt AMPure XP PCR Purification Beads (Beckman Coulter) to purify DNA throughout the protocol. We estimated the concentration of our DNA library using the Qubit DNA BR Assay (Thermo Fisher Scientific) and evaluated the quality and final concentration using the Bioanalyzer with an HS DNA Kit following the manufacturer's instructions (Agilent Technologies). KAPA qPCR was conducted at the Florida State University Molecular Cloning Facility to determine the amplifiable concentration of each sample. For the samples sequenced on the Illumina MiSeq in the Florida State University Department of Biological Science DNA Sequencing Facility, we used the MiSeq Version 2 Reagent Kit and sequenced the individuals separately except for CLP 2136 and CLP 2142 which were pooled using equal amounts of amplifiable cDNA for each. If the sample was sequenced on the Illumina HiSeq 2500 in the Florida State University College of Medicine Translational Science Laboratory, we used the KAPA result to pool the samples to a final concentration of 10 nM so that each library was equally represented. We assessed the concentration and quality of the pooled DNA sample on the Bioanalyzer with the HS DNA Kit and performed an additional round of KAPA PCR before sequencing. Both the MiSeq and HiSeq runs were 150 bp paired-end reads (Table 3.1) [28].

Transcriptome assembly

We first used the Illumina quality filter to remove low-quality reads, and because the insert sizes for cDNA libraries were approximately 370 bp (170 bp without adapters), we were able to merge the remaining overlapping 150-nt paired-end reads at their 3' ends with PEAR v 0.9.6 [48]. Merged reads were then used for subsequent assemblies and analyses. We assembled venom gland transcriptomes as described in Rokyta *et al.* [44]. To maximize the number of transcripts recovered, we assembled the merged reads using DNASTar SeqMan NGen version 12.3.1 (DNASTAR Inc., Madison, WI, USA) with default settings except that only contigs with ≥ 200 assembled reads were kept and Extender [49]. We then compared all assembled transcripts through both methods to a venom toxin and nontoxin sequence database based on transcripts identified previously [27, 28, 43, 44, 49]. This database included 1047 toxins and 4516 nontoxins. Toxins were annotated if they matched at $\geq 90\%$ to the local database based on similarity in Cd-hit-est v.4.6 [50] and had at least 10X coverage across the transcript. The remaining unmatched contigs were compared via blastx v. 2.2.30+ searches (minimum e-value of 10^{-4}) to the curated Uniprot animal toxins database (downloaded 16 November 2017) and annotated in Geneious v 10.1.2 (Biomatters Ltd., Auckland, New Zealand) by checking for open reading frames that matched the blastx search [8, 27, 28, 49]. Signal peptides were checked for and identified with SignalP v 4.1 [51, 52].

The final set of possible toxin and nontoxin transcripts for each of the nine individuals included anything that matched our local database or the Uniprot database that also had a complete protein coding sequences. From these, we extracted the coding sequence from each transcript in Geneious and then removed duplicates using the BBtools package Dedupe (Joint Genome Institute, Department of Energy, Walnut Creek, CA, USA) implemented in Geneious. We screened for and removed chimeric sequences in our toxins within each individual in two ways. First, we mapped the merged reads to the identified transcripts using Bowtie2 v 2.3.0 [53] implemented in Geneious and identified ones with irregular coverage including zero, multimodal, or uneven

coverage. Second, we checked for recombination within each of the toxin families through the ClustalW alignment algorithm [54] implemented in Geneious and identified transcripts that matched exactly with two or more other sequences. Transcripts identified in both were removed. The toxin and nontoxin transcripts remaining were clustered using Cd-hit at $\geq 98\%$ (-c .98) and a representative transcript was retained for each cluster. We then combined all toxin and nontoxins from the nine individuals into two sets based on their venom type. We removed duplicates, chimeras, and clustered as described above to get a master transcriptome for both Type A and Type B *C. scutulatus*. Finally, these were combined, had duplicates and chimeras removed, and clustered to get a master *C. scutulatus* transcriptome for final analyses.

Expression analysis

Relative expression of toxin and nontoxin genes was calculated by mapping merged reads to the final transcript set with Bowtie2 [53, 55] in RSEM v 1.3.0 [56] on the Stokes HPC on the UCF Advanced Research Computing Center using the default parameters. We used the transcripts per million reads (TPM) data for each individual as our abundance estimates [11]. We imported the dataset into RStudio v. 1.1.383 using R v. 3.4.2 (R Development Core Team 2006). We then created a 10th and 11th “individual” by averaging the TPM values for the five A animals (Average A) and four B animals (Average B). To eliminate 0 values from the dataset while preserving the compositional nature of expression data, we used the `cmultRepl` function of the R package `zCompositions` [57] then performed a centered log-ratio (clr) transformation to linearize the compositional dataset while preserving rank order of the transcripts [11, 58].

To determine if there was differential expression of toxin genes within and among venom types, we performed pairwise comparisons of each of the nine individuals as well as the Average A and Average B transcriptomes. We used the nontoxin expression values to generate a null distribution of expression divergence [44]. This was done by taking the absolute value of the difference in the transformed data for the two individuals being compared and finding the 99th

percentile value. Any toxin outside of this value was identified as an outlier to the null distribution. For each pairwise comparison, we used the Spearman's rank correlation coefficient (ρ), Pearson's correlation coefficient (R), and a coefficient of determination (R^2) to look at how similar the individuals being compared were. Finally, we tested for differential expression between our two venom types and between males and females using DESeq v 1.26.0 [59] and DESeq2 v. 1.14.1 [60] with a false discovery rate threshold of 0.1. We used the expected counts generated in RSEM and used the effective length to normalize the data. Toxins were assigned names based on toxin family and then ranked in order of highest average expression across all individuals for all families except PLA₂s. These were named to match the PLA₂s identified by Dowell *et al.* [7].

In RSEM, it is possible for a transcript that is not in the transcriptome to have a non-zero value. This is due to unmapped reads mapping to dissimilar sequences or, more commonly, highly similar regions among toxins within a specific family having reads dispersed among the different representatives. To determine which toxin transcripts were present in each individual and not an artifact of poor mapping, we aligned merged reads for each individual to the *C. scutulatus* master transcriptome using Bowtie2 implemented in Geneious. Any toxin that had more than 10% of the sequence with less than 5× coverage were considered absent in the transcriptome for that individual.

PLA₂ diversity

We tested the Dowell *et al.* [7] hypothesis of PLA₂ gene loss in *C. scutulatus* by assessing the diversity of PLA₂s in the nine individuals. We downloaded the four published genome fragments from GenBank (KX211993-KX21996) and extracted the mRNA sequence from all annotated PLA₂s. We clustered all sequences using Cd-hit at $\geq 90\%$ (-c .90) and a representative transcript was retained for each cluster. We then used Bowtie2 to align merged reads from each individual to the 10 resulting PLA₂s to determine which of the PLA₂s described in Dowell *et al.* [7] were present in *C. scutulatus*. We followed the same rule as above where there could not be

less than 5 coverage over 10% of the transcript.

Data availability

Raw data for the nine venom gland transcriptomes were submitted to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) accession SRP011323. BioSample accession numbers are provided in Table 3.1 and are under BioProject PRJNA88989. The consensus transcriptome was submitted to the NCBI Transcriptome Shotgun Assembly (TSA) database. This TSA project has been deposited at DDBJ/EMBL/GenBank under the accession GGIP00000000. The version described in this paper is the first version, GGIP01000000.

Results

Venom gland transcriptomes of C. scutulatus

We sequenced the venom gland transcriptomes of nine *C. scutulatus* individuals from the Sonoran Desert in the U.S.A. (Table 3.1, Figure 3.1). These individuals were chosen after determining their venom phenotype using RP-HPLC (Figure 3.2) to maximize the venom variation in *C. scutulatus*. Using 150 bp paired-end transcriptome sequencing on the Illumina MiSeq and HiSeq platforms, we generated over 131 million raw read pairs that yielded over 113 million merged reads that passed the quality filter, and where the 3' ends overlapped (Table 3.1). The nine individuals had an average of 12.6 ± 3.7 million merged reads. After assembly, annotation, duplicate and chimera removal, and clustering, our consensus *C. scutulatus* transcriptome consisted of 1889 putative nontoxins and 75 putative toxins from 17 toxin families (Table 3.2). To be considered present in the transcriptome, toxins had to have at least 5X read coverage over 90% of the total transcript sequence after mapping. Of the 75 toxins identified, three were exclusively found in animals with Type A venom and 17 were exclusively found in individuals with Type B venom (Table 3.3). Type A individuals had an average of 48.6 ± 6.1

toxins and Type B individuals had 65.6 ± 2.4 toxins. This difference was found to be significant based on a Mann–Whitney–Wilcoxon test ($W = 0$, $df = 7$, $p = 0.019$). Only 33 toxins were in all individuals and another 22 toxins were found in both Type A and Type B individuals but not in all nine individuals (Table 3.3).

Toxin diversity in C. scutulatus

We found high intraspecific diversity in the venom gland transcriptome both within and between Type A and Type B venoms in *C. scutulatus* (Figure 3.3). Much of the diversity was due to the presence and expression differences between Type A and Type B individuals in three toxin families: PLA₂s (including MTX), SVMPs, and CTLs. However, we document the first case of CTLs being highly expressed in the transcriptome of Type A animals. When comparing the average transcriptomes for Type A and Type B, the dichotomy in toxins is clear (Figure 3.4). However, there were toxins that demonstrated high variability in presence and expression both within and among venom types. Myotoxins had the most variation in expression of a toxin family not associated with venom type (Figure 3.3). In two Type A individuals (Christopher L. Parkinson field number CLP1972 and CLP1936) and one Type B individual (CLP1831), MYO-1 was the most highly expressed toxin in the transcriptome and was the second highest in CLP1835 (Figure 3.3). The most diverse toxin family was the CTL family, with 23 different putative toxins identified, followed by snake venom serine proteases (SVSPs) with 14, SVMPIIIs with 10, and myotoxins with 8 (Table 3.2).

Myotoxin a diversity

Myotoxins were the fourth most diverse toxin family with eight different toxins identified and presence and expression levels were highly variable among individuals (Figure 3.3). The four most highly expressed myotoxins were not associated with venom type nor were myotoxins associated with contact zones between the two types. CLP1835 had 10.1% of its toxin expression

Table 3.2: Transcripts per million reads (TPM) values for 75 toxins identified in the nine *C. scutulatus* individuals. TPM values were generated in RSEM [56] with Bowtie2 [53]. CRISP: cysteine-rich secretory protein; BPP: bradykinin potentiating peptide; CTL: C-type lectin; HYAL: hyaluronidase; KUN: Kunitz peptide; LAAO: L-amino-acid oxidase; MYO: myotoxin; NGF: nerve growth factor; NUC: 5' nucleotidase; PDE: phosphodiesterase; PLB: phospholipase B; SVMP: snake venom metalloproteinase; SVSP: snake venom serine protease; VEGF: vascular endothelial growth factor.

Toxin	Type A					Type B			
	CLP1930	CLP1936	CLP1959	CLP1961	CLP1972	CLP1831	CLP1835	CLP2136	CLP2142
BPP-1	14,0915.97	42,576.13	83,149.6	63,491.95	45,783.43	15,2393.13	12,6252.15	20,6001.59	14,0706.64
CRISP-1	4,835.29	4,825.49	15,386.6	7,108.63	5,075.55	4,047.71	4,625.53	2,879.87	3,386.38
CTL-1	0.00	3.65	14,629.20	25,897.00	0.00	17,494.71	26,507.89	18,545.42	44,489.80
CTL-2	0.00	1.98	12,439.65	20,441.04	0.00	18,825.37	22,039.29	18,679.67	43,789.90
CTL-3	1.12	0.00	0.31	10.11	0.00	12,805.18	13,046.93	13,965.96	21,899.71
CTL-4	0.00	0.45	0.00	3.20	0.00	12,605.98	15,574.61	11,584.92	16,750.81
CTL-5	0.41	2.52	7,965.07	10,951.63	0.00	4,218.11	11,546.02	10,057.03	10,667.89
CTL-6	0.40	1.26	11,399.06	12,437.39	0.00	3,461.60	3,578.61	9,394.12	10,374.16
CTL-7	0.38	1.44	7,423.12	8,508.27	0.00	4,653.94	10,065.04	9,327.24	10,662.63
CTL-8	22.39	8.78	7,498.38	9,084.97	232.76	2,273.74	4,494.79	6,799.15	11,587.90
CTL-9	3,859.43	4,539.36	9,932.82	10,002.06	3,012.54	1,561.25	2.48	1.52	2,634.56
CTL-10	10.63	4.93	7,658.06	9,102.48	84.69	2,499.90	3,293.65	6,962.90	3,823.74
CTL-11	0.00	0.47	2,744.48	3,752.74	0.00	2,814.96	6,914.87	4,523.03	5,131.21
CTL-12	3,604.05	3,074.72	6,797.37	6,941.60	2,123.74	1,975.31	14.78	24.72	40.48
CTL-13	0.00	0.41	2,379.11	2,630.23	0.00	947.76	3,008.94	1,948.93	2,122.28
CTL-14	0.88	171.20	883.68	555.05	147.92	637.66	1,994.47	2,157.31	2,572.39
CTL-15	995.02	292.59	809.42	726.11	910.53	1,168.87	748.84	1,791.77	1,309.00
CTL-16	1,246.38	418.83	2,132.33	2,473.36	472.09	433.29	0.55	234.84	342.16
CTL-17	0.00	0.00	0.41	8.06	0.00	4.34	10.41	51.78	6,202.50
CTL-18	325.45	305.49	4,026.88	247.89	201.87	279.93	274.45	36.75	527.21
CTL-19	28.46	1,045.78	2,968.50	60.46	395.23	49.87	213.36	285.33	172.05
CTL-20	0.00	0.00	0.00	0.87	0.00	0.00	0.00	1.60	1,233.72
CTL-21	0.40	0.00	0.53	11.61	0.00	59.06	27.62	511.53	361.34
CTL-22	0.00	0.00	0.38	29.57	0.00	51.73	54.86	280.15	438.68
CTL-23	3.52	9.49	15.25	7.90	45.56	5.09	13.42	1.34	9.87
HYAL-1	557.54	304.13	1,309.25	379.62	221.07	468.14	408.71	436.41	222.96
KUN-1	178.35	48.25	168.92	123.70	71.34	107.63	147.09	195.49	122.11
KUN-2	14.96	4.32	13.36	14.48	5.19	20.25	20.77	20.98	12.38
LAAO-1	5,188.99	2,817.21	20,588.57	9,120.95	1,626.36	5,792.61	10,578.38	6,868.00	5,503.11
MYO-1	1,220.73	549,957.26	1,072.89	27.84	457,306.12	209,616.39	44.93	44.14	29,489.77
MYO-2	6,119.03	49,381.65	2,467.88	3,325.84	95,421.49	27,536.10	2,903.51	6,987.97	3,925.07
MYO-3	0.00	0.00	0.00	34,019.55	0.00	0.00	87,549.28	30.87	33,694.44
MYO-4	5,843.47	3,702.71	12,274.75	19,693.00	497.74	1,611.10	10,899.95	2,894.17	4,002.29
MYO-5	0.00	82.14	0.00	0.00	0.00	44.63	0.00	43.19	35,485.73
MYO-6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2,706.33
MYO-7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1,805.60
MYO-8	3.83	960.47	8.14	0.00	0.00	0.00	0.00	0.00	0.00
NGF-1	2,021.69	761.53	5,039.75	2,527.51	2,023.71	1,592.10	2,240.15	2,459.36	2,108.54
NUC-1	779.38	366.04	1,526.31	945.96	368.05	610.32	960.93	1,068.46	502.67
PDE-1	445.18	229.67	691.74	312.13	236.13	209.66	277.73	208.12	249.61
Pla2gA1	25,306.99	39,764.45	61.41	46.36	2.19	4.48	50,036.23	0.00	0.61
Pla2gB1	3.44	7.16	29,405.58	14.34	1.43	37,548.76	46,634.43	51,949.04	31,992.11
Pla2gK	0.00	0.30	1,793.61	3.04	0.00	1,356.66	2,971.25	1,123.48	807.59
Pla2gA2-MTXA	274,390.23	80,733.96	102,588.14	199,107.20	102,677.33	4.93	65.65	0.00	0.00
Pla2gB2-MTXB	133,822.75	43,410.96	48,028.34	138,470.84	42,571.19	3.73	30.28	0.00	0.00
PLA ₂ -6	21,306.42	22.86	106,639.29	90,911.79	51,893.58	60,760.91	56,428.98	129,974.00	93,064.76
PLB-1	837.02	189.30	1,170.66	1,059.23	597.70	509.37	745.22	346.61	182.83
SVMP1I-1	0.00	0.84	0.60	122.86	3.30	23,801.48	32,375.22	24,762.53	24,369.30
SVMP1I-1	0.00	1.20	1.83	262.28	9.58	58,107.40	69,839.79	69,573.48	39,410.80
SVMP1I-2	8.60	3,074.27	61.30	2,524.10	479.43	3,581.65	14,465.14	2,364.36	4,089.81
SVMP1I-3	0.01	0.79	1.32	206.34	6.17	6,920.71	8,968.12	7,129.26	6,456.84
SVMP1I-4	1,497.61	1,018.71	4,893.89	2,711.96	276.26	3,637.08	4,729.55	5,060.59	2,488.36
SVMP1I-5	6.65	4.39	4.51	2.02	0.00	1,293.10	2,689.88	3,646.35	1,942.60
SVMP1I-6	19.52	12.50	6.89	9.20	2.20	274.19	3,250.08	2,739.21	663.84
SVMP1I-7	0.02	0.17	0.57	16.36	0.53	643.82	2,557.44	2,534.21	994.18
SVMP1I-8	0.00	0.00	0.00	0.07	0.00	44.33	50.02	4,977.26	25.70
SVMP1I-9	0.00	0.00	0.00	0.07	0.00	279.91	2,000.74	1,017.78	1,092.48
SVMP1I-10	0.00	0.01	0.00	0.19	0.00	181.26	751.33	220.24	404.67
SVSP-1	25,158.51	15,910.07	30,344.55	17,038.01	10,432.56	21,732.91	32,652.46	20,711.71	26,686.11
SVSP-2	42,979.80	20,190.56	35,325.26	21,389.97	14,037.74	16,208.71	10,861.87	6,624.20	6,521.70
SVSP-3	24,402.09	7,206.88	34,812.15	9,121.20	3,774.14	19,942.19	20,070.68	20,385.25	15,388.82
SVSP-4	29,001.28	5,094.40	18,056.16	9,305.84	1,916.43	8,654.88	9,431.62	7,042.42	8,412.05
SVSP-5	37,090.47	5,414.68	21,446.17	8,874.54	1,346.38	1,863.68	5,619.86	2,373.32	3,563.03
SVSP-6	11,830.88	7,170.10	15,765.87	9,603.42	4,485.31	8,500.16	12,274.97	6,642.73	10,481.74
SVSP-7	7,289.06	2,382.92	5,895.10	3,081.90	1,513.33	4,024.78	2,685.70	1,659.02	1,379.27
SVSP-8	4,813.04	2,087.12	5,900.27	2,021.75	407.33	2,169.99	2,985.67	1,994.54	1,945.35
SVSP-9	4,048.20	724.15	1,391.55	1,288.30	1,231.13	3,097.60	3,252.51	2,154.34	3,169.89
SVSP-10	4,038.45	1,189.72	4,293.59	1,766.03	84.61	1,208.22	2,752.40	2,430.80	2,459.42
SVSP-11	5,777.06	1,464.77	2,503.34	1,599.56	962.33	2,311.84	3,050.76	478.40	1,664.01
SVSP-12	3,674.93	1,456.06	3,014.74	2,310.02	565.32	2,209.80	1,897.49	1,761.75	1,281.51
SVSP-13	3,238.56	798.05	3,159.21	1,118.98	287.65	1,467.22	709.22	982.38	1,198.93
SVSP-14	1,558.37	620.40	1,082.12	667.47	38.63	762.05	289.19	599.21	693.61
VEGF-1	4,255.45	10,557.63	14,271.01	11,732.75	7,493.65	20,763.31	9,620.66	12,596.93	10,796.52
VEGF-2	134.91	18.84	82.27	95.20	3.90	96.12	72.87	313.81	91.89
Vespryn-1	4,976.55	718.69	3,827.06	1,024.92	1,089.92	2,106.99	2,068.47	922.31	582.28

Table 3.3: Presence and absence data for toxin transcripts that were not found in all individuals. Toxins highlighted in dark blue were found in all Type B individuals but never in Type A individuals. Toxins highlighted in dark red were found in all Type A individuals but never in Type B individuals. Toxins highlighted in light blue or light red were only found in individuals of that venom type but were not found in all individuals. The last row is the number (out of 75) of toxins present in total, which includes 33 toxins present in all individuals not listed in this table. To be present in the transcriptome, toxins had to have at least 5× coverage over 90% of the transcript. MTXA: acidic subunit of Mojave toxin; MTXB: basic subunit of Mojave toxin.

Toxin	Type A					Type B			
	CLP1930	CLP1936	CLP1959	CLP1961	CLP1972	CLP1831	CLP1835	CLP2136	CLP2142
CTL-1	-	-	+	+	-	+	+	+	+
CTL-2	-	-	+	+	-	+	+	+	+
CTL-3	-	-	-	-	-	+	+	+	+
CTL-4	-	-	-	-	-	+	+	+	+
CTL-5	-	-	+	+	-	+	+	+	+
CTL-6	-	-	+	+	-	+	+	+	+
CTL-7	-	-	+	+	-	+	+	+	+
CTL-8	-	-	+	+	+	+	+	+	+
CTL-9	+	+	+	+	+	+	-	-	+
CTL-10	-	-	+	+	+	+	+	+	+
CTL-11	-	-	+	+	-	+	+	+	+
CTL-12	+	+	+	+	+	+	+	-	-
CTL-13	-	-	+	+	-	+	+	+	+
CTL-14	-	+	+	+	+	+	+	+	+
CTL-16	+	+	+	+	+	+	-	+	+
CTL-17	-	-	-	-	-	-	-	+	+
CTL-19	+	+	+	+	+	-	+	-	+
CTL-20	-	-	-	-	-	-	-	-	+
CTL-21	-	-	-	-	-	+	+	+	+
CTL-22	-	-	-	-	-	+	+	+	+
CTL-23	+	+	+	+	+	+	+	-	+
MYO-3	-	-	-	+	-	-	+	+	+
MYO-5	-	+	-	-	-	+	-	+	+
MYO-6	-	-	-	-	-	-	-	-	+
MYO-7	-	-	-	-	-	-	-	-	+
MYO-8	-	+	+	-	-	-	-	-	-
Pla2gA1	+	+	+	-	-	-	+	-	-
Pla2gB1	-	-	+	-	-	+	+	+	+
Pla2gK	-	-	+	-	-	+	+	+	+
Pla2gA2-MTXA	+	+	+	+	+	-	-	-	-
Pla2gB2-MTXB	+	+	+	+	+	-	-	-	-
PLA ₂ -6	+	-	+	+	+	+	+	+	+
SVMPII-1	-	-	-	-	-	+	+	+	+
SVMPIII-1	-	-	-	-	-	+	+	+	+
SVMPIII-2	-	+	+	+	+	+	+	+	+
SVMPIII-3	-	-	-	-	-	+	+	+	+
SVMPIII-5	-	-	-	-	-	+	+	+	+
SVMPIII-6	-	-	-	-	-	+	+	+	+
SVMPIII-7	-	-	-	-	-	+	+	+	+
SVMPIII-8	-	-	-	-	-	-	-	+	-
SVMPIII-9	-	-	-	-	-	+	+	+	+
SVMPIII-10	-	-	-	-	-	+	+	+	+
Toxins Present	42	45	56	53	45	64	66	64	69

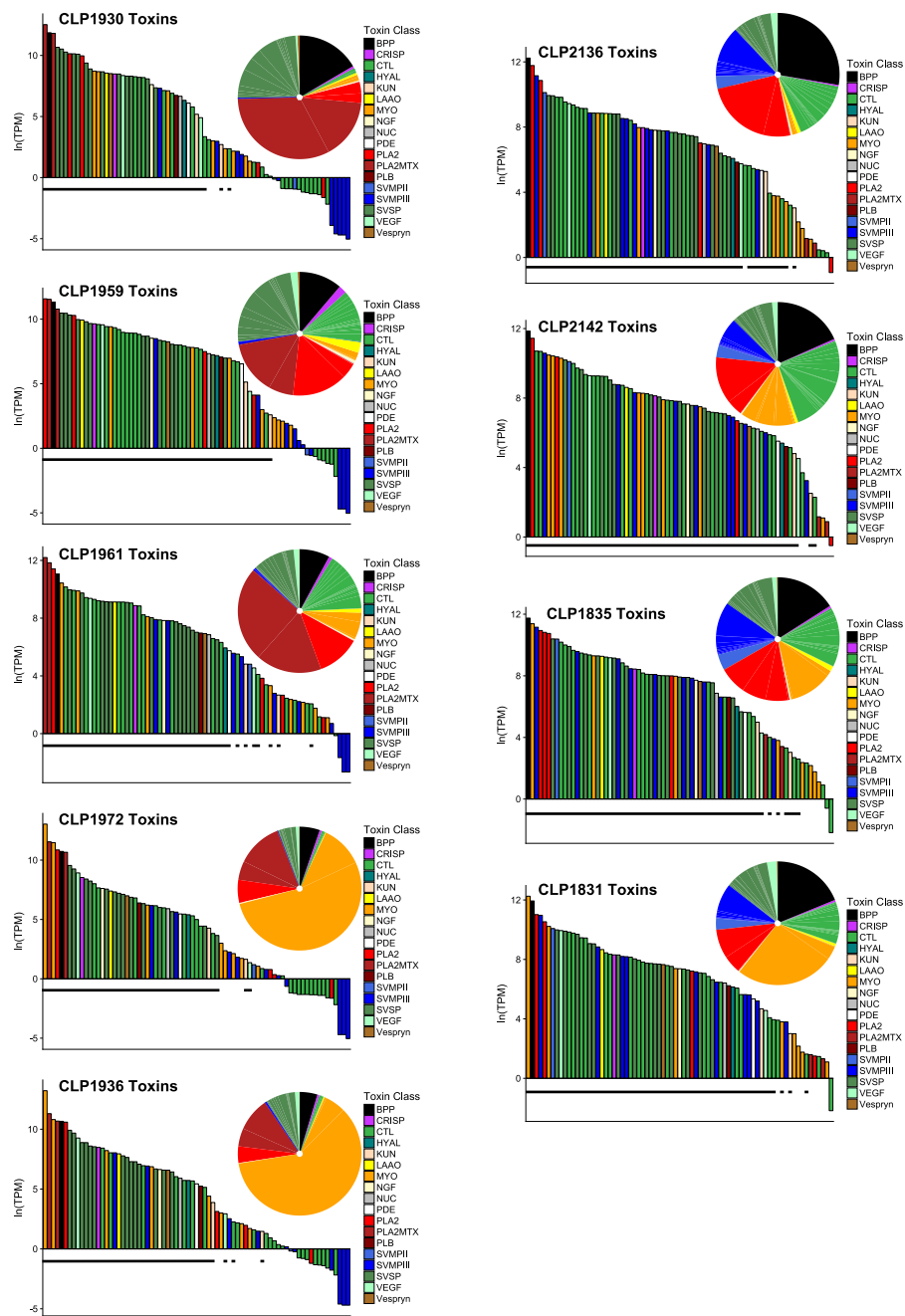


Figure 3.3: Representation of the toxins in the venom gland transcriptome for nine *C. scutulatus* specimens. Type A are on the left and Type B are on the right, and are ranked by the increasing amount of myotoxins. The bar graphs represent each of the 75 toxins identified. Any toxin that does not have a black bar under it did not meet the criteria for presence in the transcriptome. The pie charts represent the proportion of each toxin family in the venom gland transcriptome.

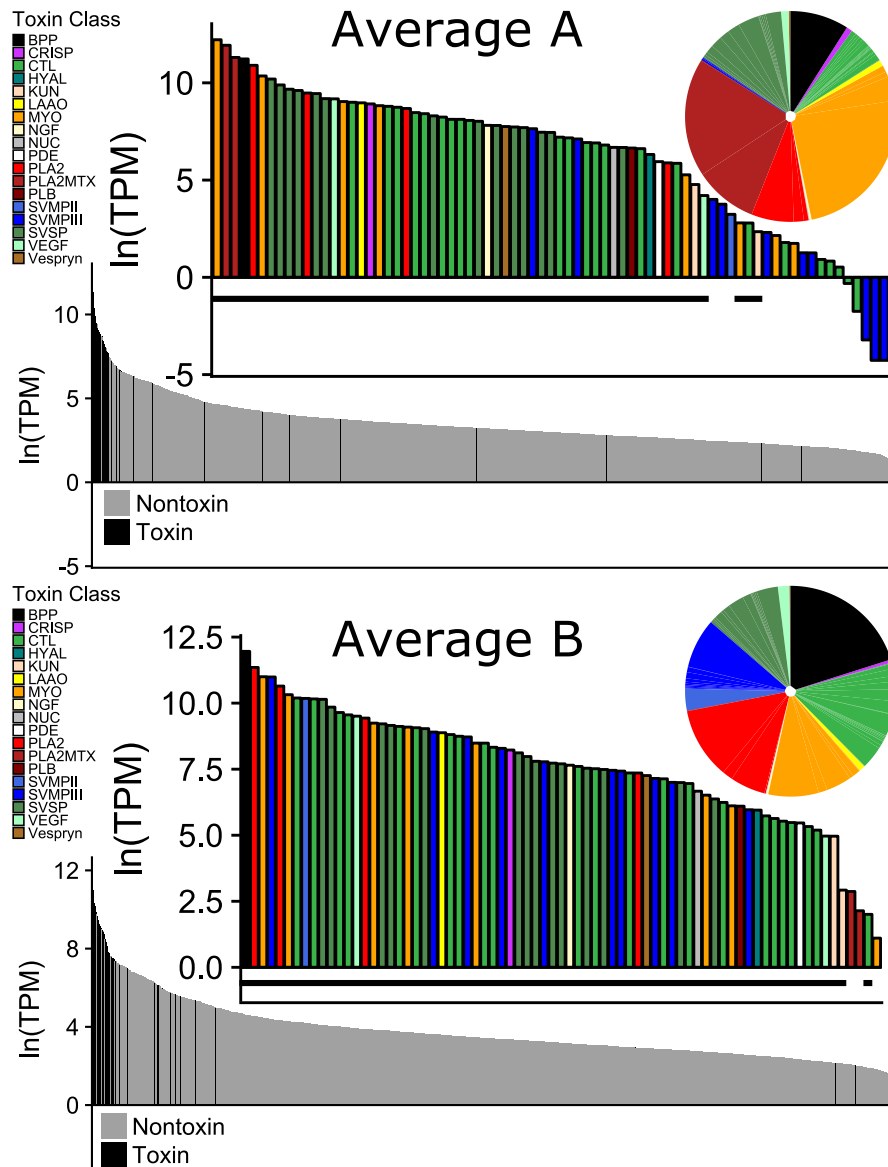


Figure 3.4: Representation of the average Type A and Type B venom gland transcriptome of *C. scutulatus* from the Sonoran Desert. These were generated by averaging the TPM values of each individual within a venom type. For both Type A and Type B, the majority of the highly expressed transcripts are toxins. The bar graphs with colors represent each of the 75 toxins identified. Any toxin that does not have the black bar under it did not meet the criteria for being present in the transcriptome of any individual of that venom type. The pie charts represent the proportion of each toxin family in the venom gland transcriptome. Type A individuals have very few SVMPIs and the Type B individuals are lacking neurotoxic phospholipase A₂ (PLA₂), Mojave toxin.

comprised of myotoxins and was closest to the center of the Type B distribution. Additionally, CLP2136 and CLP2142 were found very close to each other but CLP2136 had little if any myotoxin expression. The percentage of myotoxins in the toxin transcriptome ranged from 1.0% to 60.4% and was not associated with size or sex.

Expression differences in Type A and Type B C. scutulatus

Toxin family expression levels were highly variable among families and many differed between venom types. (Figure 3.5). The PLA₂s that comprise MTX were highly expressed in Type A, with very low expression in Type B as expected. The opposite was true for the SVMPII and the SVMPIII. These correspond to the major differences between Type A and Type B venom (Figure 3.4 and 3.5). For most of the toxin families, there was variability within venom types as illustrated by the error bars in Figure 3.5. The non-MTX PLA₂s, CTLs, MYOs, and vascular endothelial growth factors (VEGFs) all had almost completely overlapping standard deviations. Other toxin families such as bradykinin potentiating peptide (BPP), cysteine-rich secretory protein (CRISP), L-amino-acid oxidase (LAAO), and nerve growth factor (NGF) did differ between venom types slightly but had much tighter variation in expression levels within each venom type.

Using the pairwise comparisons of all Type A individuals to all Type B individuals (20 comparisons), we found many toxins that differed between Type A and Type B individuals (Table 3.4). Many of these corresponded with the presence/absence variation between PLA₂s and SVMPs but others were in toxins expressed in all individuals. Our DESeq1 and DESeq2 analyses were more conservative in identifying toxins that differed between Type A and Type B venoms, primarily identifying toxins that had dramatic presence/absence differences in the transcriptome. DESeq1 identified 20 toxins and DESeq2 identified 26 toxins that were significantly different but 13 and 19 were toxins that were exclusively found in one venom type by each program, respectively. Only Pla2gB1, SVSP-2, and CTL-12 were identified among the remaining toxins by both analyses. Of the 33 toxins present in all nine individuals, only CRISP-1, phospholipase

B (PLB-1), and SVSP-5 were identified as differentially expressed in DESeq2. By comparison, the only toxin identified as having significantly different expression between males and females in DESeq1 and DESeq2 was PLA2gA1, which is likely an artifact of sampling rather than an indication of sexual dimorphism.

For the pairwise comparison of the Average A and Average B transcriptomes, the relative expression of each of the nontoxins was highly correlated (Figure 3.6). However, toxin expression between venom types was poorly correlated. This was driven by the two subunits of MTX, SVMPs, CTLs, and myotoxins. The remaining toxins within the toxin families were highly correlated between the venom types. This includes BPP, CRISP, hyaluronidase (HYAL), Kunitz peptide (KUN), LAAO, NGF, 5' nucleotidase (NUC), phosphodiesterase (PDE), PLB, SVSPs, VEGF, and Vespryn.

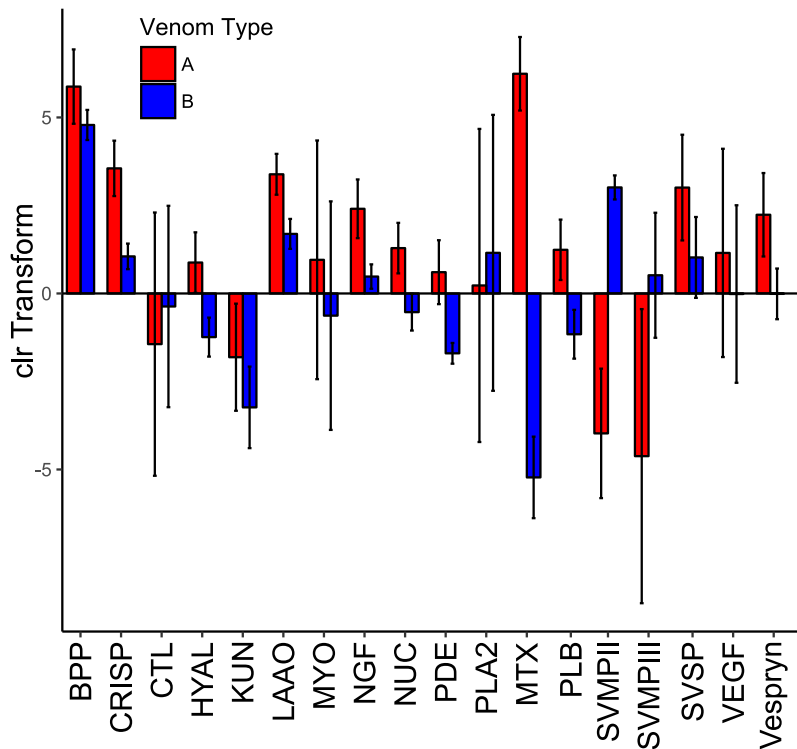


Figure 3.5: Average expression of toxin families with respect to Type A and Type B *C. scutulatus* after centered log-ratio (clr) transform. Error bars are the standard deviation around the mean of the clr-transformed data.

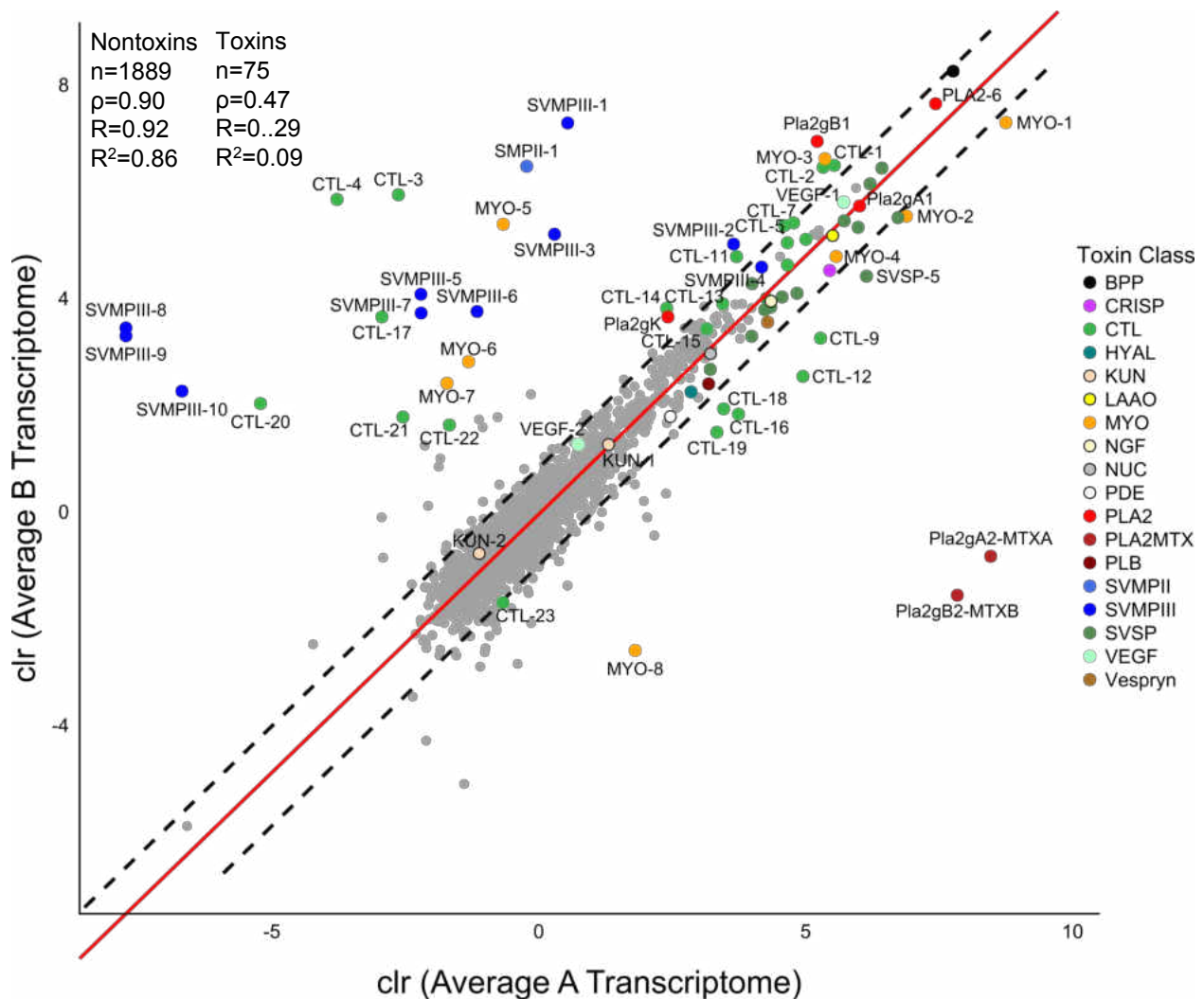


Figure 3.6: Pairwise comparison of the average Type A and Type B venom gland transcriptomes using the centered log-ratio (clr)-transformed TPM data. SVSPs and toxin families that only have one toxin are not labeled. SVSPs did not differ between the two transcriptomes. The red line is the line of best fit through the non toxins and the dashed black lines are the 99% confidence around that line. Any transcript outside the dashed black lines was identified as an outlier. Anything above the upper line is overexpressed in Type B and anything below the lower line is overexpressed in Type A.

PLA₂ diversity

Dowell *et al.* [7] annotated nine PLA₂s with mRNA coding sequence in the four sequenced genome fragments. Of these, only five were expressed in the venom gland transcriptome and the remaining four were exclusively found in the genomes of the species they sequenced (see Figure

Table 3.4: Differential expression analyses for toxins between the nine Type A and Type B *C. scutulatus* as well as the average Type A (AveA) and Type B (AveB) transcriptomes. The UpB and UpA count data were generated by identifying outlier transcripts in the pairwise comparisons of the Type A and Type B individuals (maximum of 20 comparisons). The last four columns are the data from DESeq and DESeq2 identifying differential expression between the two venom types. Toxins highlighted in dark red were found in all Type A individuals but never in Type B individuals. Toxins highlighted in light blue or light red were exclusively found in individuals of that venom type but were not found in all individuals. Toxins highlighted in green were found in all nine individuals. Toxins with NA for P_{adj} only had one individual in one of the treatments with the toxin so it was not possible to calculate in DESeq2.

Toxin	UpB	UpA	Δ_{B-A}	B to A	AveB to AveA	$\text{Log}_2\Delta$ DESeq1	P_{adj} DESeq1	$\text{Log}_2\Delta$ DESeq2	P_{adj} DESeq2
CTL-3	20	0	20	Up	Up	12.76	6.55E-42	12.73	1.11E-20
CTL-4	20	0	20	Up	Up	14.14	5.91E-96	14.13	1.03E-111
SVMPH-1	20	0	20	Up	Up	9.94	1.96E-71	9.97	N/A
SVMPH-1	20	0	20	Up	Up	9.97	1.01E-71	10.00	N/A
SVMPH-10	20	0	20	Up	Up	13.17	6.48E-16	13.28	4.27E-40
SVMPH-3	20	0	20	Up	Up	7.35	2.89E-48	7.38	N/A
SVMPH-5	20	0	20	Up	Up	8.96	8.76E-42	8.96	4.42E-16
SVMPH-6	20	0	20	Up	Up	6.86	5.44E-04	6.86	5.30E-22
SVMPH-7	20	0	20	Up	Up	8.75	6.91E-16	8.79	1.00E-08
SVMPH-9	20	0	20	Up	Up	16.23	1.76E-20	15.89	2.41E-48
CTL-21	19	0	19	Up	Up	6.65	2.60E-02	6.58	3.92E-03
Pla2gB1	19	0	19	Up	Up	3.06	6.07E-08	3.04	N/A
CTL-22	18	0	18	Up	Up	-	-	-	-
CTL-14	16	0	16	Up	Up	2.22	9.75E-02	-	-
Pla2gK	16	0	16	Up	Up	-	-	2.32	N/A
CTL-1	15	0	15	Up	Up	-	-	-	-
CTL-2	15	0	15	Up	Up	-	-	-	-
CTL-11	14	0	14	Up	Up	2.01	7.05E-02	-	-
SVMPH-2	14	0	14	Up	Up	-	-	-	-
CTL-7	12	0	12	Up	Up	-	-	-	-
CTL-17	10	0	10	Up	Up	-	-	10.07	N/A
MYO-5	9	1	8	Up	Up	-	-	8.04	N/A
MYO-3	9	2	7	Up	Up	-	-	-	-
CTL-20	5	0	5	Up	Up	-	-	11.01	N/A
MYO-6	5	0	5	Up	Up	-	-	23.74	N/A
MYO-7	5	0	5	Up	Up	-	-	23.49	N/A
SVMPH-8	5	0	5	Up	Up	-	-	15.96	N/A
MYO-2	4	3	1	Up	Down	-	-	-	-
CRISP-1	0	0	0	No Difference	-	-	-	-1.32	3.16E-03
CTL-19	6	8	-2	Down	Down	-	-	-	-
MYO-1	7	9	-2	Down	Down	-	-	-	-
PLB-1	0	2	-2	Down	-	-	-	-1.07	7.76E-02
SVSP-2	1	4	-3	Down	Down	-1.77	9.50E-03	-1.77	5.11E-04
MYO-8	0	4	-4	Down	Down	-	-	-23.22	N/A
SVSP-5	1	5	-4	Down	Down	-	-	-2.32	5.29E-03
CTL-18	0	5	-5	Down	Down	-	-	-	-
CTL-23	0	7	-7	Down	Down	-	-	-	-
CTL-16	0	9	-9	Down	Down	-2.49	5.78E-04	-	-
CTL-9	0	11	-11	Down	Down	-2.73	2.56E-06	-	-
CTL-12	0	15	-15	Down	Down	-3.36	1.14E-09	-3.35	7.88E-02
Pla2gA2-MTXA	0	20	-20	Down	Down	-13.55	8.76E-42	-13.51	1.18E-35
Pla2gB2-MTXB	0	20	-20	Down	Down	-13.65	5.75E-38	-13.61	2.10E-34

1 in [7]). We found evidence for all five of the expressed PLA₂s in our transcriptomes as well as a sixth PLA₂ (PLA₂-6) not recovered in Dowell *et al.* [7] (Table 3.5). As expected, the two ancestral mammal homolog PLA₂s, Pla2-e and Pla2-f, were not expressed in the venom gland. In agreement with Dowell *et al.* [7] we did not find evidence for Pla2-gC1 identified in all three genomes nor Pla2-d identified in the *C. adamanteus* genome being expressed. No PLA₂ was found in all individuals and one individual (CLP1959) had evidence for all six PLA₂s being present although Pla2gA1 had very low expression compared to other individuals and PLA₂s (Table 3.5). In CLP1835, the sequence for the Pla2gA1 was not the same as the other *C. scutulatus* individuals that expressed it. When blasted against the non-redundant nucleotide database in GenBank, it matched the sequence of *Crotalus viridis* (Accession AF403134) which is in the sister species complex to *C. scutulatus*. The sequence had eight nonsynonymous nucleotide changes compared to the other three *C. scutulatus* individuals that expressed Pla2gA1 (Table 3.5).

Table 3.5: Presence or absence data for PLA₂s identified by Dowell *et al.* [7] (the first five PLA₂s) and the sixth PLA₂ identified in this study. The first four individuals were the specimen used in Dowell *et al.* [7] and only presence/absence is indicated. The last nine individuals are *C. scutulatus* specimens sequenced in this study with venom type indicated and transcripts per million reads (TPM) values given when that PLA₂ is present. The one TPM value denoted by an * had eight nonsynonymous nucleotide changes in the sequence compared to the other three *C. scutulatus* specimens and matched that of *Crotalus viridis* (Genbank accession AF403134).

Specimen	Pla2gA1	Pla2gB1	Pla2gK	Pla2gA2-MTXA	Pla2gB2-MTXB	PLA ₂ -6
<i>C. atrox</i>	+	+	+	-	-	-
<i>C. atrox</i>	+	+	+	-	-	-
<i>C. adamanteus</i>	+	+	-	-	-	-
<i>C. scutulatus</i>	+	-	-	+	+	-
CLP1930A	55212.77	-	-	603261.66	294142.52	47256.79
CLP1936A	242536.57	-	-	492462.39	264770.94	-
CLP1959A	214.27	101885.11	6162.67	355273.29	166449.59	369588.5
CLP1961A	-	-	-	464139.64	323205.86	212285.6
CLP1972A	-	-	-	520518.57	216000.5	263309.5
CLP1831B	-	370853.15	13305.52	-	-	600310.86
CLP1835B	320034.14*	297244.14	18761.21	-	-	58732.73
CLP2136B	-	283162.92	6014.01	-	-	709003.82
CLP2142B	-	253634.76	6285.41	-	-	738317.63

Discussion

The major differences in the venom gland transcriptome within and between Type A and Type B *C. scutulatus* from the Sonoran Desert were driven by the presence or absence of PLA₂s including Mojave toxin (MTX), snake venom metalloproteinases (SVMPs), C-type lectins (CTLs), and myotoxins (MYO) (Figure 3.3). Myotoxin expression was not associated with the contact zones between the two venom types as hypothesized by Massey *et al.* [39]. We did not find evidence for distinct sets of PLA₂s for Type A and Type B venom types as hypothesized by Dowell *et al.* [7] using three species (two Type B/one Type A). Additionally, one individual expressed all six PLA₂s (Table 3.5). We also found evidence for a PLA₂ allele in one *C. scutulatus* individual that was more similar to a congener, *C. viridis*, than the other eight *C. scutulatus* specimens. The allele could have originated from interspecific hybridization with a member of the *C. viridis* species complex. Other works have suggested introgression through hybridization as a mechanism for propagating toxin genes among species. For example, it is hypothesized that the *C. horridus* MTX homolog, canebrake toxin, was introduced by intergeneric hybridization with *Sistrurus catenatus* [28]. These data, taken together, support non-allelic homologous recombination (NAHR) as an important mechanism driving PLA₂ diversity in *C. scutulatus* venom and rattlesnakes more broadly [7].

Expression differences between Type A and Type B venom were due to the presence/absence of specific toxins (Table 3.4). Type A venoms were simpler in that they contained fewer toxins overall, driven by the lack of SVMP and CTL expression. The presence/absence expression difference between MTX and SVMPs is the characteristic difference between the Type A (neurotoxic) and Type B (hemorrhagic) venom dichotomy seen within rattlesnakes. As expected, both subunits of the neurotoxic PLA₂ (Mojave toxin) were exclusively found in Type A individuals, and Type B individuals had a high diversity of SVMPs [61, 62]. All but two SVMPs (SVMPIII-2 and SVMPIII-4) were absent from all Type A individuals which is similar to what was found in

the Type A *C. horridus* [28]. SVMPIII-4 was found in all nine individuals and SVMPIII-2 was missing in CLP1930 (Table 3.3).

C-type lectins (CTLs) had the highest number of unique toxins within a family, with 23 different toxins (Table 3.2). Only two were found in all nine individuals, six were exclusively found in Type B individuals, and the remaining 15 were found in different combinations regardless of venom type (Table 3.3). CTLs are diverse in Type B species and affect coagulation factors, increasing hemorrhaging [27, 63]. Interestingly, two of the Type A individuals, CLP1959 and CLP1961, had high expression of CTLs at 10.2% and 12.4%, respectively where the other three individuals had almost no expression ($\leq 1.0\%$, Figure 3.3). This is the first time that CTLs have been documented to be highly expressed in a Type A individual. In the Type A *C. horridus*, CTLs only accounted for 0.2% of the toxin reads [27].

Massey *et al.* [39] documented high intraspecific variability of myotoxin in the proteome of *C. scutulatus*, independent of venom type. Because of this, they suggested *C. scutulatus* venoms be further divided into six venom types: Type A, Type A + M, Type B, Type B + M, Type A + B, and Type A + B + M. The addition of these myotoxins in the venom decreased the lethal dose 50 (LD_{50}) values and work to disrupt sodium channels in muscle cells causing muscle paralysis [39, 64]. Our transcriptome data support further differentiating venom types to account for the diversity in myotoxin expression levels. We found differences in myotoxin expression among the nine individuals but they were not associated with the contact zone between the two venom types as predicted. Of the eight myotoxins we identified, one was found exclusively in Type A individuals and two were exclusively found in type B individuals (Table 3.3). However, these are likely a function of sampling rather than fixed toxins in those venom types given the variability in myotoxins overall (Figure 3.3). When present, the four most highly expressed myotoxins (MYO-1–4) were expressed at similar levels between the venom types (Figure 3.6).

Based on the hypothesis of Dowell *et al.* [7], we expected to find distinct sets of PLA_2 s in Type A and Type B *C. scutulatus*, as was found in *C. horridus* [28] and the three species

with genomic fragments sequenced [7]. However, this was not the case and we found almost all possible combinations of the six putative PLA₂ toxins among the nine individuals we sequenced (Table 3.5). Other than the acidic and basic subunits of MTX which were exclusively found in Type A individuals, the remaining four PLA₂s were not specific to a venom type (Table 3.5). Further genomic analysis is needed to determine why the two subunits of MTX appear to be consistently found together whereas the other four can be inherited in different combinations. Additionally, it is possible for homologs of MTX, like Crotoxin, to be present in the genome and not be expressed as *Crotalus simus* undergoes an ontogenetic shift where Crotoxin is expressed in juveniles, but not expressed in adults [65].

One individual, CLP1959, expressed all six PLA₂s including the acidic subunit of Mojave toxin (MTXA) and the basic unit of Mojave toxin (MTXB) and the PLA₂s associated with Type B individuals, although one, Pla2gA1, was expressed at a significantly lower level than the other individuals that expressed it. That same toxin, Pla2gA1, in CLP1835 was 100% identical to a known allele found in *C. viridis*. All eight nucleotide changes from the other *C. scutulatus* that expressed it were non-synonymous, thus changing the amino acid sequence. The *C. viridis* species complex is sister to *C. scutulatus* so this could be a shared ancestral allele or introduced through hybridization. Hybridization between *C. scutulatus* and *C. viridis* has been documented but they are not syntopic in this region [19]. However, *Crotalus cerberus*, a member of the *C. viridis* species complex, is co-distributed with *C. scutulatus* in this region and could be the origin of this allele if it is shared within the complex.

Using two Type B species (*Crotalus atrox* and *C. adamanteus*) and one Type A individual (*C. scutulatus* from the Chihuahuan phylogeographic lineage), Dowell *et al.* [7] predicted that the region in the genome that contains the PLA₂ genes is prone to NAHR and hypothesized that there would likely be diversity within a species as we document in *C. scutulatus*. If NAHR is the mechanism for gene movement, then it might explain how the different PLA₂ genes, particularly MTX, can be reintroduced into populations that lose it through hybridization as hypothesized by

Rokyta *et al.* [28] for *C. horridus*. *Crotalus scutulatus* is known to hybridize with other species of rattlesnakes and, when this occurs, MTX can be found in the resulting hybrids [19]. Dowell *et al.* [7] agree with Lynch [66] in that PLA₂s could go through a selective sieve after an ecological shift such as changing diet which causes the loss of the less adaptive PLA₂s. This mechanism presumes SVMPs are down regulated or lost when MTX is present, but it could be the opposite. Alternatively, the two phenotypes may represent two fitness optima that can be maintained spatially or certain venom components could be selected for or against in specific environments based on different prey availability. Regardless of the mechanism, the interplay between gene flow and selection in the Sonoran Desert is allowing individuals with the two venom types to persist spatially without an obvious ecological difference between venom phenotypes.

The 13 remaining toxin families did not show the same pattern of presence/absence and all 27 toxins were present in each transcriptome at varying levels (Table 3.3 and Figure 3.4). Though snake venom serine proteinases (SVSPs) were the second most diverse toxin family, there were no differences in expression of the toxins between venom types. This is similar to what was found in *C. horridus* (see Figure 4 in Rokyta *et al.* [28]). There was one BPP and it was among the most highly expressed proteins in the venom of all individuals and had low variability among individuals (Figure 3.5). Other than KUN and VEGF, the remaining families (CRISP, LAAO, NGF, NUC, PDE, PLB, and Vespryn) also had low variability among individuals (Figure 3.5). We did not find any expression differences associated with the size or sex of *C. scutulatus*. Nontoxin expression was strongly correlated between the two venom types (Figure 3.4).

Mojave Rattlesnakes are representative of the diversity documented in other rattlesnakes and our work illustrates the utility of sequencing multiple individuals of a species to represent the phenotypic diversity found. Both CTL and MYO diversity would have been underestimated within venom types if multiple individuals were not included (Figure 3.3). Overall, the transcriptomic differences in *C. scutulatus* matched the patterns documented between *C. adamanteus* and *C. horridus* [27] and that between Type A and Type B *C. horridus*. This included some individuals

that had high levels of myotoxins as in *C. adamanteus* [49]. However, *C. scutulatus* was different in that myotoxin expression was much higher and there were no distinct sets of PLA₂s between venom types as exhibited by the polymorphic *C. horridus*. *Crotalus scutulatus* will continue to be an exemplary model system to understand the evolution of venom particularly when the remaining phylogeographic lineages are included as well as the Type A + B individuals. Additionally, given the diversity in presence/absence of toxins within the major families, *C. scutulatus* would be useful to test NAHR in other toxin families.

Conclusions

Phenotypic diversity in *Crotalus scutulatus* is representative of venom diversity in rattlesnakes with both sequence-based and expression evolution occurring. By sampling and sequencing the transcriptome of multiple individuals of each venom phenotype, we were able to highlight the diversity throughout the Sonoran lineage, including the characteristic Type A and Type B venom phenotypes. For toxins that were not exclusively associated with venom type, different combinations occurred, particularly in PLA₂s. High gene flow in this region of *C. scutulatus*' distribution and interspecific hybridization may facilitate different combinations of these toxins, especially if NAHR occurs broadly among toxin families. Further genomic resources coupled with transcriptomics and proteomics of the remaining two lineages as well as the Type A + B phenotype may make it possible to understand the evolution of toxin gains and losses in PLA₂s, CTLs, myotoxins, and SVMPs as well as differential expression and sequence evolutions of specific toxins.

In *C. scutulatus*, PLA₂s, SVMPs, CTLs, and MYOs are primarily responsible for the transcriptomic diversity within and between the neurotoxic (Type A) and hemorrhagic (Type B) venom phenotypes. Many, but not all toxins within these four families are associated with the difference between the two phenotypes. Variation in myotoxins was found across the sampled

range of *C. scutulatus* irrespective of venom type and was not exclusively found at contact zones between the two types. Their diversity among the nine individuals supports the further division of venom types within *C. scutulatus* and potentially all rattlesnakes to account for myotoxin diversity and expression level which is obscured when only considering the relationship between MTX and SVMPs. Our work represents the first complete venom gland transcriptome analysis in *C. scutulatus* and the best representation of species polymorphic for both Type A and Type B venoms. These data support utilizing Mojave Rattlesnakes as a model for understanding the molecular mechanisms driving the evolution of phenotypic diversity.

Acknowledgments

Support and assistance in the field was provided by B. Ashley, T. Burkhardt, H. Dahn, M. Feldner, T. Fisher, R. Govreau, J. Houck, T. Jones, P. Lindsey, M. Linsalata, C. May, S. May, R. Mayerhofer, E. McCormick, J. McNally, D. Ortiz, A. Owens, R. Rautsaw, J. Slone, D. Speckin, G. Territo, B. Townsend, K. VanSooy, C. Vratil, D. Weber, and W. Wüster. We thank M. Hogan, M. Margres, J. McGivern, M. Ward in the Rokyta Lab, M. Seavy and B. Washburn (FSU Department of Biological Science Analytical Lab), and S. Miller (FSU DNA Sequencing Facility), for assistance with preparing the samples and generating data. We thank D. Greenspan, N. Lucas, J. Schnaitter, and P. Wiegand (UCF ARCC) for assistance with Stokes HPC and P. Larabee (UCF IT) for assistance with operating system compatibility. We thank T. Jones and C. Kondrat-Smith with Arizona Game and Fish Department and S. Ferguson with the New Mexico Department of Game and Fish with assistance in obtaining collecting permits. We thank C. Johnston at the Arizona State University Natural History Museum and M. Revelez at the Angelo State Natural History Collection for assistance with depositing specimens. Comments and suggestions were given by Erich Hofmann and Rhett Rautsaw which greatly improved the quality of this manuscript. Financial support was generously provided by the National Science Foundation to D.R.R. (DEB 1145987

and DEB 1638902) and C.L.P. (DUE 1161228 which funded J.L.S. and A.J.M. and DEB 1638879), Prairie Biotic Research Inc. (J.L.S.), Sigma Xi Grants-in-aid-of-research (J.L.S.), the SnakeDays Research Grant (J.L.S.), the Southwestern Association of Naturalists McCarley Research Grant (J.L.S.), and the Theodore Roosevelt Memorial Fund through the American Museum of Natural History (J.L.S. and A.J.M.).

References

1. Kussell, E.; Leibler, S. Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **2005**, *309*, 2075–2078.
2. Gallego-Romero, I.; Ruvinsky, I.; Gilad, Y. Comparative studies of gene expression and the evolution of gene regulation. *Nature Reviews Genetics* **2012**, *13*.
3. Savolainen, O.; Lascoux, M.; Merilä, J. Ecological genomics of local adaptation. *Nature Reviews Genetics* **2013**, *14*.
4. Margres, M.J.; Wray, K.P.; Hassinger, A.T.B.; Ward, M.J.; McGivern, J.J.; Moriarty-Lemmon, E.; Lemmon, A.R.; Rokyta, D.R. Quantity, not quality: rapid adaptation in a polygenic trait proceeded exclusively through expression differentiation. *Molecular Biology and Evolution* **2017**, *34*, 3099–3110.
5. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex cocktails: the evolutionary novelty of venoms. *Trends in Ecology & Evolution* **2013**, *28*, 219–229.
6. Margres, M.J.; McGivern, J.J.; Seavy, M.; Wray, K.P.; Facente, J.; Rokyta, D.R. Contrasting modes and tempos of venom expression evolution in two snake species. *Genetics* **2015**, *199*, 165–176.
7. Dowell, N.L.; Giorgianni, M.W.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. The deep origin and recent loss of venom toxin genes in rattlesnakes. *Current Biology* **2016**, *26*, 2434–2445.

8. Margres, M.J.; Aronow, K.; Loyacano, J.; Rokyta, D.R. The venom-gland transcriptome of the Eastern Coral Snake (*Micrurus fulvius*) reveals high venom complexity in the intragenomic evolution of venoms. *BMC Genomics* **2013**, *14*, 531.
9. Whittington, A.C.; Mason, A.J.; Rokyta, D.R. A single mutation unlocks cascading exaptations in the origin of a potent pitviper neurotoxin. *Molecular Biology and Evolution* **2018**, p. msx334.
10. Casewell, N.R.; Wagstaff, S.C.; Wüster, W.; Cook, D.A.N.; Bolton, F.M.S.; King, S.I.; Pla, D.; Sanz, L.; Calvete, J.J.; Harrison, R.A. Medically important differences in snake venom composition are dictated by distinct postgenomic mechanisms. *Proceedings of the National Academy of Sciences* **2014**, *111*, 9205–9210.
11. Rokyta, D.R.; Margres, M.J.; Calvin, K. Post-transcriptional mechanisms contribute little to phenotypic variation in snake venoms. *G3* **2015**, *5*, 2375–2382.
12. Glenn, J.; Straight, R. Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) venom: variation in toxicity with geographical origin. *Toxicon* **1978**, *16*, 81–84.
13. Glenn, J.L.; Straight, R.C. The rattlesnakes and their venom yield and lethal toxicity. In *Rattlesnake Venoms: Their Actions and Treatment.*; Tu, A.T., Ed.; Marcel Dekker, Inc.: New York, 1982; pp. 3–119.
14. Weinstein, S.A.; Minton, S.A.; Wilde, C.E. The distribution among ophidian venoms of a toxin isolated from the venom of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*). *Toxicon* **1985**, *23*, 825–844.
15. Mackessy, S.P. Venom composition in rattlesnakes: trends and biological significance. In *The Biology of Rattlesnakes*; Hayes, W.K.; Beaman, K.R.; Cardwell, M.D.; Bush, S.P., Eds.; Loma Linda University Press: Loma Linda, CA, 2008; pp. 495–510.
16. Doley, R.; Kini, R.M. Protein complexes in snake venom. *Cellular and Molecular Life Sciences* **2009**, *66*, 2851–2871.
17. Doley, R.; Zhou, X.; Kini, R. Snake venom phospholipase A2 enzymes. In *Handbook*

- of Venoms and Toxins of Reptiles*; Mackessy, S.P., Ed.; CRC Press: Boca Raton, Florida, 2009; pp. 173–205.
18. Wooldridge, B.; Pineda, G.; Banuelas-Ornelas, J.; Dagda, R.; Gasanov, S.; Rael, E.; Lieb, C. Mojave Rattlesnakes (*Crotalus scutulatus scutulatus*) lacking the acidic subunit DNA sequence lack Mojave toxin in their venom. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **2001**, *130*, 169–179.
 19. Zancolli, G.; Baker, T.; Barlow, A.; Bradley, R.; Calvete, J.; Carter, K.; de Jager, K.; Owens, J.; Price, J.; Sanz, L.; Scholes-Higham, A.; Shier, L.; Wood, L.; Wüster, C.; Wüster, W. Is hybridization a source of adaptive venom variation in rattlesnakes? A test, using a *Crotalus scutulatus* x *viridis* hybrid zone in southwestern New Mexico. *Toxins* **2016**, *8*, 188.
 20. Borja, M.; Neri-Castro, E.; Castañeda-Gaytán, G.; Strickland, J.; Parkinson, C.; Castañeda-Gaytán, J.; Ponce-López, R.; Lomonte, B.; Olvera-Rodríguez, A.; Alagón, A.; Pérez-Morales, R. Biological and proteolytic variation in the venom of *Crotalus scutulatus scutulatus* from Mexico. *Toxins* **2018**, *10*, 35.
 21. Bieber, A.L.; Tu, T.; Tu, A.T. Studies of an acidic cardiotoxin isolated from the venom of Mojave Rattlesnake (*Crotalus scutulatus*). *Biochimica et Biophysica Acta - Protein Structure* **1975**, *400*, 178–188.
 22. Cate, R.L.; Bieber, A.L. Purification and characterization of Mojave (*Crotalus scutulatus scutulatus*) toxin and its subunits. *Archives of Biochemistry and Biophysics* **1978**, *189*, 397–408.
 23. Gopalakrishnakone, P.; Hawgood, B.J.; Holbrooke, S.E.; Marsh, N.A.; Santana De Sa, S.; Tu, A.T. Sites of action of Mojave toxin isolated from the venom of the Mojave Rattlesnake. *British Journal of Pharmacology* **1980**, *69*, 421–31.
 24. John, T.R.; Smith, L.A.; Kaiser, I.I. Genomic sequences encoding the acidic and basic subunits of Mojave toxin: unusually high sequence identity of non-coding regions. *Gene*

- 1994**, *139*, 229–34.
25. Calvete, J.J.; Pérez, A.; Lomonte, B.; Sánchez, E.E.; Sanz, L. 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. *Journal of Proteome Research* **2012**, *11*, 1382–1390.
 26. Uetz, P. The Reptile Database. Available at <http://www.reptile-database.org/> and Accessed 24 January 2018.
 27. Rokyta, D.R.; Wray, K.P.; Margres, M.J. The genesis of an exceptionally lethal venom in the Timber Rattlesnake (*Crotalus horridus*) revealed through comparative venom-gland transcriptomics. *BMC Genomics* **2013**, *14*, 394.
 28. Rokyta, D.R.; Wray, K.P.; McGivern, J.J.; Margres, M.J. The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the Timber Rattlesnake (*Crotalus horridus*). *Toxicon* **2015**, *98*, 34–48.
 29. Minton, S.A. Observations on the amphibians and reptiles of the Big Bend region of Texas. *The Southwestern Naturalist* **1958**, *3*, 28–54.
 30. Martinez, M.; Rael, E.D.; Maddux, N.L. Isolation of a hemorrhagic toxin from Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) venom. *Toxicon* **1990**, *28*, 685–694.
 31. Wilkinson, J.A.; Glenn, J.L.; Straight, R.C.; Sites, J.W. Distribution and genetic variation in venom A and B populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Herpetologica* **1991**, *47*, 54 – 68.
 32. Borja, M.; Castañeda, G.; Espinosa, J.; Neri, E.; Carbajal, A.; Clement, H.; García, O.; Alagon, A. Mojave Rattlesnake *Crotalus scutulatus scutulatus* with Type B Venom from Mexico. *Copeia* **2014**, *2014*, 7–13.
 33. Dobson, J.; Yang, D.; Op Den Brouw, B.; Cochran, C.; Huynh, T.; Kurrupu, S.; Sánchez, E.E.; Massey, D.J.; Baumann, K.; Jackson, T.N.W.; Nouwens, A.; Josh, P.; Neri-Castro, E.; Alagón, A.; Hodgson, W.C.; Fry, B.G. Rattling the border

- wall: pathophysiological implications of functional and proteomic venom variation between Mexican and US subspecies of the desert rattlesnake *Crotalus scutulatus* **2017**. doi:10.1016/j.cbpc.2017.10.008.
34. Schield, D.R.; Card, D.C.; Adams, R.H.; Corbin, A.; Jezkova, T.; Hales, N.; Meik, J.M.; Spencer, C.L.; Smith, L.; Campillo-Garcia, G.; Bouzid, N.; Strickland, J.L.; Parkinson, C.L.; Flores-Villela, O.; Mackessy, S.P.; Castoe, T.A. Cryptic genetic diversity, population structure, and gene flow in the Mojave Rattlesnake (*Crotalus scutulatus*). *Molecular Phylogenetics and Evolution* **In Review**.
 35. Glenn, J.L.; Straight, R.C. Intergradation of two different venom populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Toxicon* **1989**, *27*, 411–8.
 36. Sánchez, E.E.; Galán, J.A.; Powell, R.L.; Reyes, S.R.; Soto, J.G.; Russell, W.K.; Russell, D.H.; Pérez, J.C. Disintegrin, hemorrhagic, and proteolytic activities of Mohave Rattlesnake, *Crotalus scutulatus scutulatus* venoms lacking Mojave toxin. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **2005**, *141*, 124–132.
 37. Githens, T.; George, I. Comparative studies on the venoms of certain rattlesnakes. *Bulletin of the Antivenin Institute of America* **1931**, *5*, 31–34.
 38. Hardy, D.L. Envenomation by the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in southern Arizona, U.S.A. *Toxicon* **1983**, *21*, 111–118.
 39. Massey, D.J.; Calvete, J.J.; Sánchez, E.E.; Sanz, L.; Richards, K.; Curtis, R.; Boesen, K. Venom variability and envenoming severity outcomes of the *Crotalus scutulatus scutulatus* (Mojave Rattlesnake) from Southern Arizona. *Journal of Proteomics* **2012**, *75*, 2576–2587.
 40. Glenn, J.; Straight, R.; Wolt, T. Regional variation in the presence of canebrake toxin in *Crotalus horridus* venom. *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology* **1994**, *107*, 337–346.
 41. Rotenberg, D.; Bamberger, E.S.; Kochva, E. Studies on ribonucleic acid synthesis in the venom glands of *Vipera palaestinae* (Ophidia, Reptilia). *The Biochemical Journal* **1971**,

- 121, 609–12.
42. Margres, M.J.; McGivern, J.J.; Wray, K.P.; Seavy, M.; Calvin, K.; Rokyta, D.R. Linking the transcriptome and proteome to characterize the venom of the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*). *Journal of Proteomics* **2014**, *96*, 145–158.
 43. Rokyta, D.R.; Wray, K.P.; Lemmon, A.R.; Moriarty-Lemmon, E.; Caudle, S.B. A high-throughput venom-gland transcriptome for the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*) and evidence for pervasive positive selection across toxin classes. *Toxicon* **2011**, *57*, 657–671.
 44. Rokyta, D.R.; Margres, M.J.; Ward, M.J.; Sanchez, E.E. The genetics of venom ontogeny in the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*). *PeerJ* **2017**, *5*, e3249.
 45. Rokyta, D.R.; Ward, M.J. Venom-gland transcriptomics and venom proteomics of the Black-back Scorpion (*Hadrurus spadix*) reveal detectability challenges and an unexplored realm of animal toxin diversity. *Toxicon : official journal of the International Society on Toxinology* **2017**, *128*, 23–37.
 46. Ward, M.J.; Ellsworth, S.A.; Rokyta, D.R. Venom-gland transcriptomics and venom proteomics of the Hentz Striped Scorpion (*Centruroides hentzi*; Buthidae) reveal high toxin diversity in a harmless member of a lethal family **2018**. *142*, 14–29.
 47. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **2008**, *5*, 621–628.
 48. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **2014**, *30*, 614–620.
 49. Rokyta, D.R.; Lemmon, A.R.; Margres, M.J.; Aronow, K. The venom-gland transcriptome of the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*). *BMC Genomics* **2012**, *13*, 312.
 50. Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659.

51. Bendtsen, J.D.; Nielsen, H.; Van Heijne, G.; Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **2004**, *340*, 783–795.
52. Petersen, T.N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **2011**, *8*, 785–786.
53. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **2012**, *9*, 357–359.
54. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **1994**, *22*, 4673–4680.
55. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Research* **2009**, *10*, R25.
56. Schrider, D.R.; Gout, J.F.; Hahn, M.W. Very few RNA and DNA sequence differences in the human transcriptome. *PLoS ONE* **2011**, *6*, e25842.
57. Palarea-Albaladejo, J.; Martín-Fernández, J.A. Software description zCompositions R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems* **2015**, *143*, 85–96.
58. Aitchison, J. *The Statistical Analysis of Compositional Data*; Vol. 44, Chapman and Hall: London, 1986; pp. 139–177.
59. Anders, S.; Huber, W. Differential expression analysis for sequence count data. *Genome Biology* **2010**, *11*, R106.
60. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **2014**, *15*, 550.
61. Dagda, R.K.; Gasanov, S.; De La Oiii, Y.; Rael, E.D.; Lieb, C.S. Genetic basis for variation of metalloproteinase-associated biochemical activity in venom of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*). *Biochemistry research international* **2013**, *2013*, 251474.

62. Dagda, R.K.; Gasanov, S.E.; Zhang, B.; Welch, W.; Rael, E.D. Molecular models of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) venom metalloproteinases reveal a structural basis for differences in hemorrhagic activities. *Journal of Biological Physics* **2014**, *40*, 193–216.
63. Arlinghaus, F.T.; Eble, J.A. C-type lectin-like proteins from snake venoms. *Toxicon* **2012**, *60*, 512–519.
64. Chang, C.C.; Tseng, K.H. Effect of crotamine, a toxin of South American rattlesnake venom, on the sodium channel of murine skeletal muscle. *British Journal of Pharmacology* **1978**, *63*, 551–559.
65. Durban, J.; Pérez, A.; Sanz, L.; Gómez, A.; Bonilla, F.; Rodríguez, S.; Chacón, D.; Sasa, M.; Angulo, Y.; Gutiérrez, J.M.; Calvete, J.J. Integrated "omics" profiling indicates that miRNAs are modulators of the ontogenetic venom composition shift in the Central American rattlesnake, *Crotalus simus simus*. *BMC Genomics* **2013**, *14*, 234.
66. Lynch, V.J. Inventing an arsenal: adaptive evolution and neofunctionalization of snake venom phospholipase A2 genes. *BMC Evolutionary Biology* **2007**, *7*, 2.

CHAPTER 4: ADDITIVE EXPRESSION SUGGESTS MENDELIAN INHERITANCE OF POLYMORPHIC VENOM PHENOTYPES IN MOJAVE RATTLESNAKES (*CROTALUS SCUTULATUS*)³

Abstract

The combination of phenotypes inherited from parents partially determines fitness of an individual. Monomorphic phenotypes under strong selection tend to move to fitness optima but polymorphic phenotypes that can be inherited in different combinations will result in offspring with fitness that differs from one or both parents depending on the complexity of the genotype to phenotype pathway. Venom is a polymorphic phenotype that is under strong selection and highly tractable from the genotype to the phenotype. In rattlesnakes, particularly Mojave Rattlesnakes (*Crotalus scutulatus*), a venom dichotomy exists. This dichotomy was initially defined based on proteomic characteristics but recently, a genomic definition has been proposed. Some *C. scutulatus* have Type A venom which is neurotoxic and lacks hemorrhagic activity (Type A) and others have high hemorrhagic activity and no neurotoxic activity (Type B). Rarely, individuals have a third venom type, Type A + B, where they have neurotoxic and hemorrhagic venom. We used the proteomic definition to identify putative Type A + B *C. scutulatus* to test the applicability of the genomic definition of venom type. To do this, we used comparative transcriptomics of the venom-gland on 15 *C. scutulatus* and tested for differential expression between venom types. We then applied the genomic definition of the two phenotypes and were able to identify the genotype for the phospholipases and snake venom metalloproteinases responsible for the venom phenotypes. The heterozygous individuals expressed the toxins at approximately half of the level found in homozygous individuals. Additive expression in the eight toxins tested suggest the phenotypes in

³Chapter Four is being prepared as Strickland et al. Additive expression suggests Mendelian inheritance of polymorphic venom phenotypes in Mojave Rattlesnakes (*Crotalus scutulatus*)

C. scutulatus are inherited in a Mendelian fashion and *cis*-regulation is responsible for differences in expression. We identified six of the nine possible genotypes for a dihybrid cross including five from the F2 generation or later. By testing for additive expression, we were able to determine that the Type A + B venom type is not a unique venom type and is generated through interbreeding between Type A and Type B individuals.

Keywords: Allele, Hemorrhagic, Mojave toxin, Neurotoxic, Phospholipase A₂, RNA-seq, Transcriptome, Snake Venom Metalloproteinases

Introduction

An individual's fitness is partially determined by the specific combination of phenotypes inherited from its parents [1, 2]. Selection moves phenotypes towards fitness optima, driving the frequency of these phenotypes up or down depending on what is selectively advantageous [3, 4]. However, individuals may also demonstrate mixed phenotypes, containing components of discrete phenotypes which can have higher, lower, or equivalent fitness to either parent based on the costs and benefits associated with the phenotype [5]. These mixed phenotypes may arise through a variety of mechanisms and differ in inheritance complexity [6]. When multiple alleles of a single trait generate a variety of phenotypes, but are tractable to a single genetic mechanism, such mixed traits may be simply inherited via co-dominance or incomplete dominance [1]. On the other hand, the genetic mechanism may be less clear when multiple loci are involved, potentially resulting in a wide range of offspring phenotypes [5]. Such phenotypes could incorporate traits from both parents, are intermediate between the two parents, or are superior to both parents (heterosis) [6].

Mixed phenotypes are often referred to as "hybrid" phenotypes, which posits an underlying assumption about their generation. If the parents are from two distinct genetic lineages (populations or species) which come together at hybrid zones and exchange genes, the resulting phenotypes are appropriately referred to as 'hybrid' phenotypes because they are the result of a

hybridization event [7, 8]. The breadth of the hybrid zone is constrained by the fitness of the hybrid in comparison to the parental lineages [9]. However, within a single polymorphic population, individuals can contain components from prevailing phenotypes (hereafter integrated phenotypes). Integrated phenotypes may be the result of mating between parents from the same population who exhibit different phenotypes or due to differential expression of the loci involved in the phenotype through regulatory changes [10]. By comparing integrated phenotypes to the primary phenotypes, it may be possible to infer inheritance patterns and regulatory mechanisms responsible for the differing phenotypes [11]. In the simplest case – where discrete phenotypes are inherited in a Mendelian fashion – heterozygous individuals should express proteins at 50% of the homozygous dominant individuals if expression variation is additive [2, 11] which would indicate *cis*-regulation [10, 12]. Unfortunately, comparing relative expression to test for additive expression is challenging because few phenotypes are easily linked to the genotype [13].

Venom is a genetically tractable trait that is well suited for testing differences in expression and inferring inheritance patterns. The Type A/B venom dichotomy found within rattlesnakes (*Crotalus* and *Sistrurus*) provides an ideal system to test for additive expression. As currently understood, one species of rattlesnake exhibits exclusively Type A venom, 38 exhibit exclusively Type B, and nine species have demonstrated cases of both Type A or Type B intraspecifically [14, 15]. Type A venoms are characterized by the presence of a potent neurotoxin that is a heterodimeric phospholipase A₂ (Mojave toxin and homologs) that acts presynaptically and is highly toxic [16–19]. Type B venoms lack the neurotoxin, are less toxic, and have high hemorrhagic activity due to the prevalence of snake venom metalloproteinases [20–23]. The maintenance of the dichotomy is hypothesized to be through balancing selection [24, 25] because of the important ecological role venom plays in prey acquisition and defense [26]. However, the venom phenotype dichotomy lacks both phylogenetic signal across rattlesnakes [15] and phylogeographic signal within polymorphic species where tested [25].

These phenotypes were originally defined by proteomic composition [19, 27–31]; however,

recent genetic work has indicated that there may be discrete genomic definitions as well [24, 32]. Dowell *et al.* [32] hypothesized that there are two distinct suites of PLA₂ loci for each venom type. Type A venom has both subunits of Mojave Toxin (Pla2-gA2-MTXA and Pla2-gB2-MTXB) and Type B venom has Pla2-gK and Pla2-gB1 [24]. Two other PLA₂s, Pla2-gA1 and Pla2-gC1 are found in both venom types and Strickland *et al.* [33] described a sixth PLA₂, PLA₂-6, also found in both venom types. Strickland *et al.* [33] also found Type A individuals with hypothesized Type B PLA₂s using transcriptomic data; however, if there are two suites of PLA₂s, it might be possible for an individual to be heterozygous at the PLA₂ loci. For the SVMP loci, Dowell *et al.* [24] found that *C. scutulatus* with Type B venom had 11 more SVMP loci than Type A individuals due to a large deletion in the metalloproteinase region of the genome in Type A individuals. By testing for the presence of the loci only found in Type B, it would be possible to genomically define Type B individuals and not rely on the more variable metalloproteinase activity assays.

In three instances, a single individual has been found to exhibit characteristics of both venom types where Mojave toxin (or homolog) is expressed and the individual has high metalloproteinase activity. The first example involves changing the proportion of the two components during ontogeny: *Crotalus simus* shifts from Type A to Type B venom as they grow in size [34]. The second example involves hybridization between a Type B individual from the *C. viridis* complex and a Type A individual of *C. scutulatus*. The hybridization event has been documented in captivity [35] and at a narrow hybrid zone [36, 37]. The resulting progeny have characteristics of both parent phenotypes [35, 37]. The final example of Type A + B phenotypes existing within an individual occur in *C. horridus* [38–40] and *C. scutulatus* [25, 41–43]. These two species have individuals with both Type A and Type B venom in their distribution with geographic structuring [27, 42]. The primary hypothesis for the occurrence of the Type A + B phenotype is that it is the result of mating between a Type A parent and a Type B parent [44]. However, Strickland *et al.* [25] found *C. scutulatus* individuals with Type A + B venom in locations that were not contact zones between Type A and Type B. Additionally, *C. scutulatus* in the Sonoran Desert, where Type

A + B individuals are found most frequently, are genetically indistinguishable even though venom type is geographically structured in the region [25, 45]. Because venom is highly tractable from the genotype to phenotype [46] it may be possible to determine the inheritance pattern of the Type A and Type B phenotypes by using Type A + B individuals as intermediates.

To determine if we could infer the inheritance pattern of venom phenotypes in Mojave Rattlesnakes (*C. scutulatus*), we compared the venom gland transcriptomes of Type A, Type B, and Type A + B individuals. We used the proteomic definition of Type A and Type B venom to identify putative Type A + B *C. scutulatus* from the panmictic population in the Sonoran Desert. Using comparative transcriptomics, we tested the applicability of the genomic definition of venom type and assessed whether expression was additive depending on the transcripts present. We first determined if each individual was homozygous or heterozygous at the PLA₂ loci based on presence/absence and the SVMP regions based on relative expression. Differential expression could be higher in the Type A + B individuals (heterosis), lower due to incompatibility, or equal to the parents. If venom phenotypes are inherited in a Mendelian fashion, we predicted that heterozygous individuals would have approximately half the toxin expression compared to pure Type A and Type B individuals. We present the first test of inheritance patterns of multiple toxin loci and test the hypothesis that the Type A + B phenotype is generated through interbreeding of Type A and Type B individuals within the same species.

Materials and methods

Ethics Statement

Scientific collecting permits were issued by the New Mexico Department of Game and Fish (3563, 3576) and the State of Arizona Game and Fish Department (SP628489, SP673390, SP673626, SP715023). All interactions with animals were approved by UCFs Institutional Animal Care and Use Committee under protocol 13-17W and followed the American Society of

Ichthyologists and Herpetologists ethical guidelines.

Sampling and venom type determination

In 2015 and 2016, we targeted the zone between Type A and Type B venom in Arizona and specifically searched in locations where Type A + B individuals had been documented [42]. We collected venom and tissue from each *C. scutulatus*, preserved each specimen, and deposited them in the appropriate museum based on permit requirements (Table 4.1). We collected venom by coercing each snake into a polycarbonate tube (Get Hooked L.L.C., Sanford, FL, USA) of the appropriate size. They would travel through the tube until their head was extended outside the tube and then they were able to inject venom into a sterile collection container covered with parafilm by voluntarily biting the container. We then collected the venom and vacuum dried it using a vacuum sealed container (Vacu Vin Saver, Indian Trail, NC, USA). We sacrificed each snake four days after venom collection when transcription was maximized [47] using 100 mg/kg of sodium pentobarbital injected intracoelomically. We harvested venom glands and stored them separately in RNAlater (Thermo Fisher Scientific, Waltham, Massachusetts, USA) following the manufacturer’s protocol. We preserved the specimen in 10% buffered formalin for five days and then transferred to 70% ethanol before depositing in a natural history museum.

Table 4.1: Sample information for the six *C. scutulatus* sequenced from the U.S. Sonoran Desert. ASNHC = Angelo State Natural History Collection, San Angelo, TX; ASU = Arizona State University Natural History Museum, Tempe, Arizona; CLP = Christopher L. Parkinson Field Number

Specimen ID	Museum ID	Venom Type	Sex	SVL (mm)	Mass (g)	State	County	Sequencing Platform	Read Pairs	Merged Reads	BioSample ¹ Accession
CLP1823	ASU36060	A+B	F	597	146	AZ	Yavapai	HiSeq	10,153,085	8,985,119	SAMN08596264
CLP1829	ASU36074	A	M	802	364	AZ	Maricopa	HiSeq	13,392,954	11,447,955	SAMN08596265
CLP1830	ASU36088	A+B	M	736	270	AZ	Pima	HiSeq	12,588,955	10,927,477	SAMN08596266
CLP1832	ASU36090	A+B	M	765	248	AZ	Pima	HiSeq	13,444,339	11,657,787	SAMN08596268
CLP1929	ASNHC14996	A	F	283	18	NM	Hidalgo	HiSeq	15,312,266	13,215,275	SAMN08596270
CLP2111	ASU36063	A+B	F	618	139	AZ	Graham	HiSeq	13,190,119	11,038,527	SAMN08596276

¹ National Center for Biotechnology Information under BioProject PRJNA88989

To putatively identify individuals with Type A + B venom, we used reverse-phased high performance liquid chromatography (RP-HPLC) following Strickland *et al.* [33]. Briefly, we

resuspended the venom in millipore filtered water and then centrifuged to remove cellular debris and insoluble particles. We used the Qubit Protein Assay (Thermo Fisher Scientific) and Qubit 3.0 Fluorometer (Thermo Fisher Scientific) following the manufacture's protocol to determine venom concentration. Then, 100 μg of venom was injected onto a Jupiter C18 column (250 x 2 mm; Phenomenex, Torrance, California, USA) using two solvents: A = 0.1% trifluoroacetic acid (TFA) in water and B = 0.075% TFA in acetonitrile. Run time was 180 minutes, flow rate was 0.2 mL/min, and the effluent was monitored at 220 and 260 nm using a four step linear gradient on a Beckman System Gold HPLC (Beckman Coulter, Fullerton, California, USA): 95% A and 5% B for 5 minutes, 1% per minute linear increase to 25% B, a 0.25% per minute linear increase to 55% B, a 2% per minute linear increase to 75% B, a 14% per minute linear increase to 5% B, and then 5 minutes at the initial condition to allow the column to equilibrate. Individuals that were Type A + B had both peaks corresponding to the acidic and basic subunits of MTX and snake venom metalloproteinases.

Venom gland transcriptome sequencing and assembly

We used the nine venom-gland transcriptomes sequenced in Strickland *et al.* [33] and added six Type A + B individuals that were sequenced following the same protocol. Briefly, TRIzol-based RNA extraction was conducted on both both glands separately [33, 48]. Total RNA concentration was determined using the Qubit RNA BR Assay (Thermo Fisher Scientific) and quality assessed using an Agilent Bioanalyzer 2100 with an RNA 6000 Pico Kit (Agilent Technologies, Santa Clara, California, USA) following the manufacturer's protocol. RNA was combined from both glands in equal concentration [49].

To generate our cDNA libraries, 1 μg of total RNA was used initially for mRNA isolation with the New England Biolabs (NEB, Ipswich, Massachusetts, USA) NEBNext Poly(A) mRNA magnetic isolation kit (E7490S). We moved directly into cDNA library preparation with the NEBNext Ultra RNA Library Prep Kit for Illumina (E7530), and the NEBNext Multiplex Oligos

for Illumina (E7335 -Index primer set 1). We followed the protocol as described and targeted a mean insert size of 370bp by fragmenting for 15.5 minutes and using 14 PCR cycles for the last enrichment step. DNA library concentration was determined using the Qubit DNA BR Assay (Thermo Fisher Scientific) and quality assessed using a HS DNA Kit on the Bioanalyzer. To determine the amplifiable concentration for each sample, we used KAPA qPCR in the Florida State University Molecular Cloning Facility [33]. We combined samples in equal amounts, verified the amplifiable concentration again using KAPA qPCR, and then sequenced the pooled samples on an Illumina HiSeq 2500 at the Florida State University College of Medicine Translational Science Laboratory. The final concentration was ~ 10 nM for the pooled library and they were sequenced using 150 bp paired-end reads (Table 4.1) [40].

Strickland *et al.* [33] found 75 toxins and 1889 nontoxins in the consensus transcriptome of nine individuals from this region that were either Type A or Type B. To check for new transcripts in the six Type A + B individuals added, we used the same assembly protocol described by Rokyta *et al.* [49]. Briefly, we merged reads using PEAR v 0.9.6 [50] that passed the Illumina quality filter. Merged reads were assembled using Extender [51] and DNASTar SeqMan NGen version 12.3.1 (DNASTAR Inc., Madison, Wisconsin, USA) with default settings except that only contigs with ≥ 200 assembled reads were kept [33]. We annotated transcripts using the initial consensus transcriptome as well as a local database of compiled toxins published previously [39, 40, 48, 49, 51]. To be annotated, they had to match at $\geq 90\%$ using Cd-hit-est v.4.6 [52] and had at least 10X coverage across the transcript. Remaining unannotated contigs were compared against the curated Uniprot animal toxins database (downloaded 16 Nov 2017) via blastx v. 2.2.30+ searches (minimum e-value of 10^{-4}) [39, 40, 51, 53]. All annotation was done in Geneious v 10.1.2 (Biomatters Ltd., Auckland, New Zealand) and signal peptides were added using SignalP v 4.1 [54, 55]. Only transcripts with complete coding sequences were kept and we removed duplicates using the BBtools package Dedupe (Joint Genome Institute, Department of Energy, Walnut Creek, California, USA) implemented in Geneious. We removed chimeric sequences that had irregular

coverage after mapping in Bowtie2 v 2.3.0 [56] implemented in Geneious or if recombination was detected using the ClustalW alignment algorithm [57] implemented in Geneious. Remaining transcripts were clustered using Cd-hit at $\geq 98\%$ (-c .98) and any new transcripts were added to the initial consensus. Toxin names are the same as in Strickland *et al.* [33] and new toxins were added to the series in order of expression levels.

Gene expression patterns among venom types

We determined the expression level of all loci using Bowtie2 [56, 58] implemented in RSEM v1.3.0 [59] with default parameters on the Stokes HPC cluster at the UCF Advanced Research Computing Center. Merged reads were mapped to the consensus transcriptome and transcripts per million reads (TPM) values were used for further analyses as our estimate of transcript abundance [60]. We used R v. 3.4.2 (R Development Core Team 2006) implemented in RStudio v. 1.1.383 for further analyses. To remove 0 values and keep the compositional structure of the TPM data, we used the R package zCompositions and the cmultRepl function [61]. We used the pheatmap [62] package in R to create a heatmap of the toxins identified for all individuals. Both the individuals and the toxins were hierarchically-clustered based on similarity of expression (ln-transformed TPM) to examine overall toxin expression differences among venom types. To determine which toxins were responsible for the variability in the toxin transcriptome, we conducted a Principle Component Analysis (PCA) using the PCA function in the FactoMineR [63] R package. Individuals were grouped based on venom type and the centroid and 95% ellipse was calculated for each venom type.

Differential expression in Type A + B individuals

To determine if any loci were up- or down-regulated in Type A + B individuals, we tested for differential expression among venom types. Because of the lack of consensus on which method is best and because very few tests/programs allow for including three groups at once, we tested

differences between Type A, Type B, and Type A + B using four methods [64]. When it was not possible to look at all three venom types at once, we conducted three pairwise tests: Type A vs Type B, Type A vs Type A + B, and Type B vs Type A + B. For all analyses, we used TPM output from RSEM.

First, we utilized DESeq2 v. 1.14.1 [65] implemented in R with a false discovery rate (FDR) threshold of 0.05. DESeq2 can only be used for pairwise comparisons so all three comparisons were done independently. We used the TPM values generated in RSEM and venom type as the factor. The Wald significance tests with default settings and a local fit of dispersions was used to determine toxins that were differential expressed between venom types. The log-fold changes were used to determine direction of the expression differences.

Second, we used edgeR [66, 67] implemented in R with negative binomial generalized linear models. These models were fit to each gene and likelihood ratio tests were used to determine differential expression between venom types. We used the calcNormFactors function to normalize the TPM values and the estimateDisp function to estimate dispersion globally and locus-by-locus simultaneously. Our model included comparing all three venom types at once and pairwise log-fold changes were used to determine differential expression with an FDR threshold of 0.05.

Third, we used EBseq [68] implemented in R. Venom type was the factor and all three venom types were compared at once. We had replication for each treatment so variance was measured for each locus independently. We used the MedianNorm function to normalize the TPM values and then the EBMultiTest function to compare the three conditions. For a three treatment comparison, there can be five possible expression patterns: all treatments have the same expression, all treatments have different expression, and then each of the treatments can be unique in comparison to the other three treatments. EBseq does not directly measure the direction of the difference so that was calculated post-hoc by averaging TPM values for loci that had differential expression.

Fourth, we used RNentropy [64] implemented in R to determine which venom types were

expressed higher or lower compared to overall expression. We used the RN_calc function to determine which loci diverge from the uniform background distribution by comparing relative expression as opposed to absolute expression. This method calculates both global (among treatments) and local p-values (among replicates) and tests for over- or under-expression at each locus. The estimated FDR for this method is 0.01 [64].

Finally, we compared the Average A and Average B transcriptomes to the Average A + B transcriptome to identify differentially expressed toxins. We linearized the compositional data and preserved the rank order of the transcripts by centered log-ratio (clr) transforming the expression dataset [60, 69]. We used the nontoxin expression values to generate a null distribution of expression divergence [49]. This was done by taking the absolute value of the difference in the transformed data for the two individuals being compared and finding the 99th percentile value. Any toxin outside of this value was identified as an outlier to the null distribution. For each pairwise comparison, we used the Spearman's rank correlation coefficient (ρ), Pearson's correlation coefficient (R), and coefficient of determination (R^2) to examine the similarity of individuals being compared.

PLA₂ and SVMP diversity and inheritance testing

To determine which PLA₂s and SVMPs were expressed in the transcriptomes of *C. scutulatus* that were also found in the genome by Dowell *et al.* [32] and Dowell *et al.* [24], we compiled the coding sequence identified and added them to the PLA₂ and SVMP loci in our consensus transcriptome. We clustered all sequences using Cd-hit at $\geq 90\%$ ($-c .90$) and a representative transcript was retained for each cluster. We then used Bowtie2 implemented in RSEM to align merged reads from each individual to 10 PLA₂s and 20 SVMPs. Presence in the transcriptome required at least 5X coverage over 90% transcript.

We conducted a PCA analysis as above on the four PLA₂ loci and the four SVMP loci that are in the genome and where expression was found to be high by Dowell *et al.* [24]. Individuals

that had both subunits of MTX (Pla2-gA2 and Pla2-gB2), Pla2-gK, and Pla2-gB1 expressed were identified as heterozygotes at the PLA₂ loci. For the SVMP loci, if SVMPIII-1 (mp238 in [24]) was not expressed, the individual was considered homozygous Type A. To differentiate between putative heterozygotes and homozygous Type B individuals, we compared relative expression of the four loci found exclusively in Type B individuals and were expressed at high levels: SVMPIII-1/mp238, SVMPII-1/mp237, SVMPIII-3/mp240, SVMPIII-5/mp242. We calculated the average expression for these four loci when present and then used a ratio of the individual to the average expression. If the ratio was approximately 2:1, they were classified as heterozygous.

Data availability

Raw data for the six venom gland transcriptomes added in this study were submitted to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) accession SRP011323. BioSample accession numbers are provided in Table 4.1 and are under BioProject PRJNA88989. The consensus transcriptome was submitted to the NCBI Transcriptome Shotgun Assembly (TSA) database. This TSA project has been deposited at DDBJ/EMBL/GenBank under the accession GGIP00000000. The version described in this paper is the first version, GGIP01000000.

Results

Consensus transcriptome of Crotalus scutulatus

We sequenced the venom-gland transcriptome of six *C. scutulatus* that were putatively identified as venom Type A + B based on RP-HPLC profiles (Figure 4.1). We generated over 78 million raw read pairs and used over 67 million quality reads with an average number of reads per individual of 11.2 million (Table 4.1). After we assembled, annotated, removed chimeric sequences and duplicates for each individual, these transcriptomes were combined with the

consensus transcriptome of Strickland *et al.* [33]. We removed duplicates and had a final consensus transcriptome of 2202 putative nontoxins and 85 putative toxins. This was an increase of 313 nontoxins and nine toxins from Strickland *et al.* [33]. These nine toxins included two SVSPs (SVSP-15 and -16), SVMPI-1, three SVMPIIs (2-4), Pla2-gC1, and two 3FTxs which is a toxin family not recovered by Strickland *et al.* [33]. SVSP-15 had the third highest average expression for all fifteen individuals and was missed previously [33]. SVSP-16 was highly expressed in CLP1930 newly added in this study but present in at least four other individuals and had the tenth highest average expression overall. The 3FTxs were very lowly expressed. 3Ftx-1 was only expressed in CLP1823 and CLP1829 whereas 3Ftx-2 was only expressed in CLP1829. The SVMPIs and Pla2-gC1 are discussed below.

Venom type expression patterns

Of the six individuals sequenced that were putatively Type A + B, CLP1829 and CLP1929 expressed SVMPIs consistent with Type A individuals and not Type B individuals so they were reclassified as Type A. The remaining four individuals expressed SVMPIs consistent with Type B individuals and MTX consistent with Type A individuals, thus were classified as Type A + B (Figure 4.2). The average Type A + B transcriptome was created using CLP1823, CLP1830, CLP1832, and CLP2111 (Figure 4.3).

The clustered heatmap of toxin expression separated the individuals into two clades based on similarity in expression profiles. These were essentially a Type A and a Type B clade. Three of the four Type A + B individuals clustered with the Type B individuals and one, CLP2111, clustered with the Type A individuals (Figure 4.4). SVMPIs and CTLs were primarily responsible for the differentiation between the two clades. CLP2111 clustered with Type A individuals because it was missing many of the CTLs that the remaining Type A + B individuals and the Type B individuals have. These CTLs are generally associated with Type B individuals but were also found in CLP1959 and CLP1961 so are not exclusively associated.

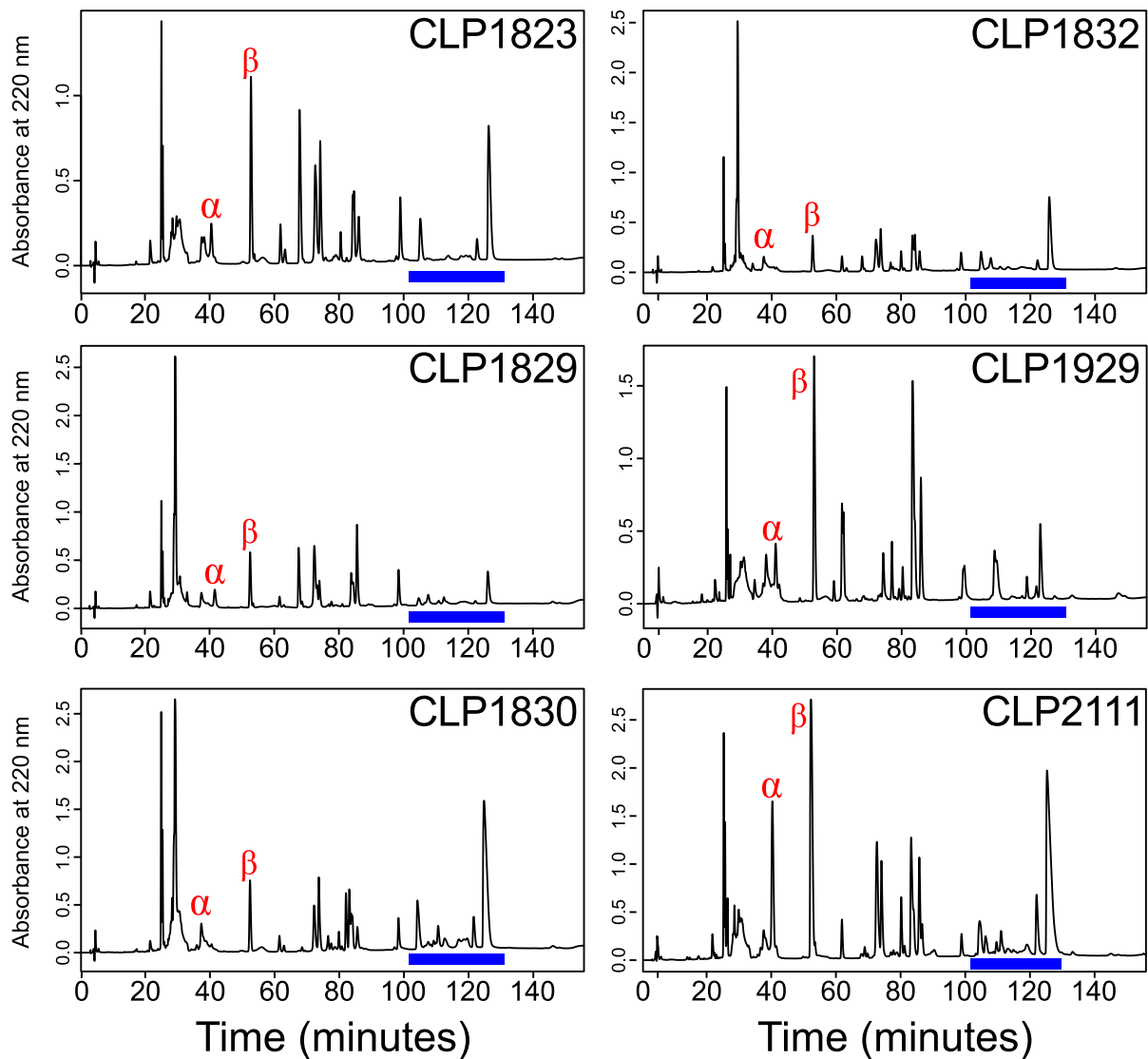


Figure 4.1: Reverse-phased high performance liquid chromatography (RP-HPLC) profiles of Mojave Rattlesnakes. The acidic (α) and basic subunit (β) peaks for Mojave toxin are marked and the region where snake venom metalloproteinases elute is marked with a blue bar.

The three venom types had significantly different clusters in our PCA analysis and none of the 95% confidence ellipses overlapped (Figure 4.5). The first axis explained 26.49% of the variation in the toxin expression differences and corresponded with the Type A and Type B dichotomy. The second axis explained 23.48% of the variation and corresponded with presence and absence of MYO-1 which is homologous with Crostamine [70]. The pattern of the Type A

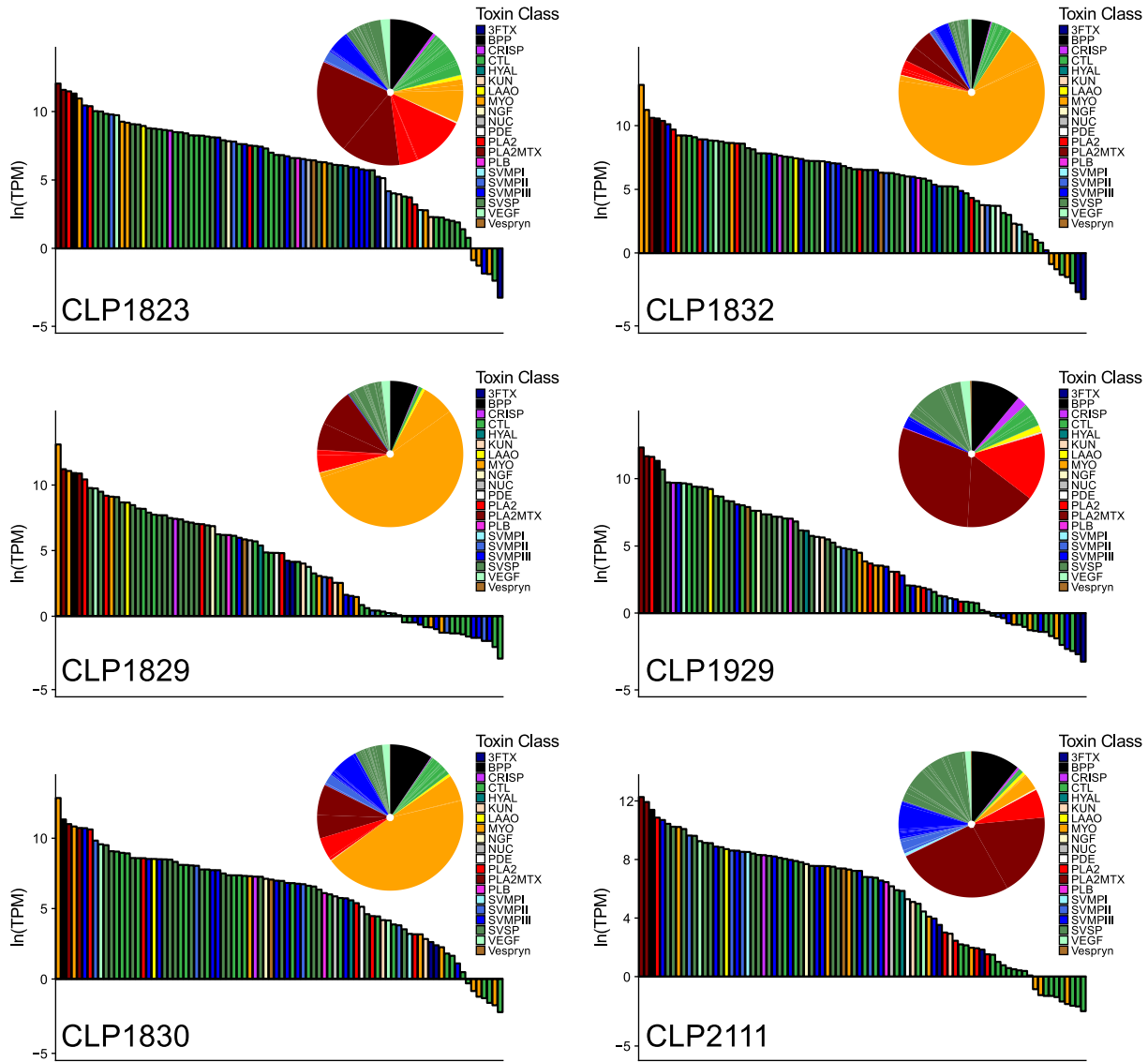


Figure 4.2: Representation of the toxins in the venom gland transcriptome for six *C. scutulatus* added in this study. The bar graphs represent each of the 84 toxins identified. The pie charts represent the proportion of each toxin family in the venom gland transcriptome. 3FTX: three-finger toxin; BPP: bradykinin potentiating peptide; CRISP: cysteine-rich secretory protein; CTL: C-type lectin; HYAL: hyaluronidase; KUN: Kunitz peptide; LAO: L-amino-acid oxidase; MYO: myotoxin; NGF: nerve growth factor; NUC: 5' nucleotidase; PDE: phosphodiesterase; PLA₂: Phospholipase A₂; PLB: phospholipase B; SVMP: snake venom metalloproteinase; SVSP: snake venom serine protease; VEGF: vascular endothelial growth factor; MTX: Mojave toxin.

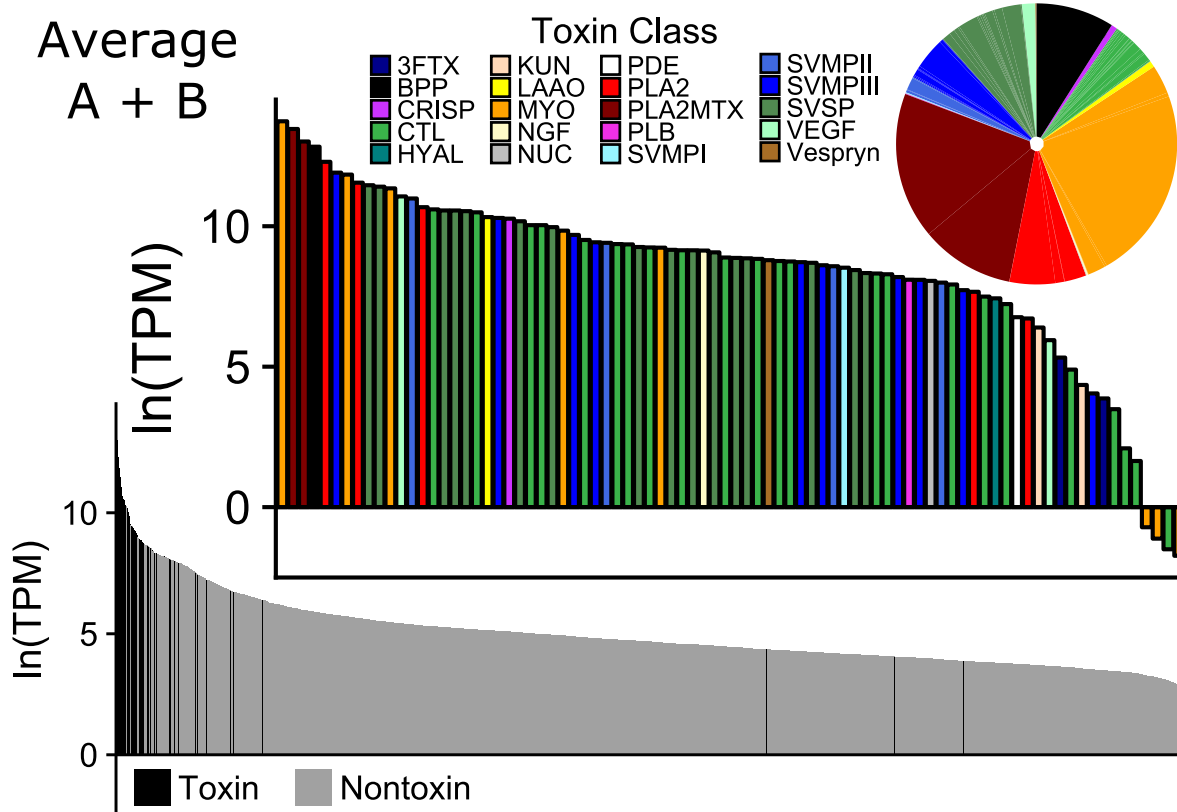


Figure 4.3: Representation of the average Type A + B transcriptome of the four *C. scutulatus* with Type A + B venom. The majority of highly expressed transcripts were the 84 toxins identified. Pie charts represent the proportion of each of the 19 toxin families identified.

+ B individuals being intermediate between the Type A and Type B clusters was driven by the presence/absence of MTX, SVMPs, and CTLs rather than intermediate expression of toxin loci per se (discussed below).

Differential expression among and between venom types

The four analyses conducted testing for differential expression among and between venom types were largely consistent (Figure 4.6). A total of 61 toxins were identified as differentially expressed by either DESeq2, edgeR, EBseq, and/or RNentropy (Supplemental Table C.1). DESeq2, edgeR, and EBSeq were more conservative and identified 18, 24, and 31 differentially

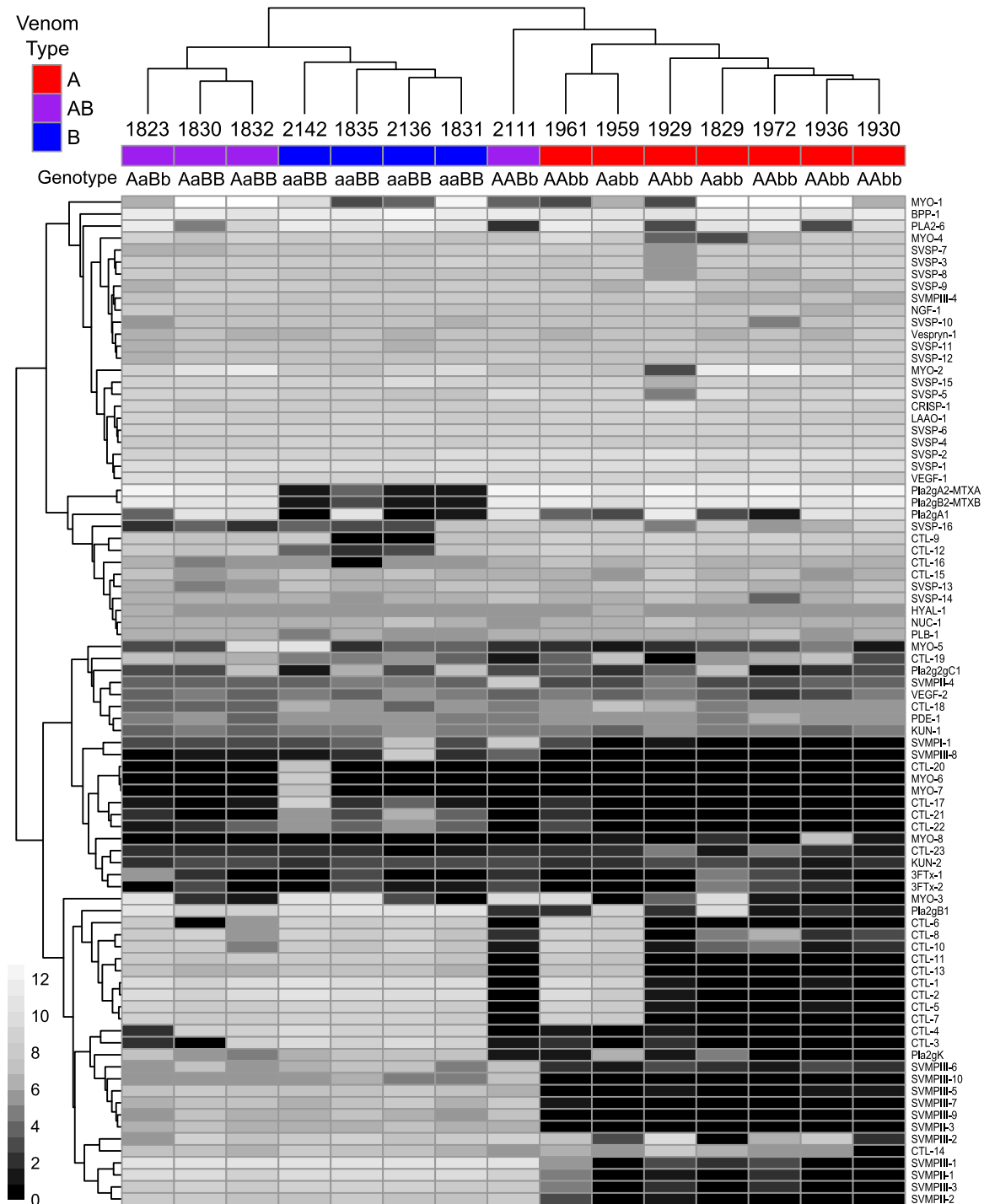


Figure 4.4: Hierarchical clustered heatmap of all 84 toxins from the 15 individuals of *C. scutellatus* with their venom gland transcriptome sequenced based on ln-transformed TPM data. Venom type is designated for each individual and dark colors are low expression and lighter colors are higher expression. The putative genotype for each individual is indicated below venom type.

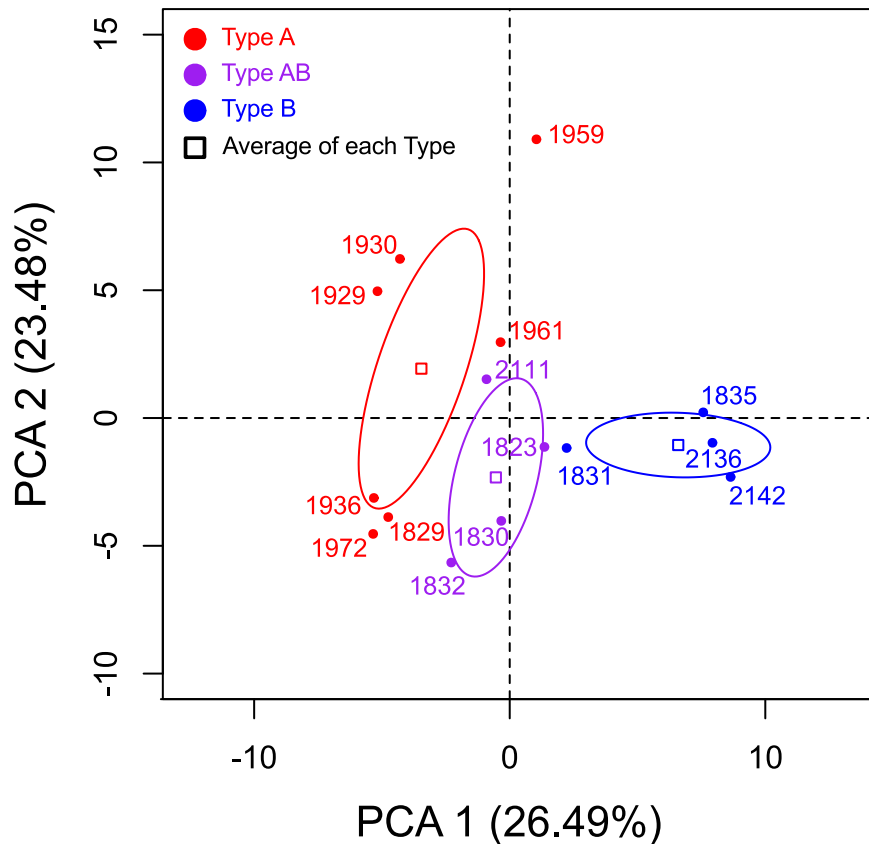


Figure 4.5: Principle component analysis (PCA) scatterplot of the 15 *C. scutulatus* from the Sonoran Desert based on TPM data for 84 toxins. The squares are the average value for each venom type and the circles are the 95% confidence ellipse for each venom type. The first two axes explain 49.97% of the variability. PCA 1 represents the Type A (left) and Type B (right) dichotomy and PCA 2 represents the presence (bottom) and absence (top) of myotoxins in the venom gland transcriptomes.

expressed toxins, respectively and RNentropy identified 59. All methods identified the two subunits of MTX as being expressed higher in Type A and Type A + B compared to Type B. Additionally, all analyses in which the comparison was possible identified all but one SVMPP as being expressed higher in Type B and Type A + B individuals compared to Type A. EBSeq and RNentropy identified SVMPII-4 as the only locus in which the Type A + B individuals expressed it at a higher level than either the Type A or Type B individuals. SVMPIII-8 was the only locus that Type A + B expressed at an intermediate level compared to Type A and Type B. The remaining differentially expressed loci were predominately identified by RNentropy which

is the least conservative in identifying differential expression. It compares each individual of the treatment to the global distribution and is sensitive to rare transcripts that do not occur in all treatments [64].

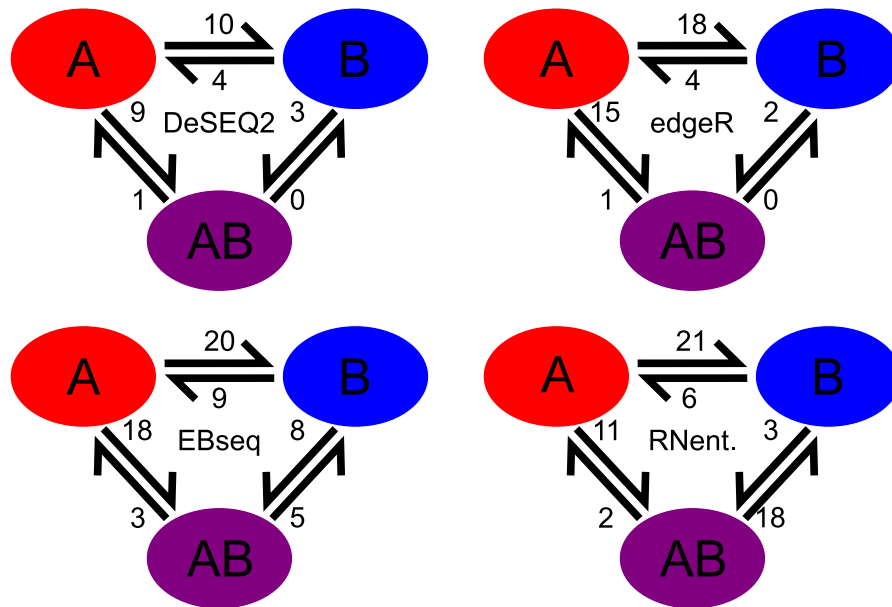


Figure 4.6: Graphical representation of differential expression based on the four methods used: DeSEQ2, edgeR, EBseq, and RNentropy. Arrow direction goes from relatively low expression to relatively high expression for each of the possible pairwise comparison between the three venom types in *C. scutulatus*. Numbers indicate number of toxins out of 84 that are differentially expressed in each comparison.

The clr transformed comparison of the average Type A + B transcriptome to the Average Type A (Figure 4.7) and Average Type B (Figure 4.8) transcriptomes were consistent in identifying the same outlier toxins as identified by the four methods above. In both comparisons, the correlation between the nontoxins was greater than 0.90 for all three measures (Figures 4.7 and 4.8). In the Type A to Type A + B comparison, SVMs and CTL-3 and CTL-4 were overexpressed in Type A + B compared to Type A (Figure 4.7). The only locus overexpressed in Type A was MYO-8 which is the most lowly expressed myotoxin and only found in CLP1936 (Figure 4.4). In comparing Type B to Type A + B, the two subunits of MTX were overexpressed in Type A + B. There were a few other toxins including MYO-1, Pla2-gA1, MYO-2, SVMPII-4, CTL-9, and

CTL-12 that were also overexpressed in Type A + B.

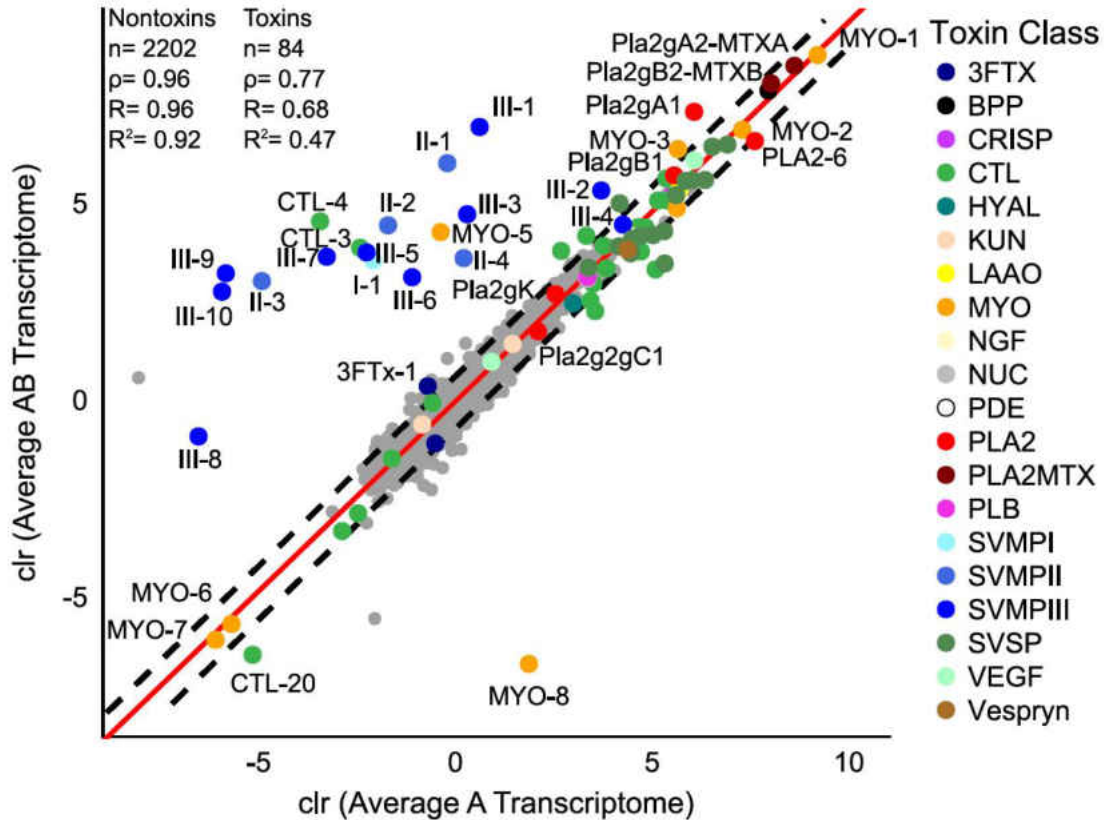


Figure 4.7: Pairwise comparison of average Type A and Type A + B venom gland transcriptomes using the centered log-ratio (clr)-transformed TPM data. The red line is the line of best fit through the non toxins and the dashed black lines are the 99% confidence around that line. Anything above the upper line is overexpressed in Type A + B and anything below the lower line is overexpressed in Type A.

Other toxins and toxin families did not have an expression pattern associated with venom type. Myotoxins were present and absent in all three venom types. The only SVSPs that were identified as differentially expressed in any of the analysis were not consistently found in all individuals and were among the lowly expressed toxins in that family.

Snake venom metalloproteinase and phospholipase A₂s in Type A + B individuals

As discussed above, we recovered SVMP and PLA₂ loci in this study that were not recovered by Strickland *et al.* [33]. SVMP-1 was recovered and expressed in CLP2111 and had

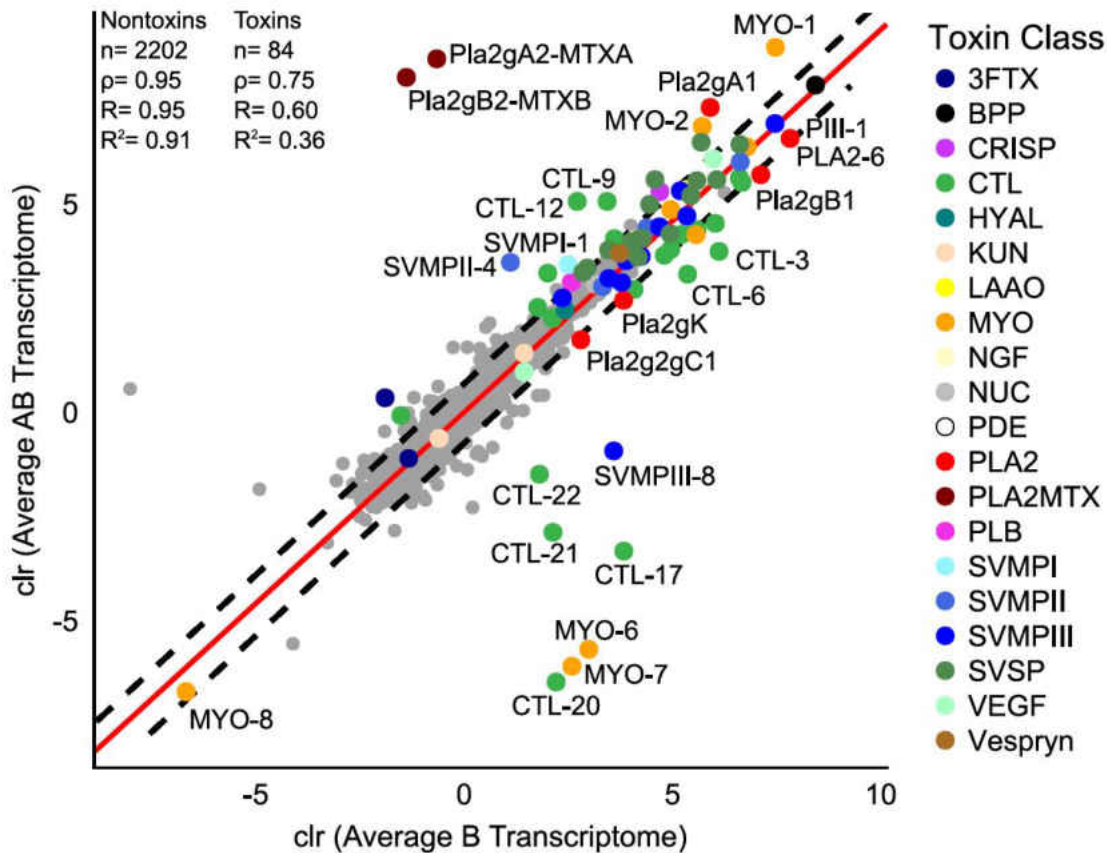


Figure 4.8: Pairwise comparison of average Type B and Type A + B venom gland transcriptomes using the centered log-ratio (clr)-transformed TPM data. The red line is the line of best fit through the non toxins and the dashed black lines are the 99% confidence around that line. Anything above the upper line is overexpressed in Type A + B and anything below the lower line is overexpressed in Type B.

low expression in CLP2136. SVMPIs are not generally recovered in venom-gland transcriptomes because they are very similar to the metalloproteinase domain of SVMPIIs and SVMPIIIs [71]. CLP2111 was the only individual with a high TPM value for SVMPI-1 and even it had poor coverage across the transcript. The three additional SVMPIIs recovered in this study were expressed at a much lower level than SVMPII-1 recovered previously [33]. The PLA₂ added to the consensus transcriptome, Pla2-gC1, was tested for by Strickland *et al.* [33] because it was recovered in *C. scutulatus*' genome by Dowell *et al.* [32]. However, neither study identified it in the transcriptomes. Here, we recovered and annotated Pla2-gC1 in CLP1829. It was also expressed

in CLP1831, CLP1832, and CLP1835 although, at low levels overall.

In testing for all PLA₂s and SVMPs that have been identified in the genome [24, 32] or transcriptome [33] of *C. scutulatus*, we found that expression levels were not different between venom types using the proteomic definition (Table 4.2). Type A + B individuals expressed MTX at the same levels as Type A individuals and expressed SVMPs at the same level as Type B individuals. The presence of the other PLA₂ loci was variable across venom types. Two of the Type A + B individuals had all six PLA₂s present in the transcriptome (Table 4.2). The two MTX subunits were the only PLA₂s not found in Type B individual and all six were found in Type A and Type A + B. For the 20 SVMP loci, there was a much clearer pattern. We identified one additional SVMPII and SVMPIII that Dowell *et al.* [24] did not find in the genome and we included SVMPI. Each of those was only found in one individual. Twelve of the SVMPs were found in all Type B and Type A + B individuals and were also only found in the genome of Type B individuals by Dowell *et al.* [24]. One of these loci, mp277, identified in Dowell *et al.* [24] was found in the genome of Type A individuals but we did not find it in Type A individuals. SVMPIII-4 was the only SVMP expressed in all 15 individuals and was also identified as mp244 in the genome of Type A and Type B individuals. The two SVMP loci (mp245 and mp2442) that were found in the genome by [24] but not in the transcriptome were not found in any individual in this study. Two SVMPs (mp232 and mp233) were found in all but two Type A individuals.

Inheritance pattern of SVMPs and PLA₂s in C. scutulatus

Using the genomic definition of venom type proposed by Dowell *et al.* [32] and Dowell *et al.* [24], we were able to determine whether each individual was homozygous or heterozygous at the PLA₂ and SVMP loci. Of the 15 individuals six were homozygous for the Type A PLA₂ loci (MTXA and MTXB), four were homozygous for the Type B PLA₂ loci (Pla2-gK and Pla2-gB1), and five were heterozygous for the PLA₂ loci and expressed all four (Table 4.2). The heterozygous individuals included three of four Type A + B individuals and two Type A individuals. Using

Table 4.2: Transcripts per million reads (TPM) values rounded to the closest integer for four PLA₂ and four SVMP loci identified as being unique to the genome of Type A and Type B *Crotalus scutulatus* [24, 32]. Values in **bold** met the criteria for presence and values in *italics* did not. Heterozygous individuals for PLA₂s express all four PLA₂s. Heterozygous individuals (CLP1823, CLP2111) at the SVMP loci were putatively identified based on relative expression of the four most highly expressed SVMP loci in the region present in Type B: mp238, mp237, mp240, mp242. The putative genotype is listed with the following code: AA–presence of MTX/absence of Pla2-gK and Pla2-gB1, Aa–expression of all four PLA₂s, aa–absence of MTX/presence of Pla2-gK and Pla2-gB1, BB–high expression of SVMPs, Bb–approximately half of the average expression of SVMPs, bb–no expression of SVMPs.

Dowell <i>et al.</i> [24]		Type A							Type A + B				Type B				
Toxin	Type A	Type B	1829	1929	1930	1936	1959	1961	1972	1823	1830	1832	2111	1831	1835	2136	2142
Pla2-gA2-MTXA	Genome	-	428109	473400	600972	479870	348757	456812	518336	358265	247506	276057	405675	24	210	0	0
Pla2-gB2-MTXB	Genome	-	308910	246797	293026	258001	163398	318099	215095	225628	187681	217362	288133	18	97	0	0
Pla2-gK	-	Genome	705	5	0	2	6050	7	0	3803	889	512	9	6493	9432	3437	3683
Pla2-gB1	-	Genome	56618	19	6	46	100015	20	0	68548	21906	37042	15	180956	149454	161831	148588
mp238/SVMPIII-1	-	Genome	28	21	0	0	0	24	0	73163	191930	175233	88172	289626	230569	232942	187000
mp237/SVMPIII-1	-	Genome	10	7	0	0	0	9	0	40495	84910	56685	32417	125979	112314	85461	126060
mp240/SVMPIII-3	-	Genome	3	4	0	1	2	103	13	7205	21588	18006	10815	34480	28586	23626	30223
mp242/SVMPIII-5	-	Genome	1	1	0	0	0	1	0	4297	9741	7930	1278	6365	8983	12283	9253
Putative Genotype	AAbb	aaBB	Aabb	AAbb	AAbb	AAbb	Aabb	AAbb	AAbb	AaBb	AaBB	AaBB	AABb	aaBB	aaBB	aaBB	aaBB

homozygous Type A, heterozygous, and homozygous Type B as our three treatments, we found that the heterozygous individuals expressed the four PLA₂s at approximately half of the parent type. For the SVMP loci, all seven individuals proteomically identified as Type A were also genomically Type A. They did not express any of the metalloproteinases found in the deletion region identified by Dowell *et al.* [24] and are homozygous Type A (Table 4.2). The remaining eight individuals each expressed the four metalloproteinases expressed in Type B individuals, two of the Type A + B individuals had approximately half the expression level of the remaining individuals and were classified as heterozygous (Table 4.2). The two other Type A + B individual and all four Type B individuals were classified as homozygous Type B. In total, five individuals were homozygous Type A for both PLA₂ and SVMP loci, four individuals were homozygous Type B for both PLA₂ and SVMP loci, one was heterozygous for PLA₂ and SVMP loci and five were heterozygous for one locus and homozygous for the other. These last five have to be at least F2 individuals which indicates F1s can successfully reproduce.

Each genotype identified occupied a unique area in PCA space (Figure 4.9). The Type A individuals clustered tightly on one end of PCA axis 1, the Type B individuals clustered at the other end of PCA 1, and the six individuals that were heterozygous were located in between the

two pure clusters (Figure 4.9). The second PCA axis only explained 10.05% of the variation and the two toxins with the highest loading scores were Pla2-gK and Pla2-B1 which are variable in all individuals except the pure Type A individuals (Figure 4.9).

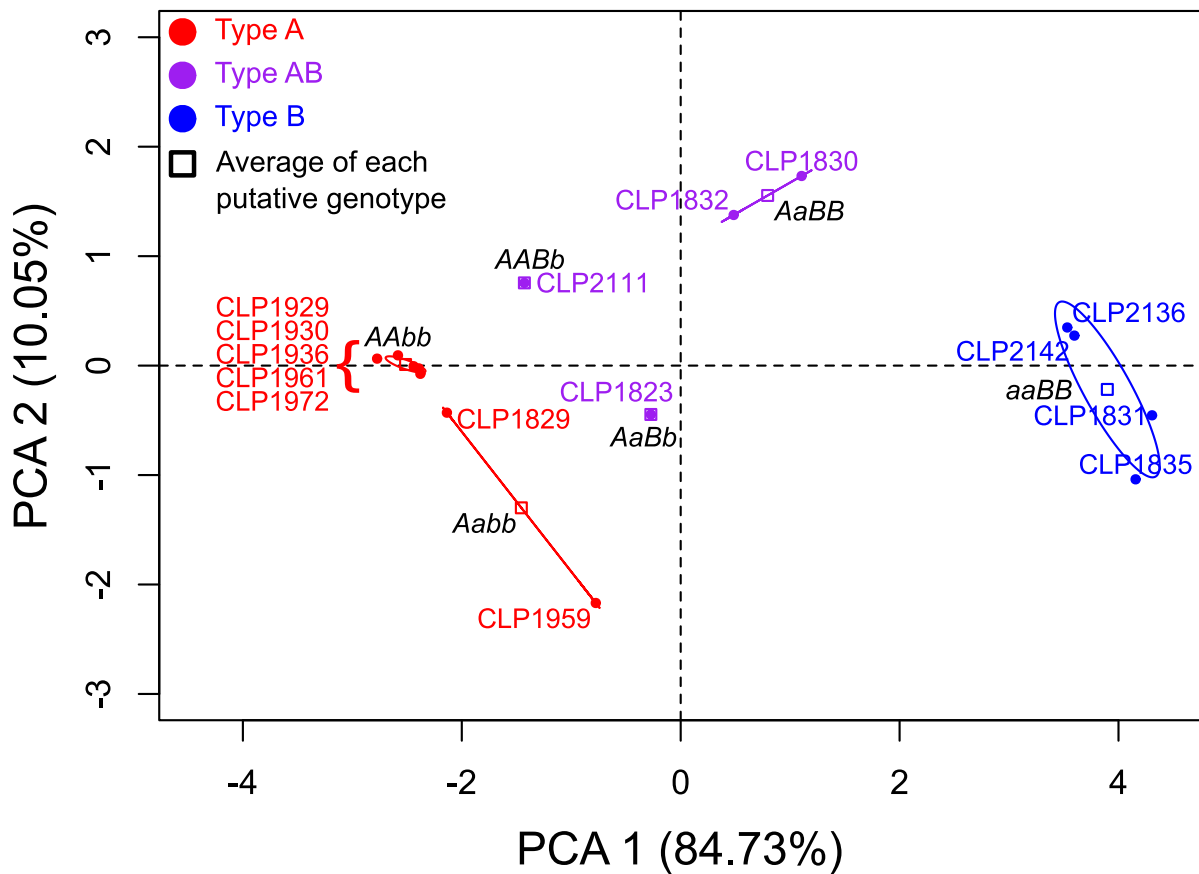


Figure 4.9: Principle component analysis (PCA) scatterplot of the 15 *C. scutulatus* from the Sonoran Desert based on TPM data for the four PLA₂S and four SVMPs in Table 4.2. The squares are the average value for each putative genotype and the circles are the 95% confidence ellipse for genotypes with more than two individuals. The first two axes explain 94.78% of the variability. PCA 1 represents MTX presence/SVMP absence (left) and MTX absence/SVMP presence (right) and PCA 2 is the variation of expression of Pla2-gK and Pla2-gB1.

Discussion

Using comparative transcriptomics of the venom glands of 15 Mojave Rattlesnakes, *C. scutulatus*, we were able to identify additive expression in a highly adapted phenotype. This

suggests that the venom phenotype dichotomy in *C. scutulatus* and rattlesnakes more broadly is Mendelianly inherited and *cis*-regulated. The integrated venom phenotype, Type A + B, is not a unique venom type and can be traced to interbreeding between a Type A individual and a Type B individual as previously hypothesized by Glenn and Straight [44]. We identified six of the nine possible genotypes using the genomic definition of venom type based on Dowell *et al.* [24]. Initially, by defining phenotypes based on the proteomic characteristics, we were unable to identify intermediate expression of the integrated phenotype which suggested heterosis and potentially *trans*-regulation. However, when we controlled for genotype, the eight loci involved that can be detected transcriptomically, had mid-parent levels of expression (Table 4.2 and Figure 4.9). To our knowledge, this is the first time that transcriptomic data of a complex phenotype was used to identify the putative genotype of individuals and demonstrates the power of RNA-seq in linking the phenotype to the genotype.

The two venom phenotypes, Type A and Type B, were initially defined based on proteomic characteristics [19, 27–31]. Using this definition, we were able to identify four *C. scutulatus* with the Type A + B phenotype (Figure 4.2). The Type A + B individuals did cluster in the middle of the Type A/Type B spectrum (Figure 4.5) but, by grouping the individuals by phenotype, we were unable to find intermediate or differential expression outside of the Type A and Type B differences (Figure 4.6). The lack of differentiation between the average Type A + B phenotype compared to the average Type A and average Type B phenotypes is due in part to not controlling for genotype (Figures 4.7 and 4.8). Using the clustering approach, the four Type A + B individuals did not have the same expression pattern (Figure 4.4). By using the proteomic definition, we were not able to detect mid-parent level expression of the Type A + B individuals. However, when we used the genomic definition of the two venom types hypothesized by Dowell *et al.* [24], we were able to detect mid-parent level expression (Table 4.2) and determine that the expression of MTX and SVMPs is additive. Therefore, considering Type A + B as unique, is not helpful. It falls along the spectrum of venom types in rattlesnakes and has a simple genetic mechanism.

We were able to identify six of the nine possible genotypes in a two allele system (Figure 4.4). The three we did not recover were aaBb, AABB, aabb. However, Glenn *et al.* [38] identified a “Type C” venom type in *C. horridus* which is a species that exhibits the same variability in venom as *C. scutulatus* [38]. Type C venoms lack Canebrake toxin (Mojave toxin homolog) and hemorrhagic activity. Using the Mendelian framework, this phenotype is likely one of the missing genotypes, aabb, not recovered in this study. Type C has not been recovered in *C. scutulatus* but, it is also one of the two rarest possible genotypes and likely has the lowest fitness of the possible genotypes because of poor venom efficacy. The aaBb phenotype likely has the second lowest fitness due to the lack of MTX and half of the SVMPs as pure Type B animals. The last genotype not recovered, AABB, has the same probability of occurring in a dihybrid cross of F1 snakes (AaBb X AaBb) as the aabb phenotype. The AABB phenotype would likely produce the most costly venom because the level of expression of MTX and SVMPs. Of the three instances of A + B venom found in rattlesnakes, Mendelian inheritance explains the phenotype of the offspring between *C. viridis* complex and *C. scutulatus* [35, 37] but the ontogenetic change from Type A to Type B in *C. simus* is not possible under this framework.

Because of the ontogenetic change in *C. simus*, one of the hypotheses for the A/B venom dichotomy in rattlesnakes was that the Type A venom was a pedomorphic characteristic retained in the one species that is monomorphic for it (*C. tigris*) and the populations that have Type A venom [15, 31, 72]. *Crotalus simus* is part of the *C. durissus* complex which has representatives of the Type A and Type B phenotypes [73, 74]. Using the genomic definition of the two venom types [24] and considering the Mendelian framework presented here, pedomorphism is not a good explanation for the retention of the Type A venom state. The two subunits of Crotoxin (Mojave toxin homolog) were the only PLA₂s recovered in the transcriptome of *C. simus* as both juveniles and adults. Additionally, no SVMPIIs were recovered in the transcriptome of either the adults and juveniles. All the SVMPIIs are located in the genome in the region that is deleted in Type A individuals. Therefore, Type A venom in *C. tigris* is due to the loss of the Type B SVMP portion

of the genome and *C. simus* should be considered Type A. There is still an ontogenetic shift in the relative proportion of Crotoxin and the small number of SVMPIIIs expressed, but it is not a change from Type A to Type B venom. The fixation of venom types is likely due to selection for specific prey and not the retention of juvenile characteristics as hypothesized.

Based on our results from other toxin classes, it may be possible that other toxins are also inherited in a Mendelian fashion. MYO-1, which is homologous with Crotoxin, showed a similar pattern of presence/absence across the 15 individuals used in this study. Additionally, CTLs may follow this pattern (Figure 4.4). High numbers of CTLs are generally associated with the Type B phenotype [33]. One Type A + B individual, CLP2111 was missing a suite of CTLs (CTL-1,2,5,6,7,11, and 13) that all of the other Type B and Type A + B individuals had. That same suite was present in two Type A individuals (CLP1959 and CLP1961) and absent in the other Type A individuals. CTLs, Myotoxins, PLA₂s, and SVMPIs are the four toxin families that vary the most in *C. scutulatus* [33] and myotoxins were also highly variable proteomically [41] and explained almost as much of the variation in venom as the A/B dichotomy (Figure 4.5). Mendelian inheritance of these loci which causes additive expression through *cis*-regulation would be a relatively simple way for fine scale local adaptation to occur.

Conclusions

We present evidence that supports Mendelian inheritance of the toxins responsible for the venom type dichotomy in rattlesnakes. The additive expression in which the heterozygous individuals expressed the toxins involved at approximately half of the parents indicating *cis*-regulation is what determines expression for these loci. Using this framework, the four phenotypes that have been described in rattlesnakes can be linked to nine possible genotypes in a dihybrid cross. The possible ways to get the integrated venom type is limited to interbreeding between a Type A individual and a Type B individual and precludes the ontogenetic shift in *C. simus*

as being “Type A + B”. Type A individuals and species should be defined based on the presence of both subunits of Mojave toxin and the absence of Pla2-gA1, Pla2-gK, and almost all SVMP loci. Type B individuals should be identified based on the presence of Pla2-gA1, Pla2-gK, and mp237, mp238, mp240, and mp242 and the lack of MTX. Because of the presence of at least one SVMPIII in Type A individuals, the proteomic definition will not always be accurate in predicting the genotype of an individual.

Using the Mendelian inheritance framework, tests of relative fitness between the genotypes will be more informative when testing if and how balancing selection is maintaining the venom polymorphism in *C. scutulatus* and rattlesnakes more broadly. The development of PCR based assay for the other PLA₂ loci and the SVMP loci will make it possible to estimate the allele frequency across populations to test for frequency dependent selection as the mechanism of balancing selection. Additionally, by using venom from animals representing the nine possible genotypes, relative efficacy could be easily assessed in prey items testing fine scale local adaptation or heterozygote advantage. This could be used as a proxy for fitness of each genotype. We hypothesize that the Mendelian inheritance pattern may extend to at least two other toxin classes which would allow for the creation of more genotypes and test how different combinations of genes change the composition of a highly adaptive phenotype.

Acknowledgments

Support and assistance in the field was provided by B. Ashley, T. Burkhardt, H. Dahn, D. Deem, M. Feldner, T. Fisher, R. Govreau, J. Houck, T. Jones, A. Mason, C. May, S. May, R. Mayerhofer, E. McCormick, J. McNally, D. Ortiz, A. Owens, A. Quillen, R. Rautsaw, J. Slone, D. Speckin, G. Territo, C. Vratil, D. Weber, W. Wüster, and G. Zancolli. We thank M. Hogan, M. Margres, J. McGivern, and M. Ward from the Rokyta Lab, M. Seavy and B. Washburn (FSU Department of Biological Science Analytical Lab), and S. Miller (FSU DNA Sequencing Facility),

for assistance with preparing the samples and generating data. We thank D. Greenspan, N. Lucas, J. Schnaitter, and P. Wiegand (UCF ARCC) for assistance with Stokes HPC and P. Larabee (UCF IT) for assistance with operating system compatibility. We thank T. Jones and C. Kondrat-Smith with Arizona Game and Fish Department and S. Ferguson with the New Mexico Department of Game and Fish with assistance in obtaining collecting permits. We thank M. Giorgianni and the Carroll lab for providing coding sequences for the *Crotalus scutulatus* snake venom metalloproteinases from Dowell et al. (2018). We thank C. Johnston at the Arizona State University Natural History Museum and M. Revelez at the Angelo State Natural History Collection for assistance with depositing specimens. Comments and suggestions were given by E. Hofmann, A. Mason, and R. Rautsaw which greatly improved the quality of this manuscript. Financial support was generously provided by the National Science Foundation to C.L.P (DUE 1161228 which funded J.L.S. and DEB 1638879) and D.R.R. (DEB 1145987 and DEB 1638902) Additional support was given to J.L.S. by Prairie Biotic Research Inc., Sigma Xi Grants-in-aid-of-research, SnakeDays Research Grant, the Southwestern Association of Naturalists McCarley Research Grant, and the Theodore Roosevelt Memorial Fund through the American Museum of Natural History.

References

1. Harrison, P.W.; Wright, A.E.; Mank, J.E. The evolution of gene expression and the transcriptome-phenotype relationship. *Seminars in Cell & Developmental Biology* **2012**, *23*, 222–229.
2. Li, L.; Petsch, K.; Shimizu, R.; Liu, S.; Xu, W.W.; Ying, K.; Yu, J.; Scanlon, M.J.; Schnable, P.S.; Timmermans, M.C.P.; Springer, N.M.; Muehlbauer, G.J. Mendelian and non-Mendelian regulation of gene expression in maize. *PLoS Genetics* **2013**, *9*, e1003202.
3. Cox, C.L.; Davis Rabosky, A.R. Spatial and temporal drivers of phenotypic diversity in polymorphic snakes. *The American Naturalist* **2013**, *182*, E40–E57.

4. Pascoal, S.; Liu, X.; Ly, T.; Fang, Y.; Rockliffe, N.; Paterson, S.; Shirran, S.L.; Botting, C.H.; Bailey, N.W. Rapid evolution and gene expression: a rapidly evolving Mendelian trait that silences field crickets has widespread effects on mRNA and protein expression. *Journal of Evolutionary Biology* **2016**, *29*, 1234–1246.
5. Swanson-Wagner, R.A.; Jia, Y.; DeCook, R.; Borsuk, L.A.; Nettleton, D.; Schnable, P.S. All possible modes of gene action are observed in a global comparison of gene expression in a maize F1 hybrid and its inbred parents. *Proceedings of the National Academy of Sciences* **2006**, *103*, 6805–10.
6. Birchler, J.A.; Auger, D.L.; Riddle, N.C. In search of the molecular basis of heterosis. *The Plant Cell* **2003**, *15*, 2236–9.
7. Wolf, J.B.; Bayer, T.; Haubold, B.; Schilhabel, M.; Rosenstiel, P.; Tautz, D. Nucleotide divergence vs. gene expression differentiation: comparative transcriptome sequencing in natural isolates from the carrion crow and its hybrid zone with the hooded crow. *Molecular Ecology* **2010**, *19*, 162–175.
8. Poelstra, J.W.; Vijay, N.; Bossu, C.M.; Lantz, H.; Ryll, B.; Müller, I.; Baglione, V.; Unneberg, P.; Wikelski, M.; Grabherr, M.G.; Wolf, J.B.W. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **2014**, *344*, 1410–4.
9. Barton, N.H.; Hewitt, G.M. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **1985**, *16*, 113–148.
10. Thiemann, A.; Fu, J.; Schrag, T.A.; Melchinger, A.E.; Frisch, M.; Scholten, S. Correlation between parental transcriptome and field data for the characterization of heterosis in *Zea mays* L. *Theoretical and Applied Genetics* **2010**, *120*, 401–413.
11. Stupar, R.M. *Cis*-transcriptional variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. *Genetics* **2006**, *173*, 2199–2210.
12. Springer, N.M.; Stupar, R.M. Allelic variation and heterosis in maize: how do two halves

- make more than a whole? *Genome Research* **2007**, *17*, 264–75.
13. Lozier, J.D.; Jackson, J.M.; Dillon, M.E.; Strange, J.P. Population genomics of divergence among extreme and intermediate color forms in a polymorphic insect. *Ecology and Evolution* **2016**, *6*, 1075–1091.
 14. Weinstein, S.A.; Minton, S.A.; Wilde, C.E. The distribution among ophidian venoms of a toxin isolated from the venom of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*). *Toxicon* **1985**, *23*, 825–844.
 15. Calvete, J.J.; Pérez, A.; Lomonte, B.; Sánchez, E.E.; Sanz, L. Snake venomomics of *Crotalus tigris*: the minimalist toxin arsenal of the deadliest nearctic rattlesnake venom. Evolutionary clues for generating a pan-specific antivenom against crotalid type II venoms. *Journal of Proteome Research* **2012**, *11*, 1382–1390.
 16. Bieber, A.L.; Tu, T.; Tu, A.T. Studies of an acidic cardiotoxin isolated from the venom of Mojave Rattlesnake (*Crotalus scutulatus*). *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1975**, *400*, 178–188.
 17. Cate, R.L.; Bieber, A.L. Purification and characterization of Mojave (*Crotalus scutulatus scutulatus*) toxin and its subunits. *Archives of Biochemistry and Biophysics* **1978**, *189*, 397–408.
 18. Gopalakrishnakone, P.; Hawgood, B.J.; Holbrooke, S.E.; Marsh, N.A.; Santana De Sa, S.; Tu, A.T. Sites of action of Mojave toxin isolated from the venom of the Mojave Rattlesnake. *British Journal of Pharmacology* **1980**, *69*, 421–31.
 19. Doley, R.; Kini, R.M. Protein complexes in snake venom. *Cellular and Molecular Life Sciences* **2009**, *66*, 2851–2871.
 20. Martinez, M.; Rael, E.D.; Maddux, N.L. Isolation of a hemorrhagic toxin from Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) venom. *Toxicon* **1990**, *28*, 685–694.
 21. Wooldridge, B.; Pineda, G.; Banuelas-Ornelas, J.; Dagda, R.; Gasanov, S.; Rael, E.; Lieb, C. Mojave Rattlesnakes (*Crotalus scutulatus scutulatus*) lacking the acidic subunit DNA

- sequence lack Mojave toxin in their venom. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **2001**, *130*, 169–179.
22. Sánchez, E.E.; Galán, J.A.; Powell, R.L.; Reyes, S.R.; Soto, J.G.; Russell, W.K.; Russell, D.H.; Pérez, J.C. Disintegrin, hemorrhagic, and proteolytic activities of Mohave Rattlesnake, *Crotalus scutulatus scutulatus* venoms lacking Mojave toxin. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* **2005**, *141*, 124–132.
 23. Borja, M.; Castañeda, G.; Espinosa, J.; Neri, E.; Carbajal, A.; Clement, H.; García, O.; Alagon, A. Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) with Type B venom from Mexico. *Copeia* **2014**, *2014*, 7–13.
 24. Dowell, N.L.; Giorgianni, M.W.; Griffin, S.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. Extremely divergent haplotypes in two toxin gene complexes encode alternative venom types within rattlesnake species. *Current Biology* **2018**.
 25. Strickland, J.L.; Smith, C.F.; Mason, A.J.; Borja, M.; Schield, D.R.; Castoe, T.A.; Spencer, C.L.; Rokyta, D.R.; Parkinson, C.L. Phylogeography and environmental variables poorly explain diverse venom phenotypes in Mojave Rattlesnakes (*Crotalus scutulatus*). *Dissertation* **2018**.
 26. Casewell, N.R.; Wüster, W.; Vonk, F.J.; Harrison, R.A.; Fry, B.G. Complex cocktails: the evolutionary novelty of venoms. *Trends in Ecology & Evolution* **2013**, *28*, 219–29.
 27. Glenn, J.; Straight, R. Mojave rattlesnake *Crotalus scutulatus scutulatus* venom: variation in toxicity with geographical origin. *Toxicon* **1978**, *16*, 81–84.
 28. Glenn, J.L.; Straight, R.C. The rattlesnakes and their venom yield and lethal toxicity. In *Rattlesnake Venoms: Their actions and treatment.*; Tu, A.T., Ed.; Marcel Dekker, Inc.: New York, 1982; pp. 3–119.
 29. Mackessy, S.P. Venom composition in rattlesnakes: trends and biological significance. In *The Biology of Rattlesnakes*; Hayes, W.K.; Beaman, K.R.; Cardwell, M.D.; Bush, S.P., Eds.; Loma Linda University Press: Loma Linda, CA, 2008; pp. 495–510.

30. Doley, R.; Zhou, X.; Kini, R. Snake venom phospholipase A2 enzymes. In *Handbook of Venoms and Toxins of Reptiles*; Mackessy, S.P., Ed.; CRC Press: Boca Raton, Florida, 2009; pp. 173–205.
31. Mackessy, S.P. Evolutionary trends in venom composition in the Western Rattlesnakes (*Crotalus viridis sensu lato*): toxicity vs. tenderizers. *Toxicon* **2010**, *55*, 1463–74.
32. Dowell, N.L.; Giorgianni, M.W.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. The deep origin and recent loss of venom toxin genes in rattlesnakes. *Current Biology* **2016**, *26*, 2434–2445.
33. Strickland, J.; Mason, A.; Rokyta, D.; Parkinson, C. Phenotypic variation in Mojave Rattlesnake (*Crotalus scutulatus*) venom is driven by four toxin families. *Toxins* **2018**, *10*, 135.
34. Calvete, J.J.; Sanz, L.; Cid, P.; de la Torre, P.; Flores-Diaz, M.; Dos Santos, M.C.; Borges, A.; Bremo, A.; Angulo, Y.; Lomonte, B.; Alape-Giron, A.; Gutierrez, J.M. Snake venomomics of the Central American Rattlesnake *Crotalus simus* and the South American *Crotalus durissus* complex points to neurotoxicity as an adaptive pedomorphic trend along *Crotalus* dispersal in South America. *Journal of Proteome Research* **2010**, *9*, 528–544.
35. Smith, C.F.; Mackessy, S.P. The effects of hybridization on divergent venom phenotypes: characterization of venom from *Crotalus scutulatus scutulatus* x *Crotalus oreganus helleri* hybrids. *Toxicon* **2016**, *120*, 110–123.
36. Glenn, J.L.; Straight, R.C. Venom characteristics as an indicator of hybridization between *Crotalus viridis viridis* and *Crotalus scutulatus scutulatus* in New Mexico. *Toxicon* **1990**, *28*, 857–862.
37. Zancolli, G.; Baker, T.; Barlow, A.; Bradley, R.; Calvete, J.; Carter, K.; de Jager, K.; Owens, J.; Price, J.; Sanz, L.; Scholes-Higham, A.; Shier, L.; Wood, L.; Wüster, C.; Wüster, W. Is hybridization a source of adaptive venom variation in rattlesnakes? A test, using a *Crotalus scutulatus* x *viridis* hybrid zone in southwestern New Mexico. *Toxins*

- 2016**, 8, 188.
38. Glenn, J.; Straight, R.; Wolt, T. Regional variation in the presence of canebrake toxin in *Crotalus horridus* venom. *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology* **1994**, 107, 337–346.
 39. Rokyta, D.R.; Wray, K.P.; Margres, M.J. The genesis of an exceptionally lethal venom in the Timber Rattlesnake (*Crotalus horridus*) revealed through comparative venom-gland transcriptomics. *BMC Genomics* **2013**, 14, 394.
 40. Rokyta, D.R.; Wray, K.P.; McGivern, J.J.; Margres, M.J. The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the Timber Rattlesnake (*Crotalus horridus*). *Toxicon* **2015**, 98, 34–48.
 41. Massey, D.J.; Calvete, J.J.; Sánchez, E.E.; Sanz, L.; Richards, K.; Curtis, R.; Boesen, K. Venom variability and envenoming severity outcomes of the *Crotalus scutulatus scutulatus* (Mojave Rattlesnake) from southern Arizona. *Journal of Proteomics* **2012**, 75, 2576–2587.
 42. Wilkinson, J.A.; Glenn, J.L.; Straight, R.C.; Sites, J.W. Distribution and genetic variation in venom A and B populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Herpetologica* **1991**, 47, 54 – 68.
 43. Borja, M.; Neri-Castro, E.; Castañeda-Gaytán, G.; Strickland, J.L.; Parkinson, C.L.; Castañeda-Gaytán, J.; Ponce-López, R.; Lomonte, B.; Olvera-Rodríguez, A.; Alagón, A.; Pérez-Morales, R.; Olvera-Rodríguez, A.; Alagón, A.; Pérez-Morales, R. Biological and proteolytic variation in the venom of *Crotalus scutulatus scutulatus* from Mexico. *Toxins* **2017**, 10, 35.
 44. Glenn, J.L.; Straight, R.C. Intergradation of two different venom populations of the Mojave Rattlesnake (*Crotalus scutulatus scutulatus*) in Arizona. *Toxicon* **1989**, 27, 411–8.
 45. Schield, D.R.; Card, D.C.; Adams, R.H.; Corbin, A.; Jezkova, T.; Hales, N.; Meik, J.M.; Spencer, C.L.; Smith, L.; Campillo-Garcia, G.; Bouzid, N.; Strickland, J.L.; Parkinson, C.L.; Flores-Villela, O.; Mackessy, S.P.; Castoe, T.A. Cryptic genetic diversity, population

- structure, and gene flow in the Mojave Rattlesnake (*Crotalus scutulatus*). *Molecular Phylogenetics and Evolution* **2018 In Review**.
46. Margres, M.J.; Wray, K.P.; Hassinger, A.T.B.; Ward, M.J.; McGivern, J.J.; Moriarty Lemmon, E.; Lemmon, A.R.; Rokyta, D.R. Quantity, not quality: rapid adaptation in a polygenic trait proceeded exclusively through expression differentiation. *Molecular Biology and Evolution* **2017**, *34*, 3099–3110.
 47. Rotenberg, D.; Bamberger, E.S.; Kochva, E. Studies on ribonucleic acid synthesis in the venom glands of *Vipera palaestinae* (Ophidia, Reptilia). *The Biochemical Journal* **1971**, *121*, 609–12.
 48. Rokyta, D.R.; Wray, K.P.; Lemmon, A.R.; Lemmon, E.M.; Caudle, S.B. A high-throughput venom-gland transcriptome for the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*) and evidence for pervasive positive selection across toxin classes. *Toxicon* **2011**, *57*, 657–671.
 49. Rokyta, D.R.; Margres, M.J.; Ward, M.J.; Sanchez, E.E. The genetics of venom ontogeny in the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*). *PeerJ* **2017**, *5*, e3249.
 50. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **2014**, *30*, 614–620.
 51. Rokyta, D.R.; Lemmon, A.R.; Margres, M.J.; Aronow, K. The venom-gland transcriptome of the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*). *BMC Genomics* **2012**, *13*, 312.
 52. Li, W.; Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659.
 53. Margres, M.J.; Aronow, K.; Loyacano, J.; Rokyta, D.R. The venom-gland transcriptome of the Eastern Coral Snake (*Micrurus fulvius*) reveals high venom complexity in the intragenomic evolution of venoms. *BMC Genomics* **2013**, *14*, 531.
 54. Bendtsen, J.D.; Nielsen, H.; Van Heijne, G.; Brunak, S. Improved prediction of signal

- peptides:SignalP 3.0. *Journal of Molecular Biology* **2004**, *340*, 783–795.
55. Petersen, T.N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **2011**, *8*, 785–786.
 56. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **2012**, *9*, 357–359.
 57. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **1994**, *22*, 4673–4680.
 58. Langmead, B.; Trapnell, C.; Pop, M.; Salzberg, S.L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **2009**, *10*, R25.
 59. Schrider, D.R.; Gout, J.F.; Hahn, M.W. Very few RNA and DNA sequence differences in the human transcriptome. *PLoS ONE* **2011**, *6*, e25842.
 60. Rokyta, D.R.; Margres, M.J.; Calvin, K. Post-transcriptional mechanisms contribute little to phenotypic variation in snake venoms. *G3* **2015**, *5*, 2375–2382.
 61. Palarea-Albaladejo, J.; Martín-Fernández, J.A. Software Description zCompositions – R package for multivariate imputation of left-censored data under a compositional approach. *Chemometrics and Intelligent Laboratory Systems* **2015**, *143*, 85–96.
 62. Kolde, R. Pretty Heatmaps. R package version 1.0.8., 2015.
 63. Lê, S.; Josse, J.; Husson, F. FactoMineR: a package for multivariate analysis. *Journal of Statistical Software* **2008**, *25*, 1–18.
 64. Zambelli, F.; Mastropasqua, F.; Picardi, E.; D’Erchia, A.M.; Pesole, G.; Pavesi, G. RNentropy: an entropy-based tool for the detection of significant variation of gene expression across multiple RNA-Seq experiments. *Nucleic Acids Research* **2018**, p. gky055.
 65. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **2014**, *15*, 550.

66. Robinson, M.D.; Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* **2010**, *11*, R25, [PMC2864565].
67. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140.
68. Leng, N.; Dawson, J.A.; Thomson, J.A.; Ruotti, V.; Rissman, A.I.; Smits, B.M.G.; Haag, J.D.; Gould, M.N.; Stewart, R.M.; Kendziorski, C. EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **2013**, *29*, 1035–1043.
69. Aitchison, J. *The statistical analysis of compositional data*; Vol. 44, Chapman and Hall: London, 1986; pp. 139–177.
70. Rádis-Baptista, G.; Oguiura, N.; Hayashi, M.A.; Camargo, M.E.; Grego, K.F.; Oliveira, E.B.; Yamane, T. Nucleotide sequence of crotoxin isoform precursors from a single South American Rattlesnake (*Crotalus durissus terrificus*). *Toxicon* **1999**, *37*, 973–984.
71. Takeda, S.; Takeya, H.; Iwanaga, S. Snake venom metalloproteinases: structure, function and relevance to the mammalian ADAM/ADAMTS family proteins. *Biochimica et Biophysica Acta - Proteins and Proteomics* **2012**, *1824*, 164–176.
72. Mackessy, S.P.; Williams, K.; Ashton, K.G. Ontogenetic variation in venom composition and diet of *Crotalus oreganus concolor*: a case of venom paedomorphosis? *Copeia* **2003**, *2003*, 769–782.
73. Durban, J.; Pérez, A.; Sanz, L.; Gómez, A.; Bonilla, F.; Rodríguez, S.; Chacón, D.; Sasa, M.; Angulo, Y.; Gutiérrez, J.M.; Calvete, J.J. Integrated ” omics ” profiling indicates that miRNAs are modulators of the ontogenetic venom composition shift in the Central American Rattlesnake, *Crotalus simus simus*. *BMC Genomics* **2013**, *14*, 234.
74. Durban, J.; Sanz, L.; Trevisan-Silva, D.; Neri-Castro, E.; Alagón, A.; Calvete, J.J. Integrated venomomics and venom gland transcriptome analysis of juvenile and adult Mexican Rattlesnakes *Crotalus simus*, *C. tzabcan*, and *C. culminatus* revealed miRNA-modulated

ontogenetic shifts. *Journal of Proteome Research* **2017**, *16*, 3370–3390.

CHAPTER 5: CONCLUSION

Mutation, selection, drift, and gene flow can generate a mosaic of phenotypic diversity across species distributions due to complex interactions between phenotypes and environment [1, 2]. Phenotypic polymorphisms within populations provide an opportunity to test the relative contributions of these four evolutionary mechanisms and examine how diversity is maintained [3–5]. Mutation is the origin of genetic variation, but the interaction between selection, drift, and gene flow determines the maintenance of polymorphisms [6, 7]. Balancing selection is the primary way in which phenotypic polymorphism can be maintained in populations. Balancing selection is rare relative to directional, stabilizing, or disruptive selection, but several cases, such as major compatibility complex, self incompatibility in plants, and butterfly mimicry, have been well documented [8–11].

The maintenance of phenotypic polymorphisms through balancing selection has several proposed mechanisms: heterozygote advantage, negative frequency-dependent selection, sexual antagonism, density-dependent selection, local adaptation to heterogeneous niches over space and/or time, and trait linkage based on the genetic architecture of the polymorphism [12, 13]. These mechanisms are not mutually exclusive and, overall, balancing selection likely includes several of these mechanisms [12]. Balancing selection is difficult to identify, but the geographic arrangement of polymorphisms in a population can provide evidence for it [14]. However, phylogeographic structure of the polymorphism must be ruled out first because local adaptation in populations where gene flow is absent can result in the same geographic pattern as balancing selection [14].

In this study, I used the venom phenotype dichotomy in Mojave Rattlesnakes, *Crotalus scutulatus*, to determine the distribution and evolution of the neurotoxic (Type A) and hemorrhagic (Type B) phenotypes (Figure 5.1). Using morphological, proteomic, transcriptomic, and genomic data, I inferred that balancing selection is likely responsible for the maintenance of the phenotype. Using DNA and/or venom from 216 individuals, I first ruled out phylogeographic structure as the

mechanism maintaining the two phenotypes in *C. scutulatus* [15]. There are at least three distinct genetic lineages within *C. scutulatus* [16] but both venom types are found in all three [17]. By ruling out phylogeographic structure, I was able to test predictions of the different mechanisms of balancing selection.

I did not find any evidence for sexual antagonism as sex did not explain any of the variability in the distribution of venom type or the expression of venom components. Additionally, the genetic architecture of the components responsible for the two venom phenotypes are not linked. There are distinct suites of PLA₂s and SVMPs for each venom type but they can be inherited in any combination and therefore not physically linked to each [18]. However, it may still be possible for these loci to be in linkage disequilibrium or that selection could be acting on both loci simultaneously. To test this would require knowing the allele frequency of the different loci and the relative efficacy of the different genotypes involved. This would also make it possible to determine if negative frequency dependent selection is occurring based on the relative frequency of the two genotypes. There is not enough data on the prey items of *C. scutulatus* so density dependent selection for snakes or different prey item was not testable in this study.

I did find evidence for local adaptation to heterogeneous niches over space and/or time that is likely mediated by fine scale adaptation to prey [17]. I determined that the ecological niches between the two venom types were not equivalent but they were similar. One climatic variable, minimum temperature of the coldest month, was the most important component of both niche models and was also significantly different between Type A and Type B venom phenotypes [17]. The Type B individuals are located in regions that are warmer for longer in the year which could mean that their active season is longer. At the beginning and end of the active season Type B venom may be advantageous for the prey being consumed. Additionally, we found functional morphological differences in the distance between the fangs which may be co-adapted with the Type A phenotype [17].

Overall, venom variability in *C. scutulatus* was higher than expected. Toxins from four

toxin families, PLA₂s, SVMPs, CTLs, and myotoxins-*a*, varied in presence/absence and as a percentage of total expression [15]. Because of the variation found even on small spatial scales, it is important to include as many individuals as possible to represent the transcriptomic diversity of highly adaptive phenotypes such as venom. Although data were only presented from 15 individuals in this study, data from an additional 12 indicate the pattern of high variability in expression patterns occurs in the remainder of *C. scutulatus* distribution (Figure 5.1). I was able to test several hypotheses regarding the venom phenotype dichotomy in *C. scutulatus*. I ruled out pedomorphism as being responsible for the venom type dichotomy as the ontogenetic change seen in *C. simus* does not meet the genomic definition of the venom types [18, 19]. Many of the new hypothesis I presented require fine-scale ecological data on the activity and diet of these snakes or testing relative efficacy on potential prey items. By collecting ecological data, it may be possible to determine the proximal cause of selection on venom phenotypes in *C. scutulatus* and other rattlesnake species.

Significance

This dissertation utilized the largest number of venom-gland transcriptomes from a single species to date. The venom phenotype dichotomy in *C. scutulatus* is likely being maintained by balancing selection and the primary toxin components responsible for the dichotomy are Mendelianly inherited. This framework extends to other species of rattlesnakes where the venom dichotomy exist within and among species [20]. Balancing selection can lead to a trans-species polymorphism which matches the pattern found within *Crotalus* and *Sistrurus* for the presence/absence of the two venom types. I support using the genomic definition of the two venom phenotypes [19, 21] because it provides for a discrete test of the efficacy of the venom which can be used to test relative fitness of the genotypes. Further utilization of *C. scutulatus* as a model for understanding the evolutionary and ecological mechanisms for phenotypic diversity will provide unique insights on how a highly adaptive trait influence fitness and the molecular

mechanisms responsible.

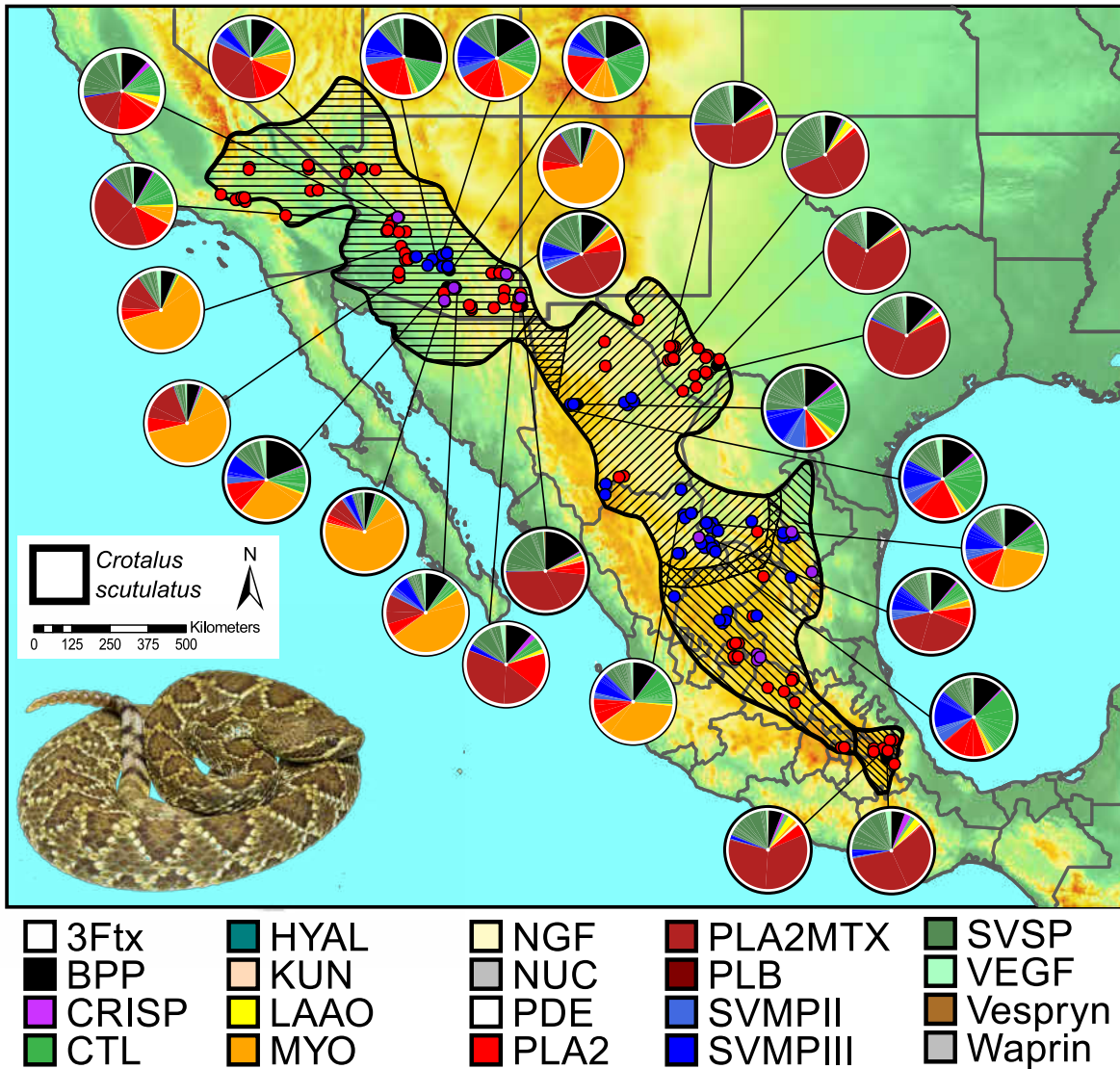


Figure 5.1: Distribution map of Mojave Rattlesnakes, *Crotalus scutulatus*, with venom type labeled with dots and pie charts used to represent the proportion of toxin family in the venom-gland transcriptome of 27 individuals.

References

1. Kawecki, T.J.; Ebert, D. Conceptual issues in local adaptation. *Ecology Letters* **2004**, *7*, 1225–1241.
2. Forester, B.R.; Jones, M.R.; Joost, S.; Landguth, E.L.; Lasky, J.R. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology* **2016**, *25*, 104–120.
3. Hoffman, E.A.; Blouin, M.S. A review of colour and pattern polymorphisms in anurans. *Biological Journal of the Linnean Society* **2000**, *70*, 633–665.
4. Cox, C.L.; Davis Rabosky, A.R. Spatial and temporal drivers of phenotypic diversity in polymorphic snakes. *The American Naturalist* **2013**, *182*, E40–E57.
5. McLean, C.A.; Stuart-Fox, D. Geographic variation in animal colour polymorphisms and its role in speciation. *Biological Reviews* **2014**, *89*, 860–873.
6. Moody, K.N.; Hunter, S.N.; Childress, M.J.; Blob, R.W.; Schoenfuss, H.L.; Blum, M.J.; Ptacek, M.B. Local adaptation despite high gene flow in the waterfall-climbing Hawaiian goby, *Sicyopterus stimpsoni*. *Molecular Ecology* **2015**, *24*, 545–563.
7. Butlin, R.K.; Saura, M.; Charrier, G.; Jackson, B.; André, C.; Caballero, A.; Coyne, J.A.; Galindo, J.; Grahame, J.W.; Hollander, J.; Kemppainen, P.; Martínez-Fernández, M.; Panova, M.; Quesada, H.; Johannesson, K.; Rolán-Alvarez, E. Parallel evolution of local adaptation and reproductive isolation in the face of gene flow. *Evolution* **2014**, *68*, 935–49.
8. Li, J.; Cocker, J.M.; Wright, J.; Webster, M.A.; McMullan, M.; Dyer, S.; Swarbreck, D.; Caccamo, M.; van Oosterhout, C.; Gilmartin, P.M. Genetic architecture and evolution of the S locus supergene in *Primula vulgaris*. *Nature Plants* **2016**, *2*, 16188.
9. Joron, M.; Frezal, L.; Jones, R.T.; Chamberlain, N.L.; Lee, S.F.; Haag, C.R.; Whibley, A.; Becuwe, M.; Baxter, S.W.; Ferguson, L.; Wilkinson, P.A.; Salazar, C.; Davidson, C.; Clark, R.; Quail, M.A.; Beasley, H.; Glithero, R.; Lloyd, C.; Sims, S.; Jones, M.C.;

- Rogers, J.; Jiggins, C.D.; Ffrench-Constant, R.H. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **2011**, *477*.
10. Klein, J. Origin of major histocompatibility complex polymorphism: the trans-species hypothesis. *Human Immunology* **1987**, *19*, 155–162.
 11. Klein, J.; Sato, A.; Nagl, S.; O’huigín, C. Molecular trans-species polymorphism. *Annual Review of Ecology and Systematics* **1998**, *29*, 1–21.
 12. Fijarczyk, A.; Babik, W. Detecting balancing selection in genomes: limits and prospects. *Molecular Ecology* **2015**, *24*, 3529–3545.
 13. Llaurens, V.; Whibley, A.; Joron, M. Genetic architecture and balancing selection: the life and death of differentiated variants. *Molecular Ecology* **2017**, *26*, 2430–2448.
 14. Holmes, I.A.; Grundler, M.R.; Davis Rabosky, A.R. Predator perspective drives geographic variation in frequency-dependent polymorphism. *The American Naturalist* **2017**, *190*, E78–E93.
 15. Strickland, J.; Mason, A.; Rokyta, D.; Parkinson, C. Phenotypic variation in Mojave Rattlesnake (*Crotalus scutulatus*) venom is driven by four toxin families. *Toxins* **2018**, *10*, 135.
 16. Schield, D.R.; Card, D.C.; Adams, R.H.; Corbin, A.; Jezkova, T.; Hales, N.; Meik, J.M.; Spencer, C.L.; Smith, L.; Campillo-Garcia, G.; Bouzid, N.; Strickland, J.L.; Parkinson, C.L.; Flores-Villela, O.; Mackessy, S.P.; Castoe, T.A. Cryptic genetic diversity, population structure, and gene flow in the Mojave Rattlesnakes (*Crotalus scutulatus*). *Molecular Phylogenetics and Evolution* **2018 In Review**.
 17. Strickland, J.L.; Smith, C.F.; Mason, A.J.; Borja, M.; Schield, D.R.; Castoe, T.A.; Spencer, C.L.; Rokyta, D.R.; Parkinson, C.L. Phylogeography and environmental variables poorly explain diverse venom phenotypes in Mojave Rattlesnakes (*Crotalus scutulatus*). *Dissertation* **2018**.
 18. Strickland, J.L. Additive expression suggests Mendelian inheritance of polymorphic

- venom phenotypes in Mojave Rattlesnakes (*Crotalus scutulatus*). *Dissertation* **2018**.
19. Dowell, N.L.; Giorgianni, M.W.; Griffin, S.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. Extremely divergent haplotypes in two toxin gene complexes encode alternative venom types within rattlesnake species. *Current Biology* **2018**, *28*, 1016–1026.
 20. Calvete, J.J.; Pérez, A.; Lomonte, B.; Sánchez, E.E.; Sanz, L. Snake venomics of *Crotalus tigris*: the minimalist toxin arsenal of the deadliest nearctic rattlesnake venom. Evolutionary clues for generating a pan-specific antivenom against crotalid type II venoms. *Journal of Proteome Research* **2012**, *11*, 1382–1390.
 21. Dowell, N.L.; Giorgianni, M.W.; Kassner, V.A.; Selegue, J.E.; Sanchez, E.E.; Carroll, S.B. The deep origin and recent loss of venom toxin genes in rattlesnakes. *Current Biology* **2016**, *26*, 2434–2445.

**APPENDIX A: CHAPTER THREE LICENSE TERMS AND
CONDITIONS**



Academic Open Access Publishing
since 1996

MDPI AG
Postfach
CH-4020 Basel
Switzerland

Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Terms of Use

- § 1 These Terms of Use govern the use of the MDPI websites or any other MDPI online services you access. This includes any updates or releases thereof. By using our online services, you are legally bound by and hereby consent to our Terms of Use and Privacy Policy. These Terms of Use form a contract between MDPI AG, registered at St. Alban-Anlage 66, 4052 Basel, Switzerland (“MDPI”) and you as the user (“User”). These Terms of Use shall be governed by and construed in accordance with Swiss Law, applicable at the place of jurisdiction of MDPI in Basel, Switzerland.
- § 2 Unless otherwise stated, the website and affiliated online services are the property of MDPI and the copyright of the website belongs to MDPI or its licensors. You may not copy, hack or modify the website or online services, or falsely claim that some other site is associated with MDPI. MDPI is a registered brand protected by the Swiss Federal Institute of Intellectual Property.
- § 3 Unless otherwise stated, articles published on the MDPI websites are labeled as “Open Access” and licensed by the respective authors in accordance with the Creative Commons Attribution (CC-BY) license. Within the limitations mentioned in §4 of these Terms of Use, the “Open Access” license allows for unlimited distribution and reuse as long as appropriate credit is given to the original source and any changes made compared to the original are indicated.
- § 4 Some articles published on this website (especially articles labeled as “Review” or similar) may make use of copyrighted material for which the author(s) have obtained a reprint permission from the copyright holder. Usually such reprint permissions do not allow author(s) and/or MDPI to further license the copyrighted material. The licensing described in §3 of these terms and conditions are therefore not applicable to such kind of material enclosed within articles. It is the User’s responsibility to identify reusability of material provided on this website, for which he may take direct contact with the authors of the article.
- § 5 You may register or otherwise create a user account, user name or password (your “Registration”) that allows you to access or receive certain content and/or to participate or utilize certain features of our online service, including features in which you interact with us or other users. You represent and warrant that the information provided in your Registration is accurate to the best of your knowledge. You are responsible for the use of any password you create as part of your Registration and for maintaining its confidentiality, and you agree that MDPI may use this password to identify you. We reserve the right to deny, terminate or restrict your access to any content or feature reached via such Registration process for any reason, at our sole discretion. MDPI reserves the right to block or to terminate the User’s access to the website at any time and without prior notice.
- § 6 The MDPI website and online services may provide links to other websites or external resources. As part of these Terms of Use, you acknowledge that MDPI is in not responsible for the availability of such external sites or resources, and that MDPI is not liable for any content, services, advertising, or materials available from such external sites or resources.

- § 7 The website may contain advertising. MDPI does not endorse any responsibility of any kind for the content of the advertisement or sponsorship or the advertised product or service, which is the responsibility of the advertiser or sponsor, unless the advertised product or service is offered by MDPI.
- § 8 There is no warranty for the website and its content, to the extent permitted by applicable law. MDPI, the copyright holders and/or other parties provide the website and its content “as is” without representations or warranties of any kind, either expressed or implied, including, but not limited to, the implied warranties of merchantability, satisfactory quality and fitness for a particular purpose relating to this website, its content or any to which it is linked. No representations or warranties are given as to the accuracy or completeness of the information provided on this website, or any website to which it is linked.
- § 9 In no event, unless required by applicable law shall MDPI, its employees, agents, suppliers, contractors or any other party, be liable to the User for any damages of any nature, including any general, special, incidental or consequential damages, loss, cost, claim or any expense of any kind arising out of the use, inability to access, or in connection with the use of the website, its content and information, even if the User has been advised of the possibility of such damages.
- § 10 MDPI reserves the right to change these Terms of Use at any time by posting changes to this page of the website without prior notice. Please check these Terms of Use periodically for any modifications. Your continued use of any Service following the posting of any changes will mean that you have accepted and agreed to the changes.
- § 11 Basel, Switzerland shall be the place of jurisdiction for all legal disputes arising of these Terms of Use, even if the Customer has her/his domicile outside of Switzerland.
- § 12 Swiss law applicable at the place of jurisdiction of MDPI shall apply exclusively.
- § 13 If any provisions of the Terms of Use should be found invalid, this shall not affect the validity of the remaining provisions. In any such case, the contracting parties shall negotiate on the invalid clause to substitute by a valid arrangement as close as possible to the original provision.

These Terms of Use were last updated on 01. March 2017
MDPI AG, St. Alban-Anlage 66, CH-4052 Basel, Switzerland

APPENDIX B: CHAPTER TWO SUPPLEMENTAL TABLE

Table B.1: Specimen information and associated data used for analyses in Chapter 2.

Specimen ID	Museum ID	Shield et al 2018	Species	Subspecies	Venom Type	Country	State	RadSeq Clade	HPLC	MTX A2-3	MTX A3-4	MTX B2-3	MTX B3-4	MP Activity	Kallikrein	SVL (mm)	TL (mm)	HL (mm)	HW (mm)	IF (mm)	RightFL (mm)	LeftFL (mm)
070401-SPUN	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0	—	—	—	—	—	—	—	—
083101-SPUN	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0	—	—	—	—	—	—	—	—
090101-SPUN	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0	—	—	—	—	—	—	—	—
318 2013-128	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	New Mexico	1	A	—	—	—	—	0	—	—	—	—	—	—	—	—
676-25490	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0	—	—	—	—	—	—	—	—
CLP2390	CHFCB-ID 302	—	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Puebla	3	A	A	A	A	A	0	—	—	—	—	—	—	—	—
CLPT598	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0	305.8879	—	—	—	—	—	—	—
CLS899	MVZ:Herp:275538	CS0223	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Guanajuato	3	A	—	—	—	—	0	—	—	—	—	—	—	—	—
CLS900	MVZ:Herp:275555	—	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Tlaxcala	3	A	—	—	—	—	0	—	—	—	—	—	—	—	—
CLS902	MVZ:Herp:275557	CS0215	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Tlaxcala	3	A	—	—	—	—	0	—	—	—	—	—	—	—	—
Css26-SPUN	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0	—	—	—	—	—	—	—	—
NMB075	MVZ:Herp:275541	CS0216	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Guanajuato	3	A	—	—	—	—	0	—	—	—	—	—	—	—	—
OFV1124	MVZ:Herp:275552	CS0146	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Puebla	3	—	A	A	A	A	0	—	—	—	—	—	—	—	—
CLP2027	ASNHC 15003	CS0176	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.0009	733.1964	—	—	—	—	—	—	—
Css24-SPUN	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0018	—	—	—	—	—	—	—	—
102-SM161	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	—	—	—	—	0.0021	—	—	—	—	—	—	—	—
117-SM165	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	—	—	—	—	0.0034	—	—	—	—	—	—	—	—
CLPT1095	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Aguascalientes	3	A	A	A	A	A	0.0046	—	—	—	—	—	—	—	—
CLP2152	ASU 36039	CS0190	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0057	290.4145	713	61	32.71	23.63	7.31	10.52	10.12
CLS898	MVZ:Herp:275542	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Guanajuato	3	A	—	—	—	—	0.0070	—	—	—	—	—	—	—	—
CLP1971	ASU 36091	CS0172	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0072	—	717	38.71	30.83	21.01	7.85	8.68	8.51
671-25359	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0083	—	—	—	—	—	—	—	—
070402-SPUN	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0084	—	—	—	—	—	—	—	—
307 2010-160	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	—	—	—	—	0.01	—	—	—	—	—	—	—	—
CLP1959	ASU 36061	CS0169	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0106	—	730	57.64	32.22	21.01	8.65	9.16	9.07
NMB056	MVZ:Herp:275546	CS0144	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	San Luis Potosi	3	A	A	B?	B?	B?	0.0113	—	—	—	—	—	—	—	—
CLP2021	ASNHC 15002	CS0175	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.0117	352.6722	—	—	—	—	—	—	—
CLP1829	ASU 36074	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0119	132.7679	801	65	35.96	29.56	11.4	11.36	10.57
CLS894	MVZ:Herp:275532	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Aguascalientes	1	A	—	—	—	—	0.0123	—	—	—	—	—	—	—	—
99-SM156	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	—	—	—	—	0.0124	—	—	—	—	—	—	—	—
CLS896	MVZ:Herp:275533	CS0218	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Aguascalientes	3	A	—	—	—	—	0.0127	—	—	—	—	—	—	—	—
672-25471	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0133	—	—	—	—	—	—	—	—
CLP2114	ASU 36065	CS0185	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0138	372.9393	519	45.3	26.46	17.15	5.84	8.66	8.66
CLP2108	CAS 259916	CS0181	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0139	387.2598	777	63	34.07	24.77	8.37	10.79	9.66
670-25358	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0154	—	—	—	—	—	—	—	—
NMB074	MVZ:Herp:275540	CS0151	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Guanajuato	3	A	—	—	—	—	0.0158	—	—	—	—	—	—	—	—
CLP1808	ASU 36069	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0162	243.4482	680	36.99	30.61	21.08	8.07	9.59	8.24

Continued on next page

Table B.1 – continued from previous page

Specimen ID	Museum ID	Shield et al 2018	Species	Subspecies	Venom Type	Country	State	RadSeq Clade	HPLC	MTX A2-3	MTX A3-4	MTX B2-3	MTX B3-4	MP Activity	Kallikrein	SVL (mm)	TL (mm)	HL (mm)	HW (mm)	IF (mm)	RightFL (mm)	LeftFL (mm)
CLS881	MVZ:Herp:275545	CS0133	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Veracruz	3	A	A	A	A	A	0.0164	—	—	—	—	—	—	—	—
673-25473	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0172	—	—	—	—	—	—	—	—
332-SM164	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	New Mexico	1	A	—	—	—	—	0.0175	—	—	—	—	—	—	—	—
674-25482	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0188	—	—	—	—	—	—	—	—
CLP1972	ASU 36092	CS0173	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0192	340.5362	635	42	26.85	20.96	6.78	8.4	8.15
CLP2112	ASU 36036	CS0183	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0197	406.9201	643	43	30.21	20.74	7.37	9.26	8.84
CLP2154	ASU 36038	CS0192	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0201	462.6243	732	62	31.94	24.68	9.62	10.61	10.61
CLP1824	ASU 36072	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0202	275.1838	696	54	30.57	23.3	8.27	9.92	9.12
CLPT548	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.0203	—	—	—	—	—	—	—	—
CLP1930	ASNHC 14997	CS0165	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	New Mexico	1	A	A	A	A	A	0.0226	360.2572	724	37.22	32.99	23.03	8.74	8.58	8.59
CLP2076	CAS 259910	CS0178	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	A	A	A	A	0.0227	466.6292	765	42.71	33.28	22	7.85	10.98	10.96
CLP2116	ASU 36067	CS0187	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0236	280.1596	558	32	29.52	25.04	7.03	8.91	8.11
CLS895	MVZ:Herp:275537	CS0222	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Aguascalientes	3	A	—	—	—	—	0.0242	—	—	—	—	—	—	—	—
CLP2077	CAS 259911	CS0179	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	A	A	A	A	0.0251	420.5731	566	32	24.7	20.2	6.8	8.49	8.12
CLPT545	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.0263	438.1703	—	—	—	—	—	—	—
CLP1957	ASU 36076	CS0168	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0293	194.0546	643	34.73	30.4	19.76	8.45	9.05	9.04
CLP1837	ASU 36034	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0296	556.739	649	52	30.03	22.68	7.84	9.96	9.96
CLS891	MVZ:Herp:275539	CS0134	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Guanajuato	3	A	B?	A	B?	A	0.0303	—	—	—	—	—	—	—	—
LomaLindaMJ	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0315	—	—	—	—	—	—	—	—
CLP2173	ASNHC 14999	CS0197	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	New Mexico	1	A	A	A	A	A	0.0321	257.89	741	43	33.81	22.45	9.44	10.69	9.82
CLP1822	ASNHC 15007	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.0329	346.7862	758	61	35.73	25.5	7.88	11.03	10.6
CLP1836	ASU 36033	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0369	617.2976	769	65	33.92	25.6	7.97	9.95	9.52
CLPT549	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.0378	—	—	—	—	—	—	—	—
CLPT963	—	—	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Puebla	3	A	A	A	A	A	0.0416	—	—	—	—	—	—	—	—
669-25081	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.0473	—	—	—	—	—	—	—	—
CLP2075	CAS 259909	CS0177	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	A	A	A	A	0.0479	467.9035	696	57.02	31.94	22.96	7.52	9.87	9.87
CLP2389	CHFCB-ID 301	—	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Puebla	3	A	A	A	B?	A	0.05	—	—	—	—	—	—	—	—
CLP1953	ASU 36075	CS0167	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0506	—	589	51	28.65	18.53	8.02	7.96	8.03
CLP1825	ASU 36073	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0565	476.9448	662	52	30.52	24.82	8.54	9.38	9.74
CLP1936	ASU 36035	CS0166	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.059	—	441	28	23.76	15.64	5.78	6.69	5.78
CLP1821	ASNHC 15006	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.0599	415.9614	561	35	28.85	21.13	7.33	9.46	7.71
CLPT965	—	—	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Puebla	3	A	A	A	A	A	0.067	—	—	—	—	—	—	—	—
CLS876	MVZ:Herp:275544	CS0217	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Veracruz	3	A	—	—	—	—	0.0694	—	—	—	—	—	—	—	—
CLP2113	ASU 36064	CS0184	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0725	178.1565	509	31	26.17	19.66	7.93	8	7.36
CLP1963	ASU 36077	CS0171	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.0738	—	464	25.98	25.09	15.79	5.38	7.7	broken
CLPT546	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.0779	—	—	—	—	—	—	—	—
CLP2333	CHFCB-ID 245	—	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Veracruz	3	A	A	B?	B?	A	0.0808	—	—	—	—	—	—	—	—
CLS875	MVZ:Herp:275543	CS0220	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Veracruz	3	A	—	—	—	—	0.0811	—	—	—	—	—	—	—	—
070601-A	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	A	—	—	—	—	0.1108	—	—	—	—	—	—	—	—

Continued on next page

Table B.1 – continued from previous page

Specimen ID	Museum ID	Shield et al 2018	Species	Subspecies	Venom Type	Country	State	RadSeq Clade	HPLC	MTX A2-3	MTX A3-4	MTX B2-3	MTX B3-4	MP Activity	Kallikrein	SVL (mm)	TL (mm)	HL (mm)	HW (mm)	IF (mm)	RightFL (mm)	LeftFL (mm)
CLP1961	ASU 36062	CS0170	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.1162	—	564	44	28	18.81	7.48	9.03	9.05
CLPT1073	—	CS0234	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Mexico State	3	A	A	A	A	A	0.1229	—	—	—	—	—	—	—	—
CLP2175	ASU 36044	CS0199	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.1433	—	258	20	16.44	11.46	3.94	4.63	4.19
CLP2153	ASU 36037	CS0191	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.1866	—	261	14	16.54	9.4	3.54	4	4.16
CLP2191	ASNHC 15004	CS0201	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	A	A	A	A	A	0.2168	453.7650	—	—	—	—	—	—	—
CLP2382	CHFCB-ID 295	—	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Puebla	3	A	A	A	B?	A	0.2317	—	—	—	—	—	—	—	—
CLPT1074	—	CS0235	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Mexico State	3	A	A	A	A	A	0.2546	535.6830	—	—	—	—	—	—	—
CLS905	MVZ:Herp:275553	CS0221	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Querataro	3	A	—	—	—	—	0.2933	—	—	—	—	—	—	—	—
CLP2174	ASU 36043	CS0198	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	0.4195	—	257	17	15.87	10.76	3.75	4.84	4.01
CLP2182	ASNHC 15000	CS0200	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	New Mexico	1	A	A	A	A	A	0.6675	—	—	—	—	—	—	—	—
CLP1798	ASNHC 15005	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	—	A	A	A	A	—	—	561	51	31.55	20.92	8.87	broken	broken
CLP1801	ASU 36068	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	671	56	30.5	22.36	7.73	broken	broken
CLP1809	ASU 36070	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	694	61	32.46	26.34	10.32	10.61	10.89
CLP1810	ASU 36071	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	645	46	30.67	23.3	8.92	8.85	8.85
CLP1816	ASU 36032	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	632	67	32.05	24.54	8.69	8.88	9.45
CLP1980	ASNHC 15001	CS0174	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLP2096	ASU 36078	CS0180	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	693	61	30.53	22.69	8.7	broken	broken
CLP2115	ASU 36066	CS0186	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	A	A	A	A	A	—	589.2028	548	43	27.09	19.9	6.89	7.94	7.87
CLP2166	ASU 36040	CS0193	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	646	41	30.13	25.31	8.57	broken	broken
CLP2167	ASU 36041	CS0194	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLP2168	ASU 36042	CS0195	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	656	53	33.43	23.12	8.68	8.45	7.9
CLP2172	ASNHC 14998	CS0196	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	New Mexico	1	—	A	A	A	A	—	—	716	56	31.96	26.41	8.93	broken	broken
CLP72	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT162	—	CS0202	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT177	—	CS0207	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT180	—	CS0208	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT181	—	CS0209	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT195	—	CS0210	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	New Mexico	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT200	—	CS0211	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	New Mexico	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT291	—	CS0212	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT518	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT528	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT531	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT540	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT587	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT589	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT590	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT594	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLS800	MVZ:Herp:265259	CS0131	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—

Continued on next page

Table B.1 – continued from previous page

Specimen ID	Museum ID	Shield et al 2018	Species	Subspecies	Venom Type	Country	State	RadSeq Clade	HPLC	MTX A2-3	MTX A3-4	MTX B2-3	MTX B3-4	MP Activity	Kallikrein	SVL (mm)	TL (mm)	HL (mm)	HW (mm)	IF (mm)	RightFL (mm)	LeftFL (mm)
CLS811	MVZ:Herp:275565	CS0132	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	California	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLS868	MVZ:Herp:275548	CS0148	<i>C. scutulatus</i>	<i>salvini</i>	A	Mexico	Puebla	3	—	B?	A	B?	A	—	—	—	—	—	—	—	—	—
CLS908	MVZ:Herp:275567	CS0136	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Nevada	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLS909	MVZ:Herp:275568	CS0137	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Nevada	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLS911	MVZ:Herp:275561	CS0138	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	B?	A	B?	A	—	—	—	—	—	—	—	—	—
CLSRB6	MVZ:Herp:286077	CS0147	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Guanajuato	3	—	A	A	A	A	—	—	—	—	—	—	—	—	—
JAC29013	—	CS0019	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Chihuahua	2	—	A	A	B?	B?	—	—	—	—	—	—	—	—	—
JAC29014	—	CS0020	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Chihuahua	2	—	A	B?	B?	A	—	—	—	—	—	—	—	—	—
JAC29076	—	CS0022	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Chihuahua	2	—	A	B?	B?	B?	—	—	—	—	—	—	—	—	—
JAC29089	—	CS0024	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Chihuahua	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
KW1107	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
MBCS010	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Zacatecas	3	—	A	A	A	?	—	—	—	—	—	—	—	—	—
MBCS011	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Coahuila	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
MM107	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
MM240	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Texas	2	—	A	A	A	A	—	—	—	—	—	—	—	—	—
NMB037	MVZ:Herp:275535	CS0142	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Aguascalientes	3	—	B?	B?	B?	A	—	—	—	—	—	—	—	—	—
NMB038	MVZ:Herp:275536	CS0143	<i>C. scutulatus</i>	<i>scutulatus</i>	A	Mexico	Aguascalientes	3	—	A	A	A	A	—	—	—	—	—	—	—	—	—
YPMR17525	YPMHERR 017525	—	<i>C. scutulatus</i>	<i>scutulatus</i>	A	USA	Arizona	1	—	A	A	A	A	—	—	—	—	—	—	—	—	—
CLPT1050	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	Mexico	Jalisco	3	A+B	A	A	A	A	0.483	—	—	—	—	—	—	—	—
CLP2111	ASU 36063	CS0182	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	USA	Arizona	1	A+B	A	A	A	A	0.6756	216.8703	623	32	27.42	21.15	7.92	8.42	8.94
CLP1823	ASU 36060	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	USA	Arizona	1	A+B	A	A	A	A	0.6913	336.8347	597	54	31.26	24.47	7.58	9.53	8.82
CLPT1052	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	Mexico	Jalisco	3	A+B	A	A	A	A	0.7258	—	—	—	—	—	—	—	—
CLP1832	ASU 36090	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	USA	Arizona	1	A+B	A	A	A	A	0.7354	298.3029	788	55.96	33.69	23.59	8.32	9.61	10.63
CLPT1051	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	Mexico	Jalisco	3	A+B	—	—	—	—	0.8518	—	—	—	—	—	—	—	—
CLP1781	CHFCB-ID 371	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	Mexico	Jalisco	2	A+B	A	A	A	A	0.856	382.7694	—	—	—	—	—	—	—
CLPT1135	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	Mexico	Nuevo Leon	3	A+B	A	A	A	A	0.9159	—	—	—	—	—	—	—	—
CLP1830	ASU 36088	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	USA	Arizona	1	A+B	A	A	A	A	0.9467	272.9386	784	59.55	34.27	23.8	7.7	10.3	10.47
CLP2272	CHFCB-ID 186	—	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	Mexico	Coahuila	3	A+B	A	B?	B?	A	1.275	—	—	—	—	—	—	—	—
CLP1929	ASNHC 14996	CS0164	<i>C. scutulatus</i>	<i>scutulatus</i>	AB	USA	New Mexico	1	A+B	A	A	A	A	—	—	283	14	17.41	11.09	3.42	3.11	3.63
CLP1834	ASU 36101	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.2081	469.2384	661	41	32.94	24.38	7.78	10.02	10.56
CLPT479	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.5862	—	—	—	—	—	—	—	—
CLPT1136	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.6738	—	—	—	—	—	—	—	—
CLPT1141	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	B	B	B	B	B	0.6876	—	—	—	—	—	—	—	—
CLPT576	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.7019	322.6962	—	—	—	—	—	—	—
CLPT449	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.7294	—	—	—	—	—	—	—	—
CLP1833	ASU 36100	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.7755	409.7720	508	24	25.26	18.72	5.93	6.87	6.65
CLPT452	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.7755	281.2518	—	—	—	—	—	—	—
CLPT456	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.7852	—	—	—	—	—	—	—	—
CLP1786	CHFCB-ID 376	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	B	B	B	B	B	0.7998	244.5404	—	—	—	—	—	—	—

Continued on next page

Table B.1 – continued from previous page

Specimen ID	Museum ID	Schild et al 2018	Species	Subspecies	Venom Type	Country	State	RadSeq Clade	HPLC	MTX A2-3	MTX A3-4	MTX B2-3	MTX B3-4	MP Activity	Kallikrein	SVL (mm)	TL (mm)	HL (mm)	HW (mm)	IF (mm)	RightFL (mm)	LeftFL (mm)	
CLS897	MVZ:Herp:275534	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Zacatecas	3	B	—	—	—	—	0.8031	—	—	—	—	—	—	—	—	
CLP1787	CHFCB-ID 377	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	B	B	B	B	B	0.8104	356.8591	—	—	—	—	—	—	—	
CLPT476	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.8204	152.9743	—	—	—	—	—	—	—	
CLP2136	ASU 36103	CS0188	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.8281	101.5176	1030	69.35	40.67	32.56	10.89	13.12	13.63	
CLPT1137	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Chihuahua	2	B	B	B	B	B	0.8411	—	—	—	—	—	—	—	—	
CLS892	MVZ:Herp:275549	CS0135	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Zacatecas	3	B	B	B	B	B	0.8483	—	—	—	—	—	—	—	—	
CLPT578	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.8583	—	—	—	—	—	—	—	—	
CLP1783	CHFCB-ID 373	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.8797	139.3213	—	—	—	—	—	—	—	
CLPT577	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.8856	229.9772	—	—	—	—	—	—	—	
CLPT1134	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.8871	—	—	—	—	—	—	—	—	
CLPT1069	—	CS0230	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	B	B	B	B	B	0.8949	—	—	—	—	—	—	—	—	
CLP1784	CHFCB-ID 374	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.8951	405.4638	—	—	—	—	—	—	—	
CLPT1091	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Chihuahua	2	B	B	B	B	B	0.9112	—	—	—	—	—	—	—	—	
CLPT1179	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.9158	—	—	—	—	—	—	—	—	
CLPT1094	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Jalisco	2	B	B	B	B	B	0.9188	—	—	—	—	—	—	—	—	
CLP1780	CHFCB-ID 370	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.9198	244.1763	—	—	—	—	—	—	—	
CLPT1071	—	CS0231	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	B	B	B	B	B	0.927	—	—	—	—	—	—	—	—	
MBCS006	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.9303	214.9286	—	—	—	—	—	—	—	
CLP2142	ASU 36104	CS0189	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.9346	353.6430	775	56.86	34.34	24.89	7.67	10.04	9.82	
CLPT1138	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Chihuahua	2	B	B	B	B	B	0.937	—	—	—	—	—	—	—	—	
CLPT1177	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.9494	—	—	—	—	—	—	—	—	
CLP1794	ASU 36096	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.9559	319.2982	556	42	28.27	21.52	6.98	8.83	8.83	
CLPT579	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.9597	—	—	—	—	—	—	—	—	
CLPT1072	—	CS0232	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.9938	—	—	—	—	—	—	—	—	
CLP1778	CHFCB-ID 369	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	0.9961	257.2832	—	—	—	—	—	—	—	
CLPT581	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	0.9969	—	—	—	—	—	—	—	—	
CLP1785	CHFCB-ID 375	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	1.0119	362.1382	—	—	—	—	—	—	—	
CLP1828	ASU 36099	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	1.0226	70.6922	766	55.55	32.77	24.69	8.05	10.49	10.39	
CLP1831	ASU 36089	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	1.0329	195.8144	795	51	35.63	24.44	9.27	8.86	9.61	
CLP2261	CHFCB-ID 177	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	1.0508	—	—	—	—	—	—	—	—	
CLPT580	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	1.0543	308.9219	—	—	—	—	—	—	—	
CLP1777	CHFCB-ID 368	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	1.06	280.2202	—	—	—	—	—	—	—	—
CLP2262	CHFCB-ID 178	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	1.0608	—	—	—	—	—	—	—	—	—
CLPT1090	—	CS0233	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Chihuahua	2	B	B	B	B	B	1.0612	—	—	—	—	—	—	—	—	—
CLP1827	ASU 36098	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	1.0792	287.9873	706	41	34.27	24.72	8.8	10.18	10.57	
CLP1835	ASU 36102	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	1.1616	385.318	685	34.97	29.26	20.86	6.35	7.76	7.73	
CLP1782	CHFCB-ID 372	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	1.2258	285.8635	—	—	—	—	—	—	—	—
CLP1826	ASU 36097	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	B	B	B	B	B	1.226	295.5723	739	65	34.57	24.17	8.7	10.79	10.14	
CLPT455	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	1.9306	294.6014	—	—	—	—	—	—	—	—

Continued on next page

Table B.1 – continued from previous page

Specimen ID	Museum ID	Schild et al 2018	Species	Subspecies	Venom Type	Country	State	RadSeq Clade	HPLC	MTX A2-3	MTX A3-4	MTX B2-3	MTX B3-4	MP Activity	Kallikrein	SVL (mm)	TL (mm)	HL (mm)	HW (mm)	IF (mm)	RightFL (mm)	LeftFL (mm)
CLP2276	CHFCB-ID 190	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	3	B	B	B	B	B	2.0433	—	—	—	—	—	—	—	—
CLP2163	ASU 36095	—	<i>C. cerberus</i>	—	B	USA	Arizona	Outgroup	—	—	—	—	—	—	—	—	—	—	—	—	—	—
CLP2046	CAS 259882	—	<i>C. oreganus</i>	—	B	USA	California	Outgroup	—	—	—	—	—	—	—	—	—	—	—	—	—	—
ANMO1369	MZFC22927	CS0141	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Zacatecas	3	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLP1813	ASU 36087	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	—	B	B	B	B	—	—	637	33	28.07	22.44	7.12	broken	8.14
CLPT1148	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	3	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT1149	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	3	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT1150	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	3	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT167	—	CS0203	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT172	—	CS0204	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT173	—	CS0205	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT174	—	CS0206	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT451	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT471	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT474	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT475	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Coahuila	2	B	B	B	B	B	—	215.1713	—	—	—	—	—	—	—
CLPT500	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT586	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLPT591	—	—	<i>C. scutulatus</i>	<i>scutulatus</i>	B	USA	Arizona	1	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLS893	MVZ:Herp:275550	CS0214	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Zacatecas	3	—	B	B	B	B	—	—	—	—	—	—	—	—	—
DAR108	—	CS0149	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	San Luis Potosi	3	—	B	B	B	B	—	—	—	—	—	—	—	—	—
JAC 29015	—	CS0021	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Chihuahua	2	—	B	B	B	B	—	—	—	—	—	—	—	—	—
JMM242	MZFC25003	CS0150	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	3	—	B	B	B	B	—	—	—	—	—	—	—	—	—
KWS294	MZFC17996	CS0139	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	Durango	2	—	B	B	B	B	—	—	—	—	—	—	—	—	—
NMB058	MVZ:Herp:275547	CS0145	<i>C. scutulatus</i>	<i>scutulatus</i>	B	Mexico	San Luis Potosi	3	—	B	B	B	B	—	—	—	—	—	—	—	—	—
CLP1914	ASNHC 15008	—	<i>C. viridis</i>	—	B	USA	New Mexico	Outgroup	—	—	—	—	—	—	—	—	—	—	—	—	—	—

APPENDIX C: CHAPTER FOUR SUPPLEMENTAL TABLE

Table C.1: Differential toxin expression output for the four methods used. Genes not shown were not differentially expressed in any analysis. For the log-fold changes, anything that is negative was higher in the first venom type being compared.

Toxin	RNentropy			RNentropy Pairwise			EBSeq	EBSeq Pairwise			DESeq2 log fold change			edgeR log fold change		
	A	AB	B	A to B	A to AB	B to AB	Pattern	A to B	A to AB	B to AB	A to B	A to AB	AB to B	A to B	A to AB	AB to B
3FTx-1						- 0	AB Unique		UpAB	UpAB						
BPP-1	0	-	+	-	+	+										
CRISP-1	0	-	-	0	-	0	A Unique	UpA	UpA		-1.20					
CTL-3	-	0	+	-	+	- 0	A Unique	DownA	DownA		12.41			12.31	10.07	
CTL-1	0	0	+	-	+	+										
CTL-11	0	0	+	-	+	+										
CTL-12	+	0	-	+	-	0	B Unique	DownB		DownB				-3.46		
CTL-2	0	0	+	-	+	+										
CTL-6	0	-	+			+										
CTL-17	-	-	0	-	0	0	B Unique	UpB		UpB				10.31		
CTL-23	0	-	0	0	-	0	A Unique	UpA	UpA							
CTL-22	-	-	+	-	+	+	B Unique	UpB		UpB				5.57		
CTL-13	0	0	+	0	+	+										
CTL-18	0	-	-	0	-	+	A Unique	UpA	UpA			-3.42			-3.40	
CTL-20	-	-	0	-	0	0								10.29		
CTL-7	0	0	+	0	+	+										
CTL-19	0	0	-	0	-	- 0										
CTL-9	+	-	-	+	-	+	B Unique	DownB		DownB						
CTL-4	-	0	+	-	+	- 0	A Unique	DownA	DownA		13.68	12.35		13.47	12.13	
CTL-21	-	-	+	-	+	+	B Unique	UpB		UpB	7.21			7.06		
CTL-5	0	0	+	0	+	+										
CTL-16	0	0	-	+	-	0	B Unique	DownB		DownB						
CTL-10	0	0	+			+										
MYO-6	-	-	0	-	0	0										
MYO-7	-	-	0	-	0	0										
MYO-5	-	0	0	-	0	- 0	A Unique	DownA	DownA							
MYO-1	0	0	-	0	-	- 0										
MYO-2	0	0	-	0	-	- 0										
MYO-8	0	-	-	0	-	0								-11.11		
NUC-1	0	-	0			0										
PDE-1	0	-	0	0	-	0										
Pla2gA2-MTXA	+	0	-	+	-	-	B Unique	DownB		DownB	-13.39		-13.76	-13.37		-13.22
Pla2gB1	0	0	+	-	+	+	A Unique	DownA	DownA							

Continued on next page

Table C.1 – continued from previous page

Toxin	RNentropy			RNentropy Pairwise			EBSeq	EBSeq Pairwise			DESeq2 log fold change			edgeR log fold change			
	A	AB	B	A to B	A to AB	B to AB	Pattern	A to B	A to AB	B to AB	A to B	A to AB	AB to B	A to B	A to AB	AB to B	
Pla2gK	0	0	+	0	+												
Pla2gB2-MTXB	0	0	-	+	-		B Unique	DownB		DownB	-13.53		-13.26	-13.51			-13.72
PLA2-6	0	0	+	0	+												
PLB-1							B Unique	UpB		UpB							
SVMPI-1	-	0	0	-	+	-	A Unique	DownA	DownA					6.77	8.58		
SVMPII-4	-	0	-			-	AB Unique		UpAB	UpAB							4.78
SVMPII-1	-	0	+	-	+	-	A Unique	DownA	DownA			10.16					10.11
SVMPII-3	-	+	+	-	+	-	A Unique	DownA	DownA		11.75	11.98		11.02	11.22		
SVMPII-2	-	+	0	-	+	-	A Unique	DownA	DownA		9.30	9.89		9.19	9.76		
SVMPIII-10	-	+	0	-	+	-	A Unique	DownA	DownA		11.77	12.77		10.49	11.46		
SVMPIII-3	-	0	+	-	+	-	A Unique	DownA	DownA					7.84	7.51		
SVMPIII-1	-	+	+	-	+	-	A Unique	DownA	DownA					10.38	10.25		
SVMPIII-4	0	0	+	0	+												
SVMPIII-9	-	0	0	-	+	-	A Unique	DownA	DownA		12.79	12.94		11.85	11.96		
SVMPIII-6	-	0	0	-	+	-	A Unique	DownA	DownA		6.70	6.23		6.72	6.22		
SVMPIII-7	-	+	+	-	+	-	A Unique	DownA	DownA		10.70			10.39	10.56		
SVMPIII-5	-	0	+	-	+	-	A Unique	DownA	DownA		9.43	9.48		9.47	9.48		
SVMPIII-8	-	-	0	-	+	-	All Different	UpB	UpAB	UpB		8.14		12.45	6.31		
SVSP-16	0	0	-	0	-												
SVSP-1	0	0	+	0	+												
SVSP-5	0	0	-	0	-								-1.85				
SVSP-10	0	-	0			0											
SVSP-3	0	-	0			0											
SVSP-2	+	0	-	+	-	+	B Unique	DownB		DownB	-1.72						
SVSP-11	0	-	0			0											
SVSP-4	0	0	+														
SVSP-15	0	0	+	0	+												
Vespryn-1	0	-	0			0											