

Article

Reduced Toxicity of *Centruroides vittatus* (Say, 1821) May Result from Lowered Sodium β Toxin Gene Expression and Toxin Protein Production

Aimee Bowman ¹, Chloe Fitzgerald ¹, Jeff F. Pummil ² , Douglas D. Rhoads ³  and Tsunemi Yamashita ^{1,*}

¹ Department of Biological Sciences, Arkansas Tech University, 1701 N Boulder Ave., Russellville, AR 72801, USA; aimeebowman13@gmail.com (A.B.); chughes11@atu.edu (C.F.)

² High Performance Computing Center, University of Arkansas-Fayetteville, 850 W Dickson St., Fayetteville, AR 7270, USA; jpummil@uark.edu

³ Department of Biological Sciences, University of Arkansas-Fayetteville, 850 W Dickson St., Fayetteville, AR 7270, USA; drhoads@uark.edu

* Correspondence: tyamashita@atu.edu

Abstract: Body tissue and venom glands from an eastern population of the scorpion *Centruroides vittatus* (Say, 1821) were homogenized and molecular constituents removed to characterize putative sodium β toxin gene diversity, RT-qPCR, transcriptomic, and proteomic variation. We cloned sodium β toxins from genomic DNA, conducted RT-qPCR experiments with seven sodium β toxin variants, performed venom gland tissue RNA-seq, and isolated venom proteins for mass spectrophotometry. We identified >70 putative novel sodium β toxin genes, 111 toxin gene transcripts, 24 different toxin proteins, and quantified sodium β toxin gene expression variation among individuals and between sexes. Our analyses contribute to the growing evidence that venom toxicity among scorpion taxa and their populations may be associated with toxin gene diversity, specific toxin transcripts variation, and subsequent protein production. Here, slight transcript variation among toxin gene variants may contribute to the major toxin protein variation in individual scorpion venom composition.

Keywords: *Centruroides*; sodium β toxin; RT-qPCR; transcriptome; proteome

Key Contribution: Putative sodium β toxin gene, transcriptome, and proteome variation in the mildly toxic *Centruroides vittatus* is reported. Transcript variation in sodium β toxin genes and toxin protein composition may account for the reduced human toxicity effects of this scorpion.



Citation: Bowman, A.; Fitzgerald, C.; Pummil, J.F.; Rhoads, D.D.; Yamashita, T. Reduced Toxicity of *Centruroides vittatus* (Say, 1821) May Result from Lowered Sodium β Toxin Gene Expression and Toxin Protein Production. *Toxins* **2021**, *13*, 828. <http://doi.org/10.3390/toxins13110828>

Received: 27 July 2021

Accepted: 9 November 2021

Published: 22 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Scorpions of the genus *Centruroides* constitute the most medically important and one of the most diverse and wide-ranging scorpion taxa in North America [1,2]. The complex evolutionary history and geographic variability of this genus has generated controversy and taxonomic confusion [2–5]. This complex evolutionary history also appears in the diversity of the venom components. Recent studies show intraspecific complexity and diversity in the protein and peptide constitution of the venom components [6,7]. In addition, age-specific and sex-specific variation has been reported in *Centruroides vittatus* and *C. hentzi* [8–10]. These studies suggest that a rich evolution of venom components as *Centruroides* species and their populations expanded and diversified into their present geographic range.

Scorpion venom toxicity results from a wide range of venom components (peptides, lipids, nucleotides, and other compounds) [11]. Most research has focused upon characterizing proteins and gene activity via proteomic and transcriptomic studies [10,12,13]. These proteins have been shown to affect sodium, potassium, chloride, and calcium ion channels primarily in muscle or nervous tissue [11,12]. Voltage-gated sodium channels (VGSCs) are critical to initiate action potentials in excitable cells [14,15]. Scorpion toxin proteins

are especially interesting as they have the ability to discriminate among subtypes of the phylogenetically conserved sodium channels, which are difficult to discriminate with most drugs [15,16]. Two classes exist with scorpion sodium toxins, an α toxin class and a β toxin class [11,12]. Although these toxins affect receptor sites on the sodium channel external pore surface (gating module) rather than the pore, both target distinct receptor sites [11,12]. The α class toxins are those that inhibit sodium channel inactivation through binding at receptor site 3. These α toxins bind sodium channels open, which allows sodium influx to compromise cell activity. The β toxins act as sodium channel modifiers, through shifting the voltage dependence of activation in a negative direction promoting spontaneous and repetitive sodium channel opening after binding to receptor site 4 allowing sodium influx [14,17]. Studies have shown that scorpion sodium β toxins affect sodium channel voltage sensors that respond to membrane depolarization [18].

Research has linked amino acid residue conservation at specific sites with toxin activity [11,12]. Other studies have conducted mutagenesis at specific sites to create toxins with reduced activity at sodium channels [11,19]. Minute changes in toxin proteins such as amino acid substitutions, deletions and modifications at short protein sequences have resulted in higher toxicity to mice [11,20]. Other recent studies have illustrated similar functional surface anatomy of β toxins in their alpha helix and at the C-terminus that confer interactions with sodium channel sites [18]. In addition, removal of residues at the N and C terminus in β toxins resulted in a non-toxic, truncated 46AA protein, unable to bind to its original sodium channel receptor but yet showed activity with the sodium channel [19]. Furthermore, more recent data have shown a multitude of variation between α and β forms and suggest that undescribed unique structural and functional relationships exist [11,12,14].

The scorpion *Centruroides vittatus* encompasses a large geographic range across the western USA and northern United Mexican States (Figure 1). Although a member of the toxic *Centruroides* genus, this species is not known as medically important [21]. However, due to coevolution with mammalian predators, evidence suggests that western *C. vittatus* populations may possess a more medically significant venom than eastern populations [22]. Throughout its geographic range, *C. vittatus* is commonly found in diverse ecological habitats, but in populations across the northern and eastern geographic distributions it appears to prefer dry, rocky south facing slopes or glade areas. Human introduction of this scorpion appears to also have created additional populations outside its known geographic range [23].

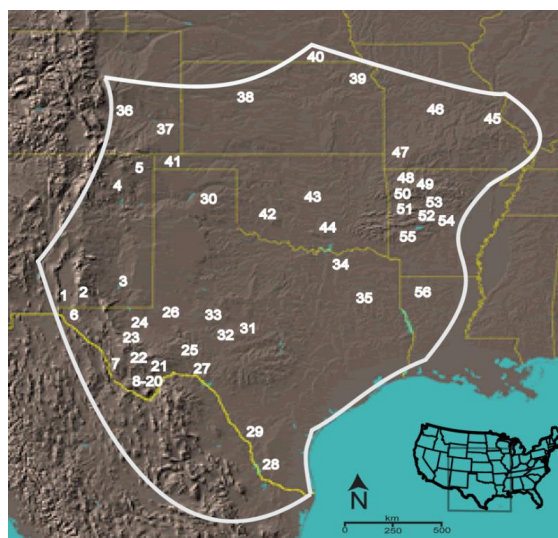


Figure 1. Approximate geographic range of *C. vittatus* with populations numbered as surveyed in a phylogeographic analysis [24]. The populations selected for the putative sodium β toxin gene

diversity analysis are the following: (1) Aguirre Springs (AgSP), (6) Hueco Tanks (HT), (7) Chinati Hot Spring (Chin), (20) Boquillas Canyon Road (BCR), (30) PaloDuro (PaDu), (37) Lake Pueblo (LPb), (39) Lawrence (KS), (40) Little Blue River (LBR), (41) Black Mesa (BMe), (46) Ha Ha Tonka (Ha), (53) Scottsville (Scv), and (56) Kisatchie (Kisa).

We hypothesize that potentially limited sodium β toxin population gene diversity and gene transcript variation along with a decreased proteome diversity and abundance may contribute to a reduced medically significant venom in this scorpion. Our findings of putative sodium β toxins and similar peptides suggest that genomic, transcriptomic, and proteomic components contribute to venom variation among populations potentially resulting in a reduced medically significant venom when compared to other medically important *Centruroides* species.

2. Results

2.1. Sodium β Toxin Variability Analysis

Through the sodium β toxin cloning experiment, we identified a total of 79 putative *C. vittatus* sodium β toxins with 72 putative new sodium β toxins and select toxins presented in Figure 2. All toxin variants are shown in Table S1. Sixteen identified sequence variants may represent pseudogenes as they housed internal stop codons or stop codons that created truncated proteins. Two toxins were commonly identified across eight and six populations (Table 1). In general, we show that western and southern populations (e.g., AgSp and Kisa) populations show greater sodium toxin gene diversity than northern and eastern populations (e.g., LBR and KS) (Table S1).

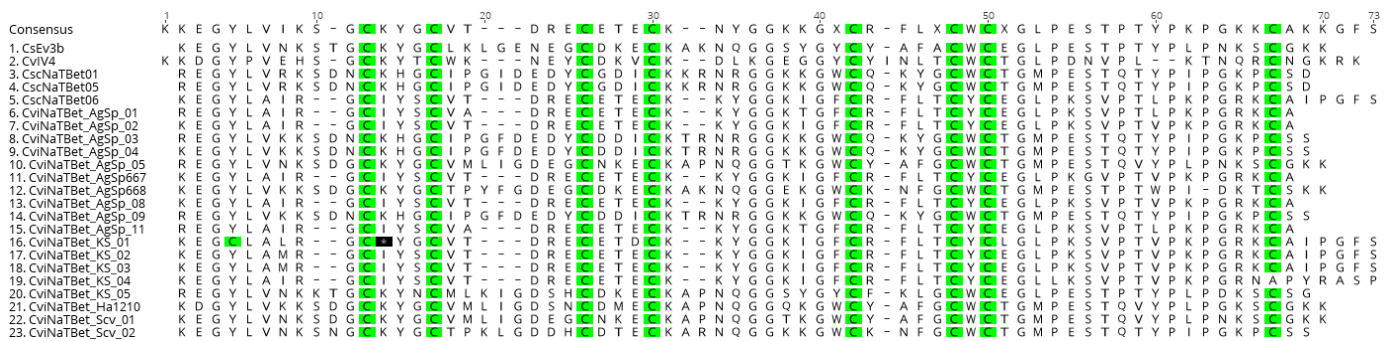


Figure 2. Select putative *C. vittatus* sodium β Toxin polypeptide sequences from western (AgSp) and eastern populations (KS, MO, AR) cataloged from a cloning survey of 10 population groups identified in a *C. vittatus* phylogeographic analysis. Also included are CsEv3b and CvIV4 toxin sequences and additional *C. sculpturatus* Beta toxin sequences obtained through additional cloning experiments. Cysteine residues are shown in green. The alignment for these sequences was conducted in Muscle (3.8.425) in the Geneious 11.1.5 package.

2.2. RT-qPCR Experiments

The data from the RT-qPCR experiments suggest that no significant expression difference occurs among males; however, there was high variation among samples (Figure 3, Table 2, Table S2). Females exhibited few significant differences, yet sodium toxin Ha1210, unique to eastern populations, showed a higher expression value than all other sampled sodium toxin genes, with significant differences to Chin654, AgSp668, and CvIV4. These data suggest that the pain-inducing toxin, CvIV4, is expressed in lower amounts and may contribute to the lower pain associated with eastern scorpions.

Table 1. The five most common putative sodium β Toxins and their AA sequences identified across populations (98-100% similarity) from a cloning survey and ranked in order of common occurrence across populations. These AA sequences were identified in multiple populations.

CviNaTBet_HT_02
REGYLAIKRCIYSCVTDRECETECKKYGGKIGFCRFLTCYCEGLPKSVPTVPKPGRKCA
8 populations: HT, AgSp, Chin, LBR, LPb, PaDu, Kisa, KS
CviNaTBet_HT_01
KEGYLVNKSDGCKYGCVMLIGDEGCNKECKAPNQGGTKGWCYAFGCWCTGMPESTQVYPLPNKSCGKK
6 populations: HT, AgSp, LBR, LPb, PaDu, Scv
CviNaTBet_Chin_05
KEGYLVKSDNCKHGCIPGIDEDFCDDICKARNRGGKKGWCQKYGCWCTGMPESTQTYPIPGKPCSS
4 populations: AgSp, Chin, BCR, Kisa
CviNaTBet_Chin_01
KEGYLVKSDGCKYGCVMLIGDSNCDMECKAPNQGGQKGCWYAFGCWCTGMPESTQVYPLPGKSCGKK
3 populations: Chin, LBR, Kisa
CviNaTBet_HT_08
REGYLVNKKTGCKYGCPTKLGDDHCDKECKAPNQGGKKGWCKNFGCWCTGMPESTQVYPIPGKSCSS
3 populations: HT, AgSp, BCR

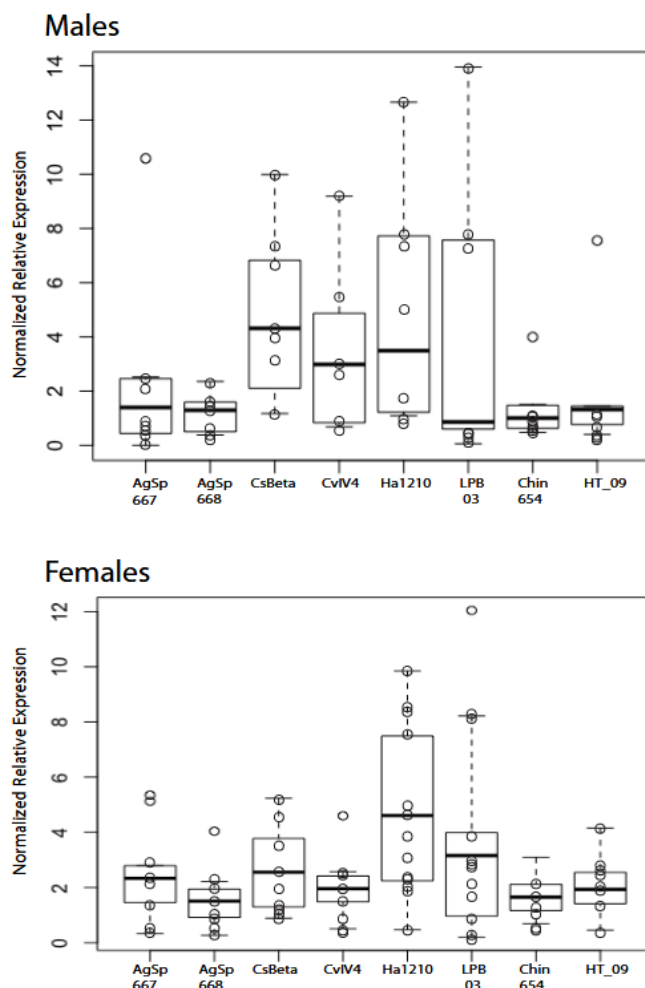


Figure 3. Box plots of normalized relative expression values for eastern population males ($n = 8$) and females ($n = 13$) with eight putative sodium β toxin genes. Normalized relative expression values (open circles for each individual) represent sodium β toxin genes adjusted to reference gene expression. The error bars represent SEMs. The only significant differences exist in females with Ha1210 to Chin654, AgSp668, and CvIV4 (see below).

Table 2. Significant values between select putative sodium β toxins for Tukey's pairwise test in females from normalized relative expression results in eight sodium β toxins and four reference genes. These three pairs represent significant values from 28 pairwise comparisons. Males did not exhibit significant differences among sodium beta toxin gene activity as determined by normalized relative expression values. See Table S2 for one-way ANOVAs and the complete Tukey's pairwise test.

	Na Tox Gene	Diff lwr	upr	p adj	
	Ha1210-Chin654	2.958	0.619	5.297	0.0039
	Ha1210-AgSp668	3.104	0.765	5.442	0.002
	Ha1210-CvIV4	2.651	0.312	4.99	0.0149

2.3. Transcriptome

The basic transcriptome contig statistics are shown in Table 3. The BLASTn query identified 1491 unique hits (Blastgrabber) with 111 total toxin gene ids and 26 sodium toxin genes (Table 4). The blast query of the transcriptome with the NCBI 2133 toxin gene sequences produced 1142 significant alignments with the following toxin categories (Table S3): NaScTx (19), Metalloproteases (10), KScTx (6), Phospholipases (2), IGFBP (2), Host Defense Peptides (2), and Serpins (1). We also discovered eleven additional sodium β toxin transcripts in our transcriptome of an eastern population when compared to the five revealed in the cloning experiment for toxin variability from the same regional population.

Table 3. Transcriptome contig statistics from an assembly from three *C. vittatus* Telsons (Venom gland).

Assembled Contig Number	226,162
Assembled contigs > 500 bp	48,463
Assembled contigs > 1000 bp	26,909
Average length of assembled contigs (bp)	463
Longest assembled contig (bp)	9430
Total length (bp)	104,750,193
Assembled contig N50	1149

Table 4. RPKM values for Na toxins from a transcriptome contig assembly of *C. vittatus* body tissue (e.g., MC1-Male Carapace 1) and telson (venom gland, e.g., MT1-Male Telson 1) for two males and one female from a NW Arkansas population. Values for housekeeping and RT-qPCR reference genes are also shown. The six putative sodium β toxin genes in the first rows represent those targeted in the RT-qPCR experiments. The second putative sodium β toxin genes are additional unique transcripts identified after a BLAST query of the transcript to a scorpion toxin gene database. The last genes presented are commonly expressed genes and reference genes utilized in the RT-qPCR experiments. Accession numbers are provided for genes identified in the NCBI database.

	MC1	MT1	MC2	MT2	FC1	FT2
CviNaTBet_Ha1210	4	425	15	7638	4	4134
CviNaTBet_Chin654	5	502	17	8923	5	4932
CviNaTBet_AgSp667	3	523	12	6482	2	329
CviNaTBet_AgSp668	4	437	16	8682	4	3476
JF938594.1 Cv Alpha Tox IV4	0	224	1	406	0	199
AF338462.1 CsEv3b	4	365	21	11,318	4	1776
L05060.1_CUDNATOXA C_noxius						
Na channel blocker	4	363	21	11,278	4	1703
HQ262494.1 Css beta neurotox	1	356	4	2030	1	738
AF491130.1 C_limp Na channel Mod tox 5c	2	303	7	3976	1	165
AF491134.1 C_limp Na channel 8	2	227	8	4031	2	1411
AY164271.1 C_nox ergtox1	1	429	4	1830	1	469
AY159340.1 C_gracil ergtox1	0	59	0	43	0	54
AY649871.1 C_ex 13 tox	5	542	17	8941	5	4344
AY649868.1 C_ex10	2	151	8	4086	3	2657

Table 4. Cont.

	MC1	MT1	MC2	MT2	FC1	FT2
AF338460.1 CsE1a	0	82	3	1471	1	1
AF338450.1 Cse8	4	386	22	12,095	4	1845
AF338448.1 CsE1x	2	268	9	4712	2	1565
AY649860.1 Cex2	4	360	21	11,258	4	1714
CviNaTBet_HT_04	2	121	11	6028	2	92
CviNaTBet_HT_08	3	176	10	5376	2	1597
CViNaTBet_HT_09	4	348	13	7181	4	3127
CviNaTBet_Kisa_01	2	200	9	4681	3	2544
CviNaTBet_LBR_03	2	161	8	3815	2	1944
CviNaTBet_LPb_04	2	122	8	3700	2	2255
CviNaTBet_Scv_02	3	139	12	6419	2	176
AF439766.1 C_limpidus 16s ribosomal	3392	2005	1894	1933	2204	1268
AH010433.1 CEX RNA poly	0	36	37	22	32	43
AY995829.1 C_nox COI	5951	2594	3837	1696	5481	2014
AY995849.1 C_Sc COII	2849	1251	1719	419	3458	1342
EU381110.1 C_vit COI	2357	1268	2393	547	2788	1352
CB334087.1 M_gib EF1	244	297	167	78	230	302
BU092001.1 M_gib EF2	1142	1103	721	369	881	944
CB334044.1 M_gib RPL19	255	260	147	105	273	244

The summary RPKM values for the two adult males and one female support the RT-qPCR experiments with higher expression values for sodium β toxins vs. the pain inducing Cv Alpha IV4 toxin (Table 4). In addition, the reference genes chosen for the RT-qPCR experiments show little deviation between body tissue and venom glands (i.e., EF1, EF2, RPL19).

2.4. Toxin Proteomics

The proteomic analysis from the telsons of three males and four females resulted in 273 proteins identified in the total spectra with 24 different toxin proteins and nine different sodium β toxins (Table 5 and Table S4). Among the sodium β toxins, CviNaTBet_Ha1210 exhibits the highest total abundance with the two other sodium β toxins targeted in the RT-qPCR experiments exhibiting much lower values. The sodium β toxins identified from other species show lower NSAF values than putative sodium β toxins in these individuals.

Table 5. Sodium α and β toxin protein abundance as identified via Normalized Spectral Abundance Factor (NSAF) in seven adult scorpions telsons from a NW Arkansas population. The first three toxins are among those targeted in the RT-qPCR experiments. The complete list of identified toxin proteins with accession numbers is in Table S4.

	Molecular Weight	Female 1	Female 2	Female 3	Female 4	Male 1	Male 2	Male 3
CviNaTBet_Ha1210	7 kDa	4.66%	15.08%	12.69%	10.03%	5.79%	5.66%	6.91%
CviNaTBet_AgSp668	7 kDa	1.09%	1.88%	0.58%	1.23%	na	na	0.35%
alpha-toxin, partial [<i>C. vittatus</i>]	11 kDa	0.75%	0.65%	1.21%	0.64%	0.81%	0.44%	0.73%
CviNaTBet_HT_01	7 kDa	2.15%	8.12%	5.48%	na	2.70%	na	na
CviNaTBet_HT_06	7 kDa	2.07%	2.94%	1.66%	1.05%	7.13%	3.62%	0.79%
CviNaTBet_Chin05	7 kDa	1.09%	1.88%	1.46%	1.23%	0.78%	3.51%	0.70%
CviNaTBet_BCR_10	7 kDa	1.09%	1.88%	1.46%	1.23%	0.78%	3.51%	0.70%
alpha-toxin CvIV4 [<i>C. sculpturatus</i>]	9 kDa	0.29%	0.95%	0.47%	0.25%	0.32%	0.52%	0.28%
RecName: Full = Toxin Cg2	8 kDa	na	na	0.87%	0.30%	na	na	0.35%
RecName: Full = Alpha-toxin Cn12	7 kDa	na	na	0.29%	0.31%	na	na	0.70%
beta-toxin Im-2-like [<i>C. sculpturatus</i>]	9 kDa	na	na	0.23%	na	na	na	0.55%
beta-toxin CeII8-like [<i>C. sculpturatus</i>]	10 kDa	na	na	0.23%	0.25%	na	na	na
RecName: Full = Alpha-toxin CsE5	7 kDa	na	na	0.31%	0.33%	na	na	0.37%
beta-neurotoxin CsxIX precursor [<i>C. suffusus suffusus</i>]	9 kDa	na	na	0.24%	na	na	na	na

3. Discussion

3.1. Sodium β Toxin Gene Diversity

Our sodium β toxin diversity analysis shows marked variability among populations in toxin gene diversity. Western populations exhibit >10 putative variants, whereas eastern populations (e.g., KS) exhibit five putative toxin variants. Western and southern populations may exhibit higher toxin gene diversity, presumably due to longer population isolation and separation. A phylogeographic analysis of *C. vittatus* populations showed western and southern populations experienced longer divergence times than those populations in the eastern and northern portions of their geographic range [24]. In addition, western scorpion populations overlap with grasshopper mice populations, a key predator of scorpions. These scorpion populations appear to have coevolved a more painful mammalian toxin to deter mice predation with subsequent increased pain tolerance from grasshopper mice [22].

Interestingly, our analysis of toxin gene variants suggests that pseudo genes may exist in the genome with increased numbers identified from samples in western populations (Table S1). Many of the sequence variants coded for polypeptides with the number of AA residues similar to that seen in known variants; however, nine putative toxin variants show increased length to ~80 AAs, with six showing 10 cysteine residues instead of the 8 typically noted for β toxins and may represent another class of venom toxins. The additional cysteine residues are unique to this study but these variants appear to show AA homology with shorter known toxin variants. As this dataset did not separate mammalian and insect toxins, nor demonstrate physiologically alpha versus beta toxins, the existence of these putative toxin proteins will need to be confirmed through further proteomic and transcriptomic analyses of other *Centruroides* and buthid species. Lastly, mild toxicity in other buthid scorpions appears to result from a substitution of the lysine from the first amino acid in the polypeptide [25]. *C. vittatus* populations do not appear to harbor this substitution; therefore, we speculate that differential gene expression explains the lowered toxicity in these populations.

3.2. Sodium β Toxin Gene Expression Variation

Our gene expression analysis via RT-qPCR shows more variation in venom gene activity in males than females as indicated by the normalized relative expression values (Figure 3). This result is echoed in the results from a proteomic analysis of venom components in populations of *C. hentzi* [7]. Male *C. vittatus* scorpions move more extensively than females and venom diversity may be a byproduct of male movement as males encounter more microclimates and potential selection pressures for venom variation [7,8,26]. The 13 females show the highest normalized relative expression values for the putative toxin CviNaTBet_Ha1210, (Figure 3) and this result is also seen in the RPKM values in the transcriptome contig assembly and the NSAF percentages in the protein abundance (Tables 4 and 5). The pain-inducing toxin, Cv Alpha Tox IV4, shows a lower normalized relative expression value along with four other toxins, and also exhibits low values in the transcriptome contig assembly and the NSAF percentages in the protein abundance. These results may explain how a less painful toxin occurs in this species.

3.3. Sodium β Toxin Transcriptome and Proteome

Our transcriptomic and proteomic data align with other studies of *Centruroides* venom transcript and proteomic studies showing the majority of toxin transcripts are those of sodium toxins [7,9,10]. If sodium toxin transcripts and protein diversity are compared to each other, we show 19 transcripts but only seven sodium toxin variants in the proteome. This finding is also reflected in *C. hentzi*, where 36 sodium toxin transcripts were revealed but nine sodium toxins were identified in the proteome [10]. Interestingly, a western *C. vittatus* population appears to show similar sodium toxin transcript numbers (13) to our study [9], which suggests that more toxin gene variants may exist in these populations. In addition, our combined datasets indicate preferential expression and production of putative

sodium β toxin variants with a reduction in the transcripts and protein abundance of the known pain-inducing toxin, Cv Alpha Tox IV4. However, we note that as entire telsons were assayed for transcriptomic and proteomic analyses, our dataset may reflect lowered toxin transcript and protein abundances. Yet, this trend of greater sodium toxin transcript diversity when compared to sodium toxin protein diversity appears to be reflected in other Buthid species (Table 6).

Table 6. A comparison among Buthid scorpion sodium channel toxin transcript and protein diversity as reported in this study and the literature. Alpha and Beta sodium variants were not separated in this Table.

Scorpion Species	Sodium Channel Toxin Transcript Diversity	Sodium Channel Toxin Protein Diversity	Reference
<i>Centruroides vittatus</i>	19	7	This paper
<i>Centruroides vittatus</i>	13	-	McElroy et al. 2017 [9]
<i>Centruroides hentzi</i>	36	9	Ward et al. 2018 [10]
<i>Centruroides sculpturatus</i> *	22	-	Corona et al. 2001 [27]
<i>Centruroides sculpturatus</i> *	-	5/6	Carcamo-Noriega et al. 2018 [6]
<i>Centruroides limpidus</i> *	59	26	Cid-Urbe et al. 2019 [13]
<i>Centruroides noxius</i> *	27	-	Rendón-Anaya et al. 2012 [17]
<i>Centruroides tecomanus</i> *	24	30	Valdez-Velázquez et al. 2013 [28]
<i>Centruroides hirsutipalpus</i> *	77	31	Valdez-Velázquez et al. 2020 [29]
<i>Tityus obscurus</i> *	48	3	Oliveira et al. 2018 [30]
<i>Tityus serrulatus</i> *	24	7	Oliveira et al. 2018 [30]
<i>Hottentotta gentili</i> *	52	-	Grashof et al. 2019 [31]
<i>Androctonus mauretanicus</i> *	42	-	Grashof et al. 2019 [31]
<i>Babycurus gigas</i>	13	-	Grashof et al. 2019 [31]
<i>Grosphus grandideri</i>	22	-	Grashof et al. 2019 [31]

* Medically important species are noted with an asterisk (*).

Our combined dataset indicate that eastern scorpions possess similar sodium toxin gene variants as those in western populations, but may express sodium toxin genes differentially than western populations, especially the known pain-inducing toxin. This conclusion is supported with studies that indicate increased sodium toxin variability may be tied to more severe scorpionism and slight variations in sequence homology can lead to venoms with decreased toxicity [7,32].

4. Conclusions

We cloned and identified over 70 putative novel sodium β toxin genes from genomic DNA and showed greater toxin gene diversity may exist in western populations when compared to eastern populations. The RT-qPCR experiments with seven sodium β toxin variants and three reference genes showed sodium β toxin gene expression variation among individuals and indicated males exhibit more variable toxin gene expression than females. The transcriptomic analysis produced 111 toxin gene transcripts with the majority of transcripts identified as sodium β toxin genes. Lastly, the proteomic data suggest a toxin protein reduction from toxin transcript numbers to 24 different toxin proteins. Our analyses correspond to the growing evidence that toxicity in the scorpion genus *Centruroides* may be associated with toxin diversity, specific toxin gene regulation with local variation, and subsequent protein production. In addition, slight variation among toxin variants in *Centruroides* may contribute to the major variation in individual scorpion venom composition.

5. Materials and Methods

5.1. Population Sodium β Toxin Gene Variability Analysis

We sampled individuals in 11 populations identified from a *C. vittatus* phylogeographic analysis (Figure 1) [24]. Total genomic DNA from up to three scorpion's pedipalps and a portion of the carapace anterior was extracted with a standard Phenol-Chloroform extraction or with the FastID genomic DNA extraction kit (GeneticIDNA, Inc., Fairfield, IA, USA). After DNA isolation, each sample was further cleaned by Spermine precipita-

tion to optimize subsequent molecular analyses. Extracted genomic DNAs were stored in molecular biology grade water (Sigma Chemical Co., St. Louis, MO, USA) at $-20\text{ }^{\circ}\text{C}$ until PCR.

PCR amplification of sodium toxin genes was performed in 25 μL aliquots with the following ingredients: 10 μL total genomic DNA, 2X *Taq* Buffer (150 mM Tris-HCl pH 8.5, 40 mM $(\text{NH}_4)_2\text{SO}_4$, 3.0 mM MgCl_2 , 0.2% Tween 20), 1 mM dNTP, 0.5 μM for each primer, 6.25 units *REDTaq* DNA polymerase (Sigma Chemical Co.), 1.6% dimethyl sulfoxide, 0.6% BSA, and 1.6% formamide. The sodium toxin primers were taken from Corona et al. [27]: forward primer 5'-GAGATGAATTCGTTGTTGATGATYA-3' and reverse primer 5'-GCAATTAAGAAGCGTTACAATA-3'. The cycling conditions consisted of an initial denaturation period of five minutes at $94\text{ }^{\circ}\text{C}$ followed with 30 one-minute cycles of $94\text{ }^{\circ}\text{C}$, $50\text{ }^{\circ}\text{C}$ annealing, $72\text{ }^{\circ}\text{C}$ extension, and a final seven-minute extension at $72\text{ }^{\circ}\text{C}$. After PCR products were verified by electrophoresis in a 0.9% agarose gel, the products were GeneCleaned (MP Biomedicals, LLC., Irvine, CA, USA).

Cleaned PCR products were cloned into the pGEM-T vector (Promega, Inc., Madison, WI, USA) with plasmid transformation into *E. coli* JM109 and blue/white selection to identify recombinant clones. From each population, ten random white colonies were selected and archived in glycerol at $-80\text{ }^{\circ}\text{C}$ for long term storage. We amplified plasmid inserts from boiled colony samples by PCR with flanking SP6 and T7 Pro primers (Promega, Inc.). Inserts were identified by agarose gel electrophoresis. Cleaned PCR products were sent to the UAMS DNA Core Sequencing Facility for DNA sequencing with SP6 and T7 Pro primers on an Applied Biosystems 3100 Genetic Analyzer, using the Big Dye Terminator Chemistry, Kit version 1.1 (Foster City, CA, USA). After sequencing, all trace files were reviewed by eye and all ambiguous bases removed from further analysis. Alignment of the sequence data was conducted with Clustal X and Geneious Pro 3.7 [33,34]. After the initial alignment, all toxin DNA sequences were converted into their amino acid sequences to compare to known sodium toxin polypeptides. We named the sodium β toxin sequences with the nomenclature outlines in Romero-Gutierrez et al. [35]. All sequences were deposited in GenBank. Accession numbers are presented in Table S1 and Table S5.

5.2. Sodium β Toxin Gene Expression (RT-qPCR)

We collected male and female scorpions from a population in northwest Arkansas for RNA extraction. Scorpions for all subsequent work were collected from the same population. Scorpions were housed in the laboratory and fed juvenile crickets weekly until harvested for RT-qPCR analysis. Venom was not extracted from scorpions prior to harvesting, thus our samples represent a resting venom gland. Individuals were flash frozen at $-80\text{ }^{\circ}\text{C}$ with tissue from the telson (venom gland) and the pedipalp (body tissue) ground for RNA extraction with the AurumTM Total RNA Mini Kit (BioRad, Inc., Hercules, CA, USA). We determined RNA quality and concentration with a NanoDrop 2000/2000 c (Thermo Fisher Scientific, Inc., Waltham, MA, USA). In addition, RNA integrity was verified via 0.9% agarose gel electrophoresis. We constructed cDNA using an iScriptTM cDNA Synthesis Kit (BioRad, Inc.) with incubation parameters of $25\text{ }^{\circ}\text{C}$ for 5 min, $42\text{ }^{\circ}\text{C}$ for 30 min, and $85\text{ }^{\circ}\text{C}$ for 5 min. To verify subsequent RT-qPCR analyses, an initial PCR on the cDNA from both tissues was conducted with RT-qPCR sodium β gene toxin primers and reference gene primers. Cycling parameters were $94\text{ }^{\circ}\text{C}$ for 1 min, $50\text{ }^{\circ}\text{C}$ for 1 min, and then $72\text{ }^{\circ}\text{C}$ for 1 min for 30 cycles, with visualization on a 0.9% agarose gel.

The sodium β toxin gene primers for RT-qPCR were selected from the population sodium β toxin gene survey with six unique to *C. vittatus* (one unique to eastern populations, five in western populations, and *C. vittatus* alpha toxin IV4) and CsEv3b. Sodium β toxin PCR products were verified by sequencing. Reference primers were identified via a blast search of insect reference genes against the *Mesobuthus gibbous* annotated genome. The identified reference genes were then blasted against a *C. vittatus* draft genome assembly (unpublished) to develop reference RT-qPCR primers (Table S6).

PCR amplification efficiency: A standard curve was created with the following conditions: 5 μ L body tissue cDNA diluted to the 5th log in three replicates, 10 μ L of iTaq™ Universal SYBR® Green Supermix (BioRad, Inc.), 10 μ Mol of RPL (Ribosomal protein-L) or EF-1 (Elongation Factor 1) primers, and 4 μ L of Sigma-Aldrich molecular biology grade water were added for a total volume of 20 μ L per well, and run in a BioRad CFX96 Touch™ Real-Time PCR Detection System with cycling parameters of 95 °C for 30s 95 °C for 5s, 60 °C for 30s with a plate read for 40 cycles.

For Experimental Trials: cDNA from body and telson tissue were analyzed on the nanodrop for concentration and integrity then standardized to 100 ng/ μ L for a 5 μ L of cDNA for all wells except the non-template controls. To all wells, 10 μ L of SYBR® Green was added as well as 1 μ L of reference primers (10 μ Mol) to standard wells and NTCs (Non-Template Controls). One μ L (10 μ Mol) of experimental primer for experimental wells and NTCs, with molecular biology grade water for a final volume of 20 μ L to all tubes. Each plate had three replicates of standard and experimental wells and two replicates of NTC wells. The well plate was run in the CFX96 Touch™ Real-Time PCR machine under the same cycling parameters as the standard curve with an added dissociation step for all samples: 95 °C for 10s, 65 °C for 5s with a gradient of 5 to 95 °C. We removed individuals from the experimental dataset if reference and experimental genes showed substantial deviations across individual replicates. The final data included eight males and 13 females.

C_t data from triplicate samples were analyzed using the $\Delta\Delta C_t$ method [36]. The normalized relative expression values were log2 transformed and post hoc comparisons used one-way ANOVA and Tukey's pairwise multiple comparison tests.

5.3. Transcriptome

Two male and one female scorpion were collected in northwest Arkansas, fed crickets with visual conformation of prey envenomation, then after three days, harvested for telson and carapace transcriptome analysis. The scorpions were flash frozen at -80 °C and total RNA extracted with a Trizol preparation. RNA sample qc was analyzed through electrophoresis with an Aligent TapeStation system. RNA-seq with 50 bp reads was conducted at the University of Delaware on an Illumina genome sequencer (Illumina, Inc., San Diego, CA, USA). The data were viewed for initial quality through FastQC (v0.11.7), trimmed with Trimmomatic (v0.36), and normalization of the data was performed using Trinity (v2.5.1) [37–39]. Assembly of the normalized reads was then performed with the following de novo assembly programs: Trinity (v2.5.1), SOAPdenovo2 (v2.4.1), Velvet (v1.2.10), and TransAbyss (v1.5.4) resulting in 4 ea individual assemblies [39–42]. The transcriptome assemblies were then aggregated together using EviGene, which deals with the many redundancies, picks the best representatives, and filters out misassemblies [43,44]. The reads were mapped to the assembly using NGen and quantified as RPKM using ArrayStar [45]. RPKMs for contigs identified by BLASTn of the assembly for key genes were extracted for each sample over all contigs matching that BLAST query. In addition, the transcriptome assembly was blasted with a query database created from NCBI scorpion toxin and our current sodium toxin databases (2133 total toxin sequences). From these Blast searches, RPKM values for the two males and female was summarized for sodium toxin RNAs with additional searches for additional scorpion toxin RNAs. The transcriptome datasets were deposited in NCBI with the following IDs: TSA: GIPT01000000, SRA: SRR11917465, BioProject: PRJNA636371, BioSample: SAMN15075759.

5.4. Toxin Proteomics

Telsons were harvested from flash frozen individual scorpions collected and maintained for the RT-qPCR analysis (four females and three males collected in NW AR), washed with 95% ethanol, and ground in lysis buffer (2% SDS in 100 mM Tris/HCl, pH 7.6). Extracted proteins were verified via gel electrophoresis on a 15% SDS-PAGE gel and sent to the UAMS Proteomics core laboratory for MS/MS spectrophotometry. Purified proteins were reduced, alkylated, and digested using filter-aided sample preparation with sequenc-

ing grade modified porcine trypsin (Promega) [46]. Tryptic peptides were then separated by reverse-phase XSelect CSH C18 2.5 μm resin (Waters) on an in-line 150 \times 0.075 mm^2 column using an UltiMate 3000 RSLCnano system (Thermo). Peptides were eluted using a 90 min gradient from 97:3 to 60:40 buffer A:B ratio (Buffer A = 0.1% formic acid, 0.5% acetonitrile; Buffer B = 0.1% formic acid, 99.9% acetonitrile). Eluted peptides were ionized by electrospray (2.15 kV) followed by mass spectrometric analysis on an Orbitrap Fusion Lumos mass spectrometer (Thermo). MS data were acquired using the FTMS analyzer in profile mode at a resolution of 240,000 over a range of 375 to 1500 m/z . Following HCD activation, MS/MS data were acquired using the ion trap analyzer in centroid mode and normal mass range with precursor mass-dependent normalized collision energy between 28.0 and 31.0. Proteins were identified by database search using Mascot (Matrix Science) with a parent ion tolerance of 3 ppm and a fragment ion tolerance of 0.5 Da. Scaffold (Proteome Software) was used to verify MS/MS based peptide and protein identifications. Protein identifications were accepted if they could be established with less than 1.0% false discovery and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm [47]. We further analyzed the proteomic data with blast searches against a fasta database of 3139 scorpion toxin sequences created from an NCBI search and our own sodium toxin diversity analyses. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [1] partner repository with the following dataset identifiers: PXD019851 (10.6019/PXD019851) and PXD019850 (10.6019/PXD019850).

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/toxins13110828/s1>, Table S1: Putative *C. vittatus* Sodium Beta Toxin polypeptide sequence diversity from a survey of 10 population groups identified in a *C. vittatus* phylogeographic analysis, Table S2: Normalized relative expression results from eight sodium beta toxins and four reference genes, Table S3: Blast alignments with NCBI accession numbers from 2133 toxin gene sequences queried against a female *C. vittatus* transcriptome assembly, Table S4: Toxin proteins identified from 273 total proteins from a proteomics analyses for four females and three males, Table S5: *C. vittatus* collection sites (USA) and GenBank Accession numbers, Table S6: NCBI Accession data for rtPCR reference gene and rtPCR primer sequences.

Author Contributions: Conceptualization, T.Y.; data curation, J.F.P.; formal analysis, A.B., C.F., J.F.P., D.D.R. and T.Y.; funding acquisition, T.Y.; methodology, D.D.R.; software, J.F.P.; writing—original draft, T.Y.; writing—review and editing, J.F.P. and D.D.R. All authors have read and agreed to the published version of the manuscript.

Funding: This project was supported by the Arkansas INBRE program, with a grant from the National Institute of General Medical Sciences, (NIGMS), P20 GM103429 from the National Institutes of Health, USA.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review and BioSafety Committee of Arkansas Tech University (8 January 2015) and the Institutional Biosafety Committee of the University of Arkansas (IBC Protocol # 15019, 9 April 2015).

Informed Consent Statement: Not applicable.

Data Availability Statement: All toxin DNA sequences were deposited in GenBank. Accession numbers are presented in Table S1 and Table S5. The transcriptome datasets were deposited in NCBI with the following IDs: TSA: GIPT01000000, SRA: SRR11917465, BioProject: PRJNA636371, BioSample: SAMN15075759. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [1] partner repository with the following dataset identifiers: PXD019851 (10.6019/PXD019851) and PXD019850 (10.6019/PXD019850).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gantenbein, B.; Fet, V.; Barker, M.D. Mitochondrial DNA reveals a deep, divergent phylogeny in *Centruroides exilicauda* (Wood, 1863) (Scorpiones: Buthidae). In *Scorpions 2001. In Memoriam Gary A. Polis*; Fet, V., Seldon, P.A., Eds.; British Arachnology Society: Burnham Beeches, Bucks, UK, 2001; pp. 235–244.
2. Santibáñez-López, C.E.; Francke, O.F.; Ureta, C.; Possani, L.D. Scorpions from Mexico: From species diversity to venom complexity. *Toxins* **2016**, *8*, 2. [[CrossRef](#)] [[PubMed](#)]
3. Sissom, W.D. Life history. In *The Biology of the Scorpions*; Polis, G.A., Ed.; Stanford University Press: Stanford, CA, USA, 1990; pp. 161–223.
4. Teruel, R.; Fet, V.; Graham, M.R. The first mitochondrial DNA phylogeny of Cuban Buthidae (Scorpiones: Buthoidea). *Boletín Soc. Entomológica Aragonesa* **2006**, *39*, 219–226.
5. Borges, A.; Miranda, R.J.; Pascale, J.M. Scorpionism in Central America, with special reference to the case of Panama. *J. Ven. Ani. Tox. Trop. Dis.* **2012**, *18*, 130–143. [[CrossRef](#)]
6. Carcamo-Noriega, E.N.; Olamendi-Portugal, T.; Restano-Cassulini, R.; Rowe, A.; Uribe-Romero, S.J.; Becerril, B.; Possani, L.D. Intraspecific variation of *Centruroides sculpturatus* scorpion venom from two Regions of Arizona. *Arch. Biochem. Biophys.* **2018**, *638*, 52–57. [[CrossRef](#)] [[PubMed](#)]
7. Ward, M.J.; Ellsworth, S.A.; Hogan, M.P.; Nystrom, G.S.; Martinez, P.; Budhdeo, A.; Zelaya, R.; Perez, A.; Powell, B.; He, H.; et al. Female-biased population divergence in the venom of the Hentz striped scorpion (*Centruroides hentzi*). *Toxicon* **2018**, *152*, 137–149. [[CrossRef](#)] [[PubMed](#)]
8. Miller, D.W.; Jones, A.D.; Goldston, J.S.; Rowe, M.P.; Rowe, A.H. Sex differences in defensive behavior and venom of the striped bark scorpion *Centruroides vittatus* (Scorpiones: Buthidae). *Integ. Comp. Biol.* **2016**, *56*, 1022–1031. [[CrossRef](#)]
9. McElroy, T.; McReynolds, C.N.; Gullede, A.; Knight, K.R.; Smith, W.E.; Albrecht, E.A. Differential toxicity and venom gland gene expression in *Centruroides vittatus*. *PLoS ONE* **2017**, *12*, e0184695. [[CrossRef](#)]
10. 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. *Toxicon* **2018**, *142*, 14–29. [[CrossRef](#)]
11. Pedraza Escalona, M.; Possani, L.D. Scorpion beta-toxins and voltage-gated sodium channels: Interactions and effects. *Front. Biosci.* **2013**, *18*, 572–587. [[CrossRef](#)]
12. Gurevitz, M. Mapping of scorpion toxin receptor sites at voltage-gated sodium channels. *Toxicon* **2012**, *60*, 502–511. [[CrossRef](#)]
13. Cid-Urbe, J.I.; Meneses, E.P.; Batista, C.V.F.; Ortiz, E.; Possani, L.D. Dissecting toxicity: The venom gland transcriptome and the venom proteome of the highly venomous scorpion *Centruroides limpidus* (karsch, 1879). *Toxins* **2019**, *11*, 247. [[CrossRef](#)] [[PubMed](#)]
14. Stevens, M.; Peigneur, S.; Tytgat, J. Neurotoxins and their binding areas on voltage-gated sodium channels. *Front. Pharm.* **2011**, *2*, 1–13. [[CrossRef](#)] [[PubMed](#)]
15. Catterall, W.A. Voltage-gated sodium channels at 60: Structure, function and pathophysiology. *J. Physiol.* **2012**, *590*, 2577–2589. [[CrossRef](#)]
16. Gordon, D.; Gilles, N.; Bertrand, D.; Molgo, J.; Nicholson, G.M.; Sauviat, M.P.; Benoit, E.; Shichor, I.; Lotan, I.; Gurevitz, M.; et al. Scorpion toxins differentiating among neuronal sodium channel subtypes: Nature’s guide for design of selective drugs. In *Perspectives in Molecular Toxinology*; Menez, A., Ed.; John Wiley & Sons, Ltd.: New York, NY, USA, 2002; pp. 215–238.
17. Rendon-Anaya, M.; Delaye, L.; Possani, L.D.; Herrera-Estrella, A. Global Transcriptome Analysis of the Scorpion *Centruroides noxius*: New Toxin Families and Evolutionary Insights from an Ancestral Scorpion Species. *PLoS ONE* **2012**, *7*, e43331. [[CrossRef](#)]
18. Cohen, L.; Karbat, I.; Gilles, N.; Froy, O.; Ilan, N.; Benveniste, M.; Gordon, D.; Gurevitz, M. Common features in the functional surface of scorpion β -toxins and elements that confer specificity for insect and mammalian voltage-gated sodium channels. *J. Biol. Chem.* **2005**, *280*, 5045–5053. [[CrossRef](#)]
19. Cohen, L.; Lipstein, N.; Karbat, I.; Ilan, N.; Gilles, N.; Kahn, R.; Gordon, D.; Gurevitz, M. Miniaturization of Scorpion β -Toxins Uncovers a Putative Ancestral Surface of Interaction with Voltage-gated Sodium Channels. *J. Biol. Chem.* **2008**, *282*, 15169–15176. [[CrossRef](#)]
20. Del Rio-Portilla, F.; Hernandez-Marin, E.; Pimienta, G.; Coronas, F.V.; Zamudio, F.Z.; Rodriguez de la Vega, R.C.; Wanke, E.; Possani, L.D. NMR solution structure of Cn12, a novel peptide from the Mexican scorpion *Centruroides noxius* with a typical β -toxin sequence but with α -like physiological activity. *Eur. J. Biochem.* **2004**, *271*, 2504–2516. [[CrossRef](#)] [[PubMed](#)]
21. Kang, A.M.; Brooks, D.E. Nationwide Scorpion Exposures Reported to US Poison Control Centers from 2005 to 2015. *J. Med. Toxicol.* **2017**, *13*, 158–165. [[CrossRef](#)] [[PubMed](#)]
22. Rowe, A.H.; Rowe, M.P. Physiological resistance of grasshopper mice (*Onychomys* spp.) to Arizona bark scorpion (*Centruroides exilicauda*) venom. *Toxicon* **2008**, *52*, 597–605. [[CrossRef](#)]
23. Shelley, R.M.; Sissom, W.D. Distributions of the scorpions *Centruroides vittatus* (say) and *Centruroides hentzi* (banks) in the United States and Mexico (scorpiones, buthidae). *J. Arach.* **1995**, *23*, 100–110.
24. Yamashita, T.; Rhoads, D.R. Species delimitation and morphological divergence in the scorpion *Centruroides vittatus* (Say, 1821): Insights from phylogeography. *PLoS ONE* **2013**, *8*, e68282. [[CrossRef](#)]
25. Pucca, M.B.; Cerni, F.A.; Peigneur, S.; Bordon, K.C.; Tytgat, J.; Arantes, E.C. Revealing the Function and the Structural Model of Ts4: Insights into the “Non-Toxic” Toxin from *Tityus serrulatus* Venom. *Toxins* **2015**, *7*, 2534–2550. [[CrossRef](#)] [[PubMed](#)]

26. Yamashita, T. Surface activity, biomass, and phenology of the striped scorpion, *Centruroides vittatus* (Buthidae) in Arkansas, USA. *Euscorpius* **2004**, *17*, 25–33.
27. Corona, M.; Valdez-Cruz, N.A.; Merino, E.; Zurita, M.; Possani, L.D. Genes and peptides from the scorpion *centruroides sculpturatus* ewing, that recognize Na⁺-channels. *Toxicon* **2001**, *39*, 1893–1898. [[CrossRef](#)]
28. Valdez-Velázquez, L.L.; Quintero-Hernández, V.; Romero-Gutiérrez, M.T.; Coronas, F.I.V.; Possani, L.D. Mass Fingerprinting of the Venom and Transcriptome of Venom Gland of Scorpion *Centruroides tecomanus*. *PLoS ONE* **2013**, *8*, e66486. [[CrossRef](#)]
29. Valdez-Velázquez, L.L.; Cid-Uribe, J.; Romero-Gutiérrez, M.T.; Olamendi-Portugal, T.; Jimenez-Vargas, J.M.; Possani, L.D. Transcriptomic and proteomic analyses of the venom and venom glands of *Centruroides hirsutipalpus*, a dangerous scorpion from Mexico. *Toxicon* **2020**, *179*, 21–32. [[CrossRef](#)]
30. De Oliveira, U.C.; Nishiyama, M.Y., Jr.; Dos Santos, M.; Santos-da-Silva, A.P.; Chalkidis, H.M.; Souza-Imberg, A.; Candido, D.M.; Yamanouye, N.; Dorce, V.; Junqueira-de-Azevedo, I. Proteomic endorsed transcriptomic profiles of venom glands from *Tityus obscurus* and *T. serrulatus* scorpions. *PLoS ONE* **2018**, *13*, e0193739. [[CrossRef](#)]
31. Grashof, D.; Kerkkamp, H.; Afonso, S.; Archer, J.; Harris, D.J.; Richardson, M.K.; Vonk, F.J.; van der Meijden, A. Transcriptome annotation and characterization of novel toxins in six scorpion species. *BMC Genom.* **2019**, *20*, 645. [[CrossRef](#)] [[PubMed](#)]
32. Oliveira, F.N.; Mortari, M.R.; Carneiro, F.P.; Guerrero-Vargas, J.A.; Santos, D.M.; Pimenta, A.M.; Schwartz, E.F. Another record of significant regional variation in toxicity of *Tityus serrulatus* venom in Brazil: A step towards understanding the possible role of sodium channel modulators. *Toxicon* **2013**, *73*, 33–46. [[CrossRef](#)]
33. Thompson, J.D.; Gibson, T.J.; Plewniak, F.; Jeanmougin, F.; Higgins, D.G. The clustal_x windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **1997**, *25*, 4876–4882. [[CrossRef](#)]
34. Geneious v4.6. Available online: <http://www.geneious.com/> (accessed on 10 October 2016).
35. Romero-Gutiérrez, T.; Peguero-Sánchez, E.; Cevallos, M.A.; Batista, C.V.F.; Ortiz, E.; Possani, L.D. A Deeper Examination of *Thorellius atrox* Scorpion Venom Components with Omic Technologies. *Toxins* **2017**, *9*, 399. [[CrossRef](#)]
36. Kuang, J.; Yan, X.; Genders, A.J.; Granata, C.; Bishop, D.J. An overview of technical considerations when using quantitative real-time PCR analysis of gene expression in human exercise research. *PLoS ONE* **2018**, *13*, e0196438. [[CrossRef](#)] [[PubMed](#)]
37. Fastqc. Available online: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 15 July 2019).
38. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
39. Haas, B.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P.; Bowden, J.; Couger, M.; Eccles, D.; Li, B.; Lieber, M.; et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **2013**, *8*, 1494–1512. [[CrossRef](#)] [[PubMed](#)]
40. Robertson, G.; Schein, J.; Chiu, R.; Corbett, R.; Field, M.; Jackman, S.D.; Mungall, K.; Lee, S.; Okada, H.M.; Qian, J.Q.; et al. De novo assembly and analysis of RNA-seq data. *Nat. Methods* **2010**, *7*, 909–912. [[CrossRef](#)]
41. Schulz, M.H.; Zerbino, D.R.; Vingron, M.; Birney, E. Oases: Robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* **2012**, *28*, 1086–1092. [[CrossRef](#)]
42. Li, R.; Yu, C.; Li, Y.; Lam, T.W.; Yiu, S.M.; Kristiansen, K.; Wang, J. SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics* **2009**, *25*, 1966–1967. [[CrossRef](#)] [[PubMed](#)]
43. Gilbert, D.G. Genes of the pig, *Sus scrofa*, reconstructed with EvidentialGene. *PeerJ* **2019**, *7*, e6374. [[CrossRef](#)] [[PubMed](#)]
44. NCGAS/de-Novo-Transcriptome-Assembly-Pipeline (Version 4.0.1). Available online: <https://zenodo.org/record/3945505#.YZthaboRVPZ> (accessed on 15 July 2019).
45. Shi, Y.; Shang, J. Long Noncoding RNA Expression Profiling Using Arraystar LncRNA Microarrays. *Methods Mol. Biol.* **2016**, *1402*, 43–61. [[CrossRef](#)] [[PubMed](#)]
46. Wiśniewski, J.R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6*, 359–362. [[CrossRef](#)] [[PubMed](#)]
47. Nesvizhskii, A.I.; Keller, A.; Kolker, E.; Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **2003**, *75*, 4646–4658. [[CrossRef](#)] [[PubMed](#)]