Toxicon 152 (2018) 137–149

Contents lists available at ScienceDirect

Toxicon

journal homepage: www.elsevier.com/locate/toxicon

Female-biased population divergence in the venom of the Hentz striped scorpion (*Centruroides hentzi*)

Micaiah J. Ward[a], Schyler A. Ellsworth[a], Michael P. Hogan[a], Gunnar S. Nystrom[a], Paul Martinez[a], Amisha Budhdeo[a], Roxana Zelaya[a], Alexander Perez[a], Barclay Powell[a], Huan He[b], Darin R. Rokyta[a,*]

[a] Department of Biological Science, Florida State University, Tallahassee, FL, 32306, USA
[b] Institute of Molecular Biophysics and College of Medicine, Florida State University, Tallahassee, FL, 32306, USA

ARTICLE INFO

Keywords:
Scorpion
Venom
Sex-biased genes
Population variation

ABSTRACT

Sex-biased genes are expressed at higher levels in one sex and contribute to phenotypic differences between males and females, as well as overall phenotypic variation within and among populations. Venom has evolved primarily for predation and defense, making venom expression a highly variable phenotype as a result of local adaptation. Several scorpion species have shown both intraspecific and intersexual venom variation, and males have been observed using venom in courtship and mating, suggesting the existence of venom-specific, sex-biased genes that may contribute to population divergence. We used reversed-phase high-performance liquid chromatography (RP-HPLC), Agilent protein bioanalyzer chips, nano-liquid chromatography mass spectrometry (nLC/MS/MS), and median lethal dose (LD_{50}) assays in fruit flies (*Drosophila melanogaster*) and banded crickets (*Gryllodes sigillatus*) to investigate proteomic and functional venom variation within and among three Florida populations of the Hentz striped scorpion (*Centruroides hentzi*). We found significant venom variation among populations, with females, not males, being responsible for this divergence. We also found significant variation in venom expression within populations, with males contributing more to within population variation than females. Our results provide evidence that male and female scorpions experience different natural and sexual selective pressures that have led to the expression of sex-biased venom genes and that these genes may be consequential in population divergence.

1. Introduction

Among metazoans, males and females within a species often exhibit unmistakable phenotypic differences. These sexually dimorphic traits are usually the result of sex-biased gene expression and include not only sex-specific reproductive tissues, but also differences in size, coloring, and behavior (Ellegren and Parsch, 2007; Mank, 2008). Sex-biased gene expression is thought to be especially prevalent in species that experience differences in maternal (or paternal) care, as well as male-male mate competition, sperm competition, female mate choice, or differences in ecology (Shine, 1989; Ellegren and Parsch, 2007; Mank, 2008), as these traits are under the strong evolutionary forces of natural and sexual selection (Chenoweth et al., 2007). Sex-biased genes may be exclusively expressed in one of the two sexes (sex-specific genes), or be expressed at higher levels in one sex compared to the other (sex-enriched genes), and are not limited to genes located on sex chromosomes (Ellegren and Parsch, 2007). The most extensive studies of sex-biased genes have included *Drosophila* (Parisi et al., 2003; Ranz et al., 2003; Zhang et al., 2004; Connallon and Knowles, 2005; Pröschel et al., 2006), mice (Torgerson et al., 2002; Yang et al., 2006), and comparisons of chimps and humans (Khaitovich et al., 2005; Nielsen et al., 2005), all of which found that male-biased genes exhibit a greater amount of rapid, positive selection and are therefore more responsible for population and/or species divergence. Most of these rapidly evolving male-biased genes, however, are primarily expressed in reproductive tissues (Ellegren and Parsch, 2007), and the degree of male or female-biased gene expression may differ among tissue types (Yang et al., 2006) or species (Ranz et al., 2003; Mank and Ellegren, 2009).

Scorpions are often sexually dimorphic in appearance and exhibit many traits associated with sex-biased gene expression, including maternal care, mate competition, female mate choice, ecology, and even differences in venom composition and lethality (De Sousa et al., 2010; Rodríguez-Ravelo et al., 2015; Miller et al., 2016; Cid-Uribe et al., 2017), and such differences seem to be widespread among arachnids

* Corresponding author. Florida State University, Department of Biological Science, 319 Stadium Dr., Tallahassee, FL, 32306-4295, USA.
E-mail address: drokyta@bio.fsu.edu (D.R. Rokyta).

https://doi.org/10.1016/j.toxicon.2018.07.026
Received 4 May 2018; Received in revised form 27 June 2018; Accepted 31 July 2018
Available online 07 August 2018
0041-0101/

(Herzig et al., 2002, 2008; Zobel-Thropp et al., 2018). Venom consists of proteins and peptides that, in most venomous animals, have evolved primarily for the purposes of predation and defense (Biardi et al., 2005, 2011). Intraspecific venom variation, which has been established in multiple scorpion species (Abdel-Rahman et al., 2009; Ruiming et al., 2010; Rodríguez-Ravelo et al., 2013; Estrada-Gómez et al., 2014; Carcamo-Noriega et al., 2017; Schaffrath et al., 2018), is thought to be a result of local adaptation, especially when differences in prey availability or predator exposure exist among populations. Some scorpion species have also been observed using venom in hygiene (D'suze et al., 2015) and mating behavior (Polis and Sissom, 1990; Lourenço, 2000; Benton, 2001), where male-male mate competition (Benton, 1992) and female mate choice (Polis and Sissom, 1990; Tallarovic et al., 2000; Contreras-Garduño et al., 2006) have been documented. Male scorpions are often faster and more mobile in search for mates (Booncham et al., 2007; Carlson and Rowe, 2009; Kaltsas and Mylonas, 2010), and females are more sedentary because they are frequently gravid or carrying young on their backs, making long-distance travel more difficult (Shaffer and Formanowicz, 1996). The differences in ecological niches between males and females has resulted in females being more apt to sting defensively than males (Williams, 1987; Shaffer and Formanowicz, 1996; Carlson and Rowe, 2009; Carlson et al., 2014). Males are prone to greater predator exposure (Polis and Farley, 1979) and are likely to consume a broad range of prey types as they travel. The differences between male and female scorpions in morphology, behavior, environment, and venom composition suggest not only that sex-biased genes are present in scorpions, but that some sex-biased genes may be venom-specific and contribute to venom variation within and among populations.

To investigate intraspecific and intrapopulation venom variation as well as explore the potential contribution of venom-specific, sex-biased gene expression, we sampled male and female Hentz striped scorpions, *Centruroides hentzi*, from three Florida populations. *Centruroides hentzi* are sexually dimorphic scorpions found in the Southeastern United States. Males have smaller bodies with longer metasomal (tail) segments and females are larger with shorter, more rounded metasomal segments (Fig. 1). *Centruroides hentzi* are commonly found in tree bark of long leaf pines, such as those found in the Apalachicola National Forest (ANF), Osceola National Forest (ONF), and the Tosohatchee Wildlife Management Area (TSO) (Fig. 1). We used a combination of reversed-phase high-performance liquid chromatography (RP-HPLC), Agilent protein bioanalyzer chips, and nano-liquid chromatography mass spectrometry (nLC/MS/MS) to determine overall venom expression variation within and among populations, as well as between sexes within and among populations. To assess the correlation between molecular variation and venom function, we performed median lethal-dose (LD_{50}) assays in banded crickets (*Gryllodes sigillatus*) and fruit flies (*Drosophila melanogaster*).

2. Materials and methods

2.1. Animal and venom collection

Adult scorpions were collected from the ANF (n_{male} = 4; n_{female} = 8), ONF (n_{male} = 4; nf_{emale} = 7), and TSO (n_{male} = 5; n_{female} = 5). Scorpions were fed banded crickets 1–2 times per week, unless they were being prepared for venom collection. Venom was collected using methods previously described (Rokyta and Ward, 2017; Ward et al., 2018). Briefly, scorpions were starved for a minimum of seven days to ensure ample venom production and anesthetized with CO_2. Electrostimulation was performed using a transcutaneous electrical nerve stimulation (TENS) unit (9 V) on a medium setting and applying a positive and negative electrode to either side of the telson to induce a muscle contraction. Venom was collected on a sterile metal spatula and pipetted into tubes containing LC/MS water. Venom samples were quickly centrifuged (two minutes at 12,000 RPM), frozen, lyophilized, and stored at −80°C until later use. Each venom sample was quantified using the Qubit Protein Assay Kit (Thermo Fisher Scientific) prior to use in our analyses. The quantities of venom present in each sample used are provided in Table S1.

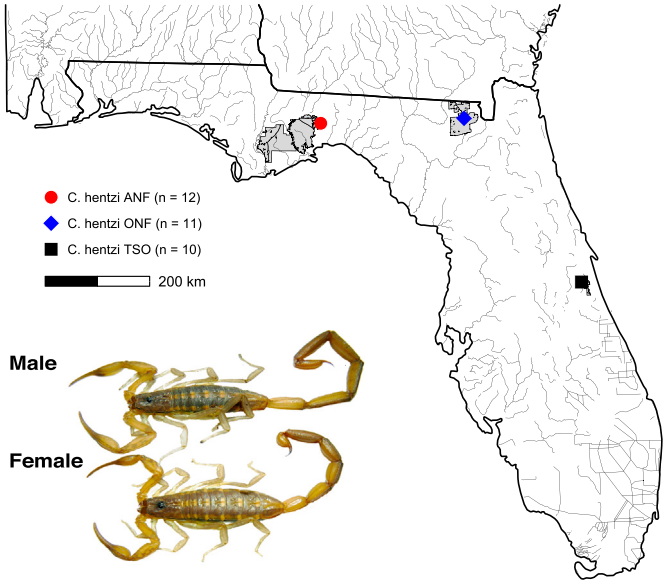

Fig. 1. Representative male and female *Centruroides hentzi* and Florida collection site map. Males have smaller bodies with longer metasomal segments and females have larger bodies with shorter, more rounded, metasomal segments. Gray areas indicate National Forest boundaries. The ANF and ONF forests are separated by roughly 257 km and the Suwannee river (SW), representing a substantial geographical boundary between the two populations. The ONF and TSO forests are separated by roughly 290 km and the ANF and TSO are separated by just over 482 km.

Banded crickets (*Gryllodes sigillatus*, size: 1.9 cm) were purchased from Ghann's Cricket Farm in Augusta, GA. Crickets were fed Ghann's Cricket Chow and housed at room temperature in a large aquarium cage with egg crates. Fruit flies (*Drosophila melanogaster*) used in this study were a single inbred line (line 11057) provided by the *Drosophila* Synthetic Population Resource (DSPR). Flies were kept on a light/dark cycle of 12 h held at a constant temperature of 25°C and fed standard cornmeal-agar media.

2.2. Reversed-phase high-performance liquid chromatography

Three venom samples for each *C. hentzi* individual were quantified using reversed-phase high-performance liquid chromatography (RP-HPLC). RP-HPLC was performed on a Waters 2695 Separations Module with a Waters 2487 Dual λ Absorbance Detector. Approximately 7 μg of protein was injected onto a Jupiter 5 μm C18 column (Phenomenex, Torrance, CA) using the standard solvent system of A = 0.1% trifluoroacetic acid (TFA) in water and B = 0.075% TFA in acetonitrile. A 125-min gradient from 10 to 75% solution B was performed at 0.2 mL/min, followed by 15 min at 10% solution B to wash the column. Data was analyzed using Empower Pro software version 5.00 (Waters Corporation, Milford, MA). Twenty-four RP-HPLC peaks were identified and quantified using methods previously described (Margres et al., 2015a; b). Briefly, the area under each peak was measured to determine the relative abundance of each peak to the total area of all protein peaks identified. According to the Lambert-Beer law, this relative abundance is related to the total percentage of peptide bonds in the sample (Gold et al., 1987) and has been shown to be representative of the relative amount of a specific protein by weight (Gibbs et al., 2009).

2.3. Protein bioanalyzer electrophoresis

A single venom sample from each *C. hentzi* individual was processed using the Agilent protein 80 assay (Santa Clara, CA). Venom samples used for this analysis were one of the three samples used in RP-HPLC and were chosen based on sample quantity. Venom samples were processed and analyzed as previously described (Ward et al., 2018) using approximately 4–6 μg of crude venom per individual and following manufacturer's instructions. The Agilent 2100 Expert software (version B.02.09 (SR1), Santa Clara, CA) was used to calculate the relative abundance and concentration of each profile peak.

2.4. Protein identification

After the most abundant variable RP-HPLC peak was identified (peak 12, see statistical analysis methods and results sections), RP-HPLC was used (following methods described above) to collect fractions of this peak from one female *C. hentzi* sample from each of the three populations. Samples were selected based on their high abundance of peak 12 to ensure enough protein would be present in each fraction for mass spectrometry (MS) analysis. Each collected fraction was spun in a SpeedVac for approximately 75 min to dry down and concentrate the sample. Both top-down (intact, reductive alkylated protein) and bottom-up (trypsin digested protein) nano-liquid chromotography mass spectrometry (nLC/MS/MS) methods were used to identify the proteins present in the collected fractions. nLC/MS/MS analyses were performed by the Florida State University College of Medicine Translational Lab. For top-down nLC/MS/MS, dried protein samples were reductive alkylated with a ProteoExtract all-in-one trypsin digestion kit (Calbiochem, Billerica, MA). Proteins were dissolved in 25 μL of digest buffer and 1 μL of the reducing agent was added. The mixture was then incubated in a thermal shaker (Thermal Shake Touch, VWR, Radnor, PA) at 37 °C and 1200 rpm for 10 min. After cooling to room temperature, 1 μL of blocking agent was added to the mixture and incubated at room temperature for 10 min. The mixture was then dried using a SpeedVac and the dried mixture was dissolved in 0.1% formic acid aqueous solution prior to nLC/MS/MS injection. For bottom-up MS, 1 μL of trypsin was added to an aliquot of the reductive alkylated protein and incubated in a thermal shaker at 37 °C and 1200 rpm for 2 h. Following this step, 25 μL of 5% formic acid was added and the mixture was dried in the SpeedVac. The dried mixture was then dissolved in 0.1% formic acid aqueous solution prior to nLC/MS/MS injection.

Dissolved protein samples were injected to a nLC (Dionex UltiMate 3000 UHPLC, Thermo Fisher Scientific, Waltham, MA) on-line coupled to a Q Exactive HF Hybrid Quadrupole-Orbitrap Mass Spectrometer MS (Thermo Fisher Scientific, Waltham, MA). A 300 μm × 5 mm trap column (C18 PepMap100, 5 μm, 100 Å, Thermo Scientific) was used as the pre-column, followed by a 75 μm × 15 cm C18 analytical column (Acclaim PepMap RSLC, 2 μm, 100 Å, Thermo Scientific). The mobile phases used were A = 99.9% H_2O and 0.1% formic acid and B = 99.9% acetonitrile and 0.1% formic acid. A linear gradient from 3% to 55% B over 1.5 h was performed with a flow rate of 300 nL/min. Eluate was on-line ionized by electrospray ionization (ESI) at positive ion mode and detected by the MS. Precursor ions were detected with a mass resolving power of 60 K (at *m/z* of 800 Da) and 3 million target ion number. For top-down analysis, the top 10 most abundant precursor ions were selected for data-dependent MS^2 by collisional induced dissociation (CID) at 27 normalized collision energy (NCE) with a mass resolving power of 15 K (at *m/z* of 800 Da) and 0.2 million target ion number. The collected raw files were viewed and analyzed by Xcalibur software (Thermo Fisher Scientific, Waltham, MA). MS signals in the chromatography region of intact proteins were averaged and compared among samples. For bottom-up analysis, the top 10 most abundant precursor ions were selected for data-dependent MS^2 by collisional induced dissociation (CID) at 27 normalized collision energy (NCE) with a mass resolving power of 15 K (at *m/z* of 800 Da) and 0.1 million target ion number. To identify the proteins present in each collected fraction of peak 12 (one from each population), collected. raw files were analyzed by Proteome Discoverer version 1.4 (Thermo Fisher Scientific, Waltham, MA) and searched against the *C. hentzi* venom protein database (Ward et al., 2018) with Cysteine carbamidomethylation, Methionine oxidation, and C-terminus amidation as dynamic modification.

2.5. Toxicity assays

One pooled venom sample for each population (ANF, ONF and TSO) was suspended in sterile phosphate-buffered saline (PBS) and used to perform median lethal dose (LD_{50}) assays in both *G. sigillatus* and *D. melanogaster*. To compare functional differences in toxicity between male and female *C. hentzi* venom, an all-male pooled venom sample and an all-female pooled venom sample from only the ANF population were used to perform LD_{50} assays in *G. sigillatus*.

2.5.1. LD_{50} assays in Gryllodes sigillatus

Adult *G. sigillatus* were divided by weight class (0.1–0.19, 0.2–0.29, and 0.3–0.39 g ± 0.01 g) and separated into groups of 10 to ensure equivalent venom dose by weight per group. Each group of 10 crickets was given one experimental dose with an injected volume of 5 μL per cricket. In some cases, higher doses were required to determine the LD_{50} value because lower doses did not have a comparable effect. The following doses were used: ANF: 0.005, 0.01, 0.1, 0.2, 0.35, 0.6, 0.8 and 1.0 μg/g (venom/cricket body weight); ONF: 0.005, 0.01, 0.1, 0.2, 0.35, 0.6, 0.8 and 1.0 μg/g; TSO: 0.005, 0.01, 0.2, 0.5, 0.8, 1.3, 2.0 and 2.7 μg/g; ANF Male: 0.01, 0.2, 0.5, 0.8, 1.3, 2.0 and 2.7 μg/g; ANF Female: 0.01, 0.2, 0.5, 0.8, 1.3, 2.0 and 2.7 μg/g. Control groups (10 crickets per assay) were injected with 5 μL of sterile PBS. Crickets were injected by methods previously described (Herzig and Hodgson, 2009), using a syringe equipped with a 29G x 1/2" needle into the ventrolateral thoracic region between legs 2 and 3. Prior to performing the injection, crickets were briefly anesthetized with CO_2 for 10–15 s. After injection, crickets were placed back into their respective containers with cricket food and water-gel globules. Crickets were observed at 2, 4, 6, 8, 12 and 24 h time points after injection and counted as either alive (normal behavior) or dead (little to no movement when perturbed). If any controls died over the course of the 24 h observation period, data analysis was performed on the last time point in which all controls were counted as alive.

2.5.2. LD_{50} assays in Drosophila melanogaster

Drosophila melanogaster (DSPR inbred line 11057, weighing 0.9 mg ± 0.1 mg) were separated into groups of 10 (in equal sex ratios) in vials with standard cornmeal-agar fly media. Each group of 10 flies was given one experimental dose with an injected volume of 18.4 nL per fly. The injected doses were: 0.1, 0.3, 0.5, 0.7, 1.0, 1.3, 2.0 and 2.8 ng/mg (venom/fly body weight). Control groups (10 flies per assay) were injected with 18.4 nL of sterile PBS. Flies were injected into the ventral-lateral portion of their abdomen using a Drummond Nanoject II Auto-Nanoliter Injector (Drummond Scientific Company, Broomall, PA) with pulled glass capillary needles, while being anesthetized on a CO_2 plate (≤ 3 L/min flow rate). Flies were not exposed to CO_2 for more than three minutes. After injection, flies were placed back into vials containing standard cornmeal-agar fly media and allowed to recover. During recovery, vials were placed on their side to avoid flies getting stuck in the media. Flies were observed at 2, 4, 6, 8, 12 and 24 h time points after injection and counted as either alive (normal behavior) or dead (little to no movement when perturbed). If any controls died over the course of the 24 h observation period, data analysis was performed on the last time point in which all controls were counted as alive.

2.6. Data analysis

2.6.1. Data sets and transformations

The data set containing 99 C. *hentzi* venom profiles (ANF $n = 36$, ONF $n = 33$, TSO $n = 30$) is referred to as the complete RP-HPLC data set. However, using the complete RP-HPLC data set was not appropriate for among-population statistical analyses due to potential effects of pseudoreplication (venom samples from the same individual are not statistically independent; Hurlbert, 1984). Therefore, one venom sample from each individual in each population was randomly chosen from the complete RP-HPLC data, resulting in a total of 33 venom profiles (ANF $n = 12$, ONF $n = 11$, TSO $n = 10$), and is referred to as the partial RP-HPLC data set. The protein bioanalyzer data contained one venom sample per individual from each *C. hentzi* population (ANF $n = 12$, ONF $n = 11$, TSO $n = 10$).

Analyses of both RP-HPLC (24 peaks) and protein bioanalyzer (11 peaks) data sets were performed following methods previously described (Margres et al., 2015a; b). Because some peaks were non-quantifiable (giving a zero value), the raw RP-HPLC and protein bioanalyzer peak data was first made compositional based on Bayesian-multiplicative replacement methods using the zCompositions package in R (Palarea-Albaladejo and Martín-Fernández, 2015), which replaces zero values with trace values that represent a non-quantifiable peak as a result of the accuracy of the measurement process and assigns the peak a positive value that is less than the smallest recorded value (*i.e.* 0.03% in our RP-HPLC data set and 0.1% in our protein bioanalyzer data set) (Aitchison, 1982). This compositional data, however, is subject to constant-sum constraints and components are biased toward negative correlation (Aitchison, 1982). The compositional data were therefore transformed to centered logratio (clr) and isometric logratio (ilr) data sets using the robCompositions package in R prior to statistical analyses (Egozcue et al., 2003; Filzmoser et al., 2009; Templ et al., 2011). As discussed by Margres et al. (2015a), the clr data retains individual peak identity, but still suffers from sum-constraints, so this data set was used for visualization purposes only. The ilr data set does not suffer from sum-constraints, so the ilr data set was used in all statistical analyses.

2.6.2. Estimating sample-to-population placement probabilities

Using ilr transformed data, a discriminant function analysis (DFA) was performed using the linear discriminant analysis (lda) function of the MASS package in R (Venables and Ripley, 2002) on both the complete RP-HPLC data set and the protein bioanalyzer data set to calculate the placement probabilities of each sample in its designated population (*e.g.,* the probability that an individual from ANF belongs to the ANF population based on its RP-HPLC or protein bioanalyzer venom profile).

2.6.3. Testing for compositional venom variation among populations

To test for variation among populations, nonparametric (permutational) MANOVA statistical tests were performed using the adonis function from the vegan package in R (Oksanen et al., 2013), which uses distance matrices to calculate variance, with linear predictors (independent factors) and a multiple-column response matrix (dependent factors) as inputs. In all statistical models used, the euclidean method was used for distance matrix calculations, the ilr transformed peak values (of either the RP-HPLC or the protein bioanalyzer data sets) represented the response matrix of dependent factors, Y, and P values were calculated under 1,000,000 permutations.

Both the partial RP-HPLC and the protein bioanalyzer data sets were used to test for variation among populations. In this model, population and sex were used as independent factors ($Y \sim Population*Sex$), to assess the overall variation among populations, between sexes, and whether these two factors interacted. If significance was detected ($P \leq 0.05$), a series of *post-hoc* pairwise tests were performed to determine the sources of variation. The *post-hoc* tests included pairwise tests between populations (three tests), as well as total and pairwise tests between populations using only females (four tests) and only males (four tests), resulting in a total of 11 *post-hoc* tests. To avoid type I error, the Bonferroni correction for multiple comparisons was applied to all *post-hoc* tests (α/total number of *post-hoc* tests), which is equal to a significance threshold of $P = 0.05/11 = 0.0045$. *Post-hoc* P-values were multiplied by the number of *post-hoc* tests and were reported as corrected values ($P\times$total number of *post-hoc* tests $= P^*$).

2.6.4. Testing for compositional venom variation within populations

To test for variation within populations, nonparametric (permutational) MANOVA statistics were performed as described in section 2.6.3. Unlike the among-population model, the within-population model used the complete RP-HPLC data set and accounted for nestedness, with animals nested within sex, and sex nested within population ($Y \sim Population/Sex/Animal$). In our initial test, permutations were constrained within each population by using the strata option in the adonis function. If significance was detected ($P \leq 0.05$), a series of *post-hoc* tests were performed within each of the three populations to determine the sources of variation. The *post-hoc* tests within each population included testing for sex differences within each population, with animal as a nested factor within sex (three tests), as well as testing for variation among animals within each sex (six tests), resulting in a total of nine *post-hoc* tests. Using the same Bonferroni correction method as previously described, the significance threshold used for the within-population *post-hoc* tests was $P = 0.0056$. *Post-hoc* P-values were multiplied by the number of *post-hoc* tests and were reported as corrected values ($P\times$total number of *post-hoc* tests $= P^*$).

2.6.5. Testing for within-individual venom variation

To assess potential variation of replicate samples from each animal, we calculated the timing of each venom sample as the number of days from the date the animal was brought into the lab to the date of venom extraction (Table S1). When possible, venom was extracted immediately upon acquisition, and venom extractions occurred regularly (approximately 1–2 times per month) throughout the life of the animal, with no guarantee of obtaining usable venom quantities from each animal with each attempted extraction. Therefore, the timing of venom samples used in this study ranged from 1 to 102 days. Plasticity in venom has previously been reported in scorpions (Gangur et al., 2017) and differences between the first venom extraction and samples collected 24 h after the first extraction have also been observed in some species (Schaffrath and Predel, 2014; Schaffrath et al., 2018). We tested for sample variation within individuals by running a nonparametric MANOVA on the complete RP-HPLC data set, using time as a nested factor within animal and constraining permutations within each animal ($Y \sim Animal/Time$, with strata on $Animal$). Additionally, we performed a robustness check to determine if choosing a sample at random from each individual (for the partial RP-HPLC data set) created any bias in our results by ignoring potential effects of plasticity or using the initial venom sample collected from an individual. To do this, we performed the same nonparametric MANOVA tests on an ilr transformed RP-HPLC data set containing only the earliest venom sample from each animal and another containing only the latest venom sample from each animal. Should either of these differ from each other or from the random sample data set, we could not rule out effects of within-individual variation in our observed variation results among populations.

2.6.6. Identifying the most variable venom components

To identify and visualize the individual peaks that contributed most to either the geographical or intersexual variation in each of the clr transformed complete RP-HPLC and protein bioanalyzer data sets, a robust principal component analysis was performed using the robCompositions R package (Templ et al., 2011) following methods described by Filzmoser et al. (2009). To determine the percent variance of each peak among populations, within populations, and between sexes, variance matrices were created using these same data sets, with

the robCompositions package in R.

2.6.7. Assessing functional variation using toxicity assays

LD_{50} values were calculated using the *tsk* package in R (Stone, 2015), which employs the Trimmed Spearman-Karber method (Hamilton et al., 1977) for calculating median effective doses from dose-response curves. This method calculates a 95% confidence interval on the LD_{50} values, but does not allow for appropriate statistical comparison between calculated values (Schenker and Gentleman, 2001). For that reason, a separate, independent dose-response value (ED_{50}) was calculated for each assay using the *drc* package in R (Ritz et al., 2015), following methods previously described (Margres et al., 2017). In brief, a dose-response model was first selected by utilizing Akaike's information criterion (Akaike, 2011) to determine the best fit model (*i.e.* lowest standard error). Both log-logistic (2–5 parameters) and Weibull (1–4 parameters) models were tested and the models with the lowest overall standard error for each cricket and fly assays were then used within the *drm* function to calculate the ED_{50}. To determine significant differences in function among ANF, ONF, and TSO venoms, as well as between male and female venoms from the ANF population, a ratio test (Wheeler et al., 2006) was implemented using the *comped* function. In this test, the ratio of ED_{50}s are calculated and determined to be significantly different if the confidence intervals for the ratios do not contain 1.

3. Results

3.1. Venom variation among populations

3.1.1. Sample-to-population placement probabilities were high

Discriminant function analysis (DFA) of the complete RP-HPLC data assigned a placement probability of 78% for ANF, 82% for ONF, and 97% for TSO, for individuals designated as being from their respective populations, with an overall average placement probability of 85%. This suggests that an individual could be assigned to its population with a 85% certainty based on its RP-HPLC venom profile alone and indicates that each population may be distinguished by their venom. The DFA of the protein bioanalyzer data assigned a placement probability of 67% for ANF, 55% for ONF and 80% for TSO, with an overall average placement probability of 67%. The RP-HPLC venom profiles have 24 distinct peak regions (Fig. 2A), whereas the protein bioanalyzer profiles have only 11 distinct peak regions (Fig. 2B), suggesting that the difference in placement probabilities between the two profiles types is due to the RP-HPLC profiles having more defined peak separation in regions that likely contribute to variation among the three populations, or that technical noise may be less of a factor in RP-HPLC profiles versus protein bioanalyzer profiles.

3.1.2. Geographic and intersexual variation contributed to overall venom variation among populations

We performed a nonparametric MANOVA on the partial RP-HPLC ilr transformed data set (one random venom sample from each animal) and found significant variation among populations ($F = 3.70$, $R^2 = 0.19$, $P = 2.2 \times 10^{-5}$) and between sexes ($F = 3.07$, $R^2 = 0.08$, $P = 4.3 \times 10^{-3}$), but the interaction between population and sex was not significant ($F = 1.23$, $R^2 = 0.06$, $P = 0.23$, Table 1). We then conducted a series of *post-hoc*, pairwise nonparametric MANOVA tests, and adjusted resulting *P* values with a Bonferroni correction, indicated by *P**, to determine the sources of variation (Table S2). We found no significant differences between the ANF and ONF populations. When comparing ONF and TSO populations, we found significant variation between populations ($F = 5.86$, $R^2 = 0.22$, $P^* = 5.5 \times 10^{-5}$) but not between sexes ($F = 2.29$, $R^2 = 0.09$, $P^* = 0.33$), and when comparing ANF and TSO, we found significant variation between populations ($F = 4.32$, $R^2 = 0.16$, $P^* = 7.4 \times 10^{-3}$) and between sexes ($F = 3.35$, $R^2 = 0.12$, $P^* = 0.046$).

We performed the same nonparametric MANOVA tests on the protein bioanalyzer ilr transformed data set (one venom sample per animal) and found similar results to the RP-HPLC data in terms of overall variation among populations ($F = 4.85$, $R^2 = 0.24$, $P = 3.6 \times 10^{-5}$), but not intersexual variation ($F = 1.53$, $R^2 = 0.04$, $P = 0.17$), and no significant interaction between population and sex ($F = 1.10$, $R^2 = 0.05$, $P = 0.34$, Table 1). In our *post-hoc* analyses using Bonferroni corrected *P* values, indicated by *P**, we found significant differences between ONF and TSO populations ($F = 4.43$, $R^2 = 0.19$, $P^* = 0.043$) and between ANF and TSO populations ($F = 8.11$, $R^2 = 0.29$, $P^* = 3.6 \times 10^{-4}$), and no difference between ANF and ONF populations, as was the case with the partial RP-HPLC data set. We did not, however, find any significant differences between the sexes in our pairwise population comparisons.

3.1.3. Venom variation among populations was female-biased

Using only female venom samples from the partial RP-HPLC data set, we found overall significant variation among populations ($F = 3.32$, $R^2 = 0.28$, $P^* = 1.8 \times 10^{-4}$), which was due to differences between females of ONF and TSO ($F = 4.81$, $R^2 = 0.32$, $P^* = 0.014$) and ANF and TSO ($F = 4.55$, $R^2 = 0.29$, $P^* = 8.7 \times 10^{-3}$), but no significant difference was found between ANF and ONF females. When using only male venom samples from the partial RP-HPLC data set, we found no significant venom variation overall or between pairs of the three populations.

Because we did not find significant differences between the sexes in our pairwise population comparisons using the protein bioanalyzer data set, we did not expect to find significant variation among populations overall using only females ($F = 2.94$, $R^2 = 0.26$, $P^* = 0.054$), but we did find significant differences between ANF and TSO females ($F = 4.80$, $R^2 = 0.30$, $P^* = 0.044$). Similar to the RP-HPLC analysis, we found no significant variation overall or between any of the three populations when using only male venom samples from the protein bioanalyzer data set.

3.1.4. Within-individual variation did not impact the among-population results

We performed a robustness check to determine if plasticity (Gangur et al., 2017) or individual sample variation affected our among population venom variation results. We observed the same patterns of variation among populations (overall and sex-based) when using both the earliest and latest extracted venom sample from each animal (data not shown). These results suggest that any potential changes in venom phenotype due to plasticity or individual variation in venom expression had no effect on our results.

3.1.5. Both low- and high-expression venom components contributed to variation among populations

The PCA analysis of the RP-HPLC peaks revealed distinct separation between TSO and the other two populations, and more overlap of the ANF and ONF populations. (Fig. 3A). The first five principal components explained 73.8% of the variation in the RP-HPLC data set making up 30.9%, 15.3%, 11.7%, 9.4% and 6.6%, respectively. The most variable peaks in the PCA analysis were: peaks 1, 11, 12, 17, and 24. The PCA analysis generated from the protein bioanalyzer peaks showed a similar result to the PCA analysis of the RP-HPLC peaks (Fig. 3B). The separation between the TSO population and the ANF and ONF populations is more distinct, and we do not see a clear distinction between the ANF and ONF populations. The first five principal components explained 92.7% of the variation in the protein bioanalyzer data set making up 45.4%, 19.2%, 13.6%, 8.5% and 6.0%, respectively. The most variable peaks were 1, 2, 3, 6 and 11.

Using a variance matrix to determine the percent variance of each RP-HPLC peak (*i.e.* which peaks were the least and most variable among populations), peaks 12 and 24 were identified as the most variable, contributing to 10.2% and 10.5% of the variance, respectively, followed

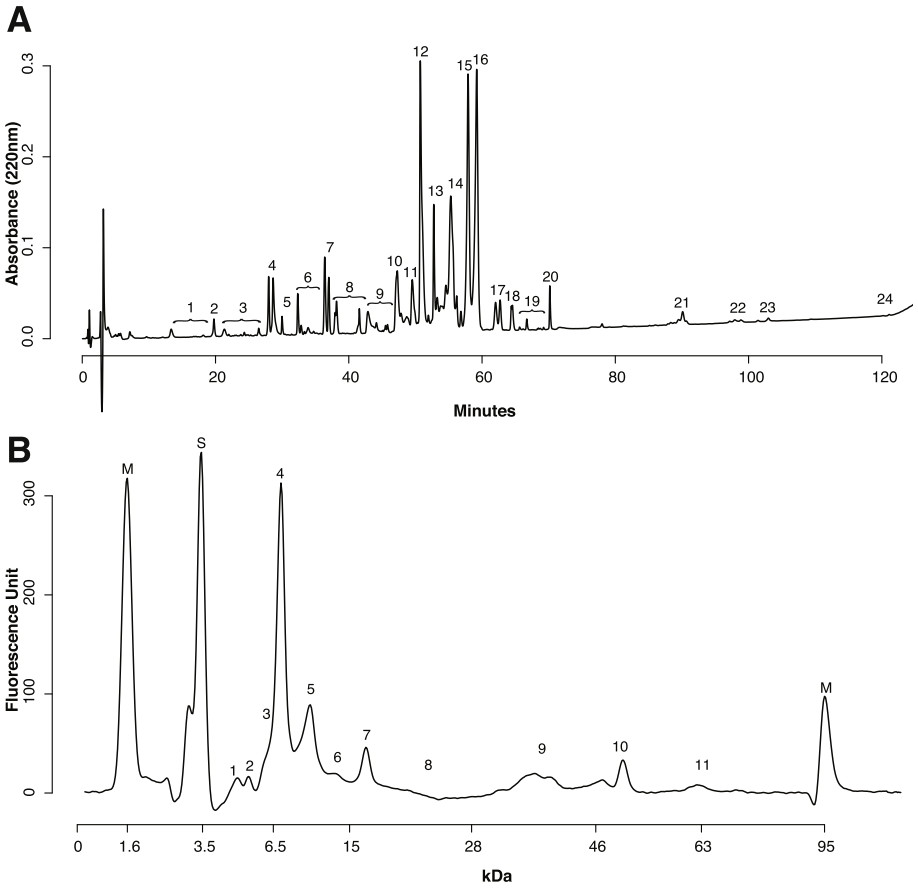

Fig. 2. Representative RP-HPLC (A) and protein bioanalyzer (B) profiles for one individual. Peaks are indicated by number. In the protein bioanalyzer profile (B), M and S indicate internal markers and system peak, respectively.

closely by peak 11, which contributed to 9.3% of the variance, and peak 1, which contributed 8.7% (Fig. 4A). Representative RP-HPLC profiles with peak identities can be seen in Fig. 2A. In the variance matrix of the protein bioanalyzer data, peak 1 was identified as the most variable peak among populations, contributing to 20.3% of the variance, followed by peaks 6, 2, 7 and 11, which contributed to 17.5%, 10.1%, 10.1% and 9.9% of the variance, respectively (Fig. 4B). Representative protein bioanalyzer profiles can be seen in Fig. 2B, where peak identities are indicated.

To better visualize the contribution of each peak to the observed

Table 1
Compositional venom variation.

RP-HPLC	Among populations	Df	SumOfSqs	MeanSqs	*F*	R^2	*P*
	Pop	2	114.33	57.17	3.7	0.19	2.2×10^{-5}
	Sex	1	47.44	47.44	3.07	0.08	4.3×10^{-3}
	Pop:Sex	2	37.93	18.97	1.23	0.06	0.23
	Residuals	27	417.39	15.46		0.68	
	Total	32	617.10			1	
Protein bioanalyzer	Among populations	Df	SumOfSqs	MeanSqs	*F*	R^2	*P*
	Pop	2	74.78	37.39	4.85	0.24	3.6×10^{-5}
	Sex	1	11.76	11.76	1.53	0.04	0.17
	Pop:Sex	2	16.70	8.50	1.10	0.05	0.34
	Residuals	27	207.95	7.70		0.67	
	Total	32	311.49			1	
RP-HPLC	Within populations	Df	SumOfSqs	MeanSqs	*F*	R^2	*P*
	Pop	2	299.55	149.77	13.88	0.17	1.0×10^{-6}
	Pop:Sex	3	164.19	54.73	5.07	0.09	1.0×10^{-6}
	Pop:Sex:Animal	27	586.61	21.73	2.01	0.33	1.0×10^{-6}
	Residuals	66	712.25	10.79		0.40	
	Total	98	1,762.60			1	

Results of nonparametric MANOVA tests run on the partial RP-HPLC data set (among populations), protein bioanalyzer data set (among populations), and complete RP-HPLC data set (within populations) reveal significant levels of variation among and within populations, between sexes, and among individuals within populations. Sources of variation were determined through *post-hoc* analyses (Tables S2–S4).

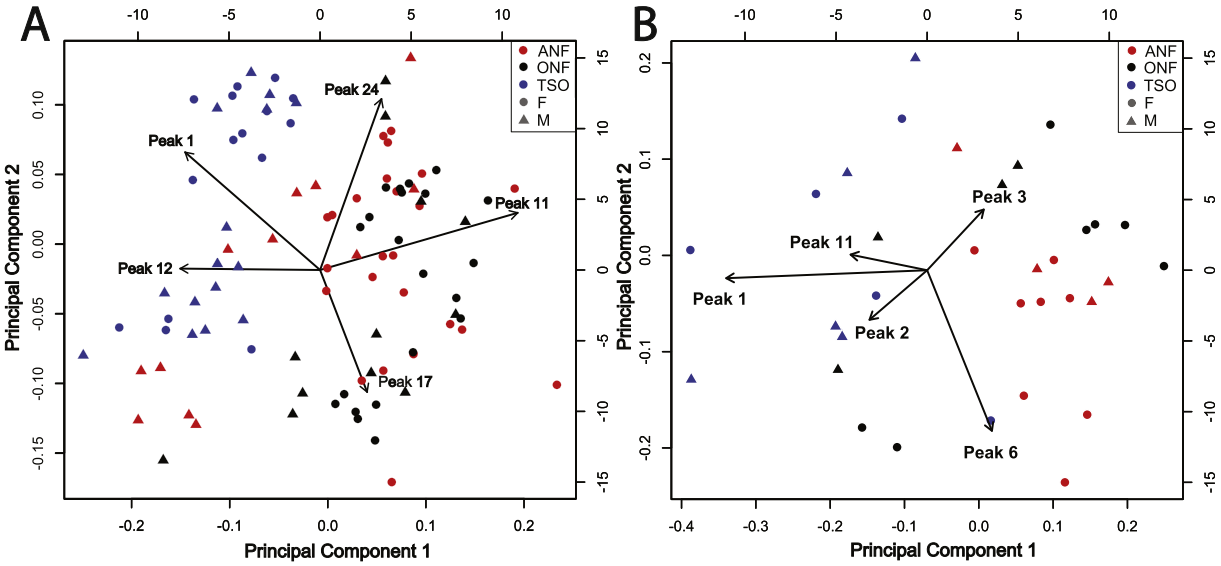

Fig. 3. The principal component analysis of the complete RP-HPLC data set (A) and the protein bioanalyzer data set (B) revealed separation of TSO from the ANF and ONF populations. The five peaks exhibiting most of the variation among populations are displayed. Populations are indicated by color and sexes are indicated by shape.

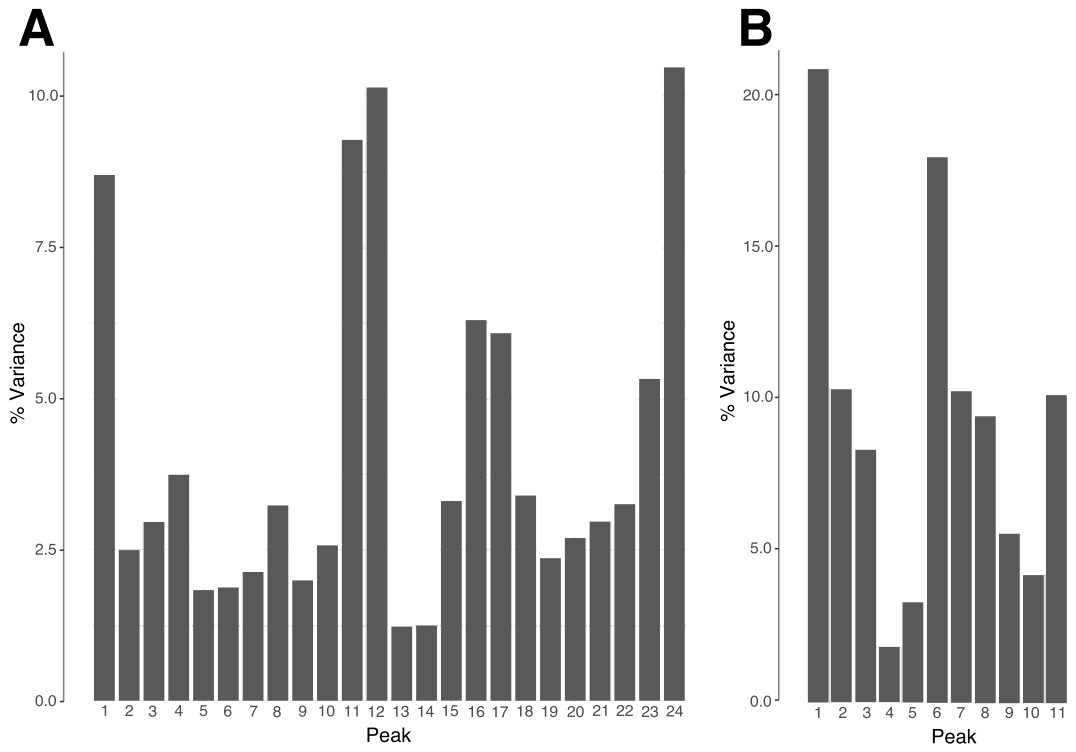

Fig. 4. Bar graph representation of among-population variance matrices of RP-HPLC peaks (A) and protein bioanalyzer peaks (B). The percent variation shown is the percent variation each peak contributes to the overall variation among all three *C. hentzi* populations.

variation among populations and between sexes, we calculated the average expression for each peak in each population and the average expression for each peak by sex, and plotted pairwise correlations of these values between populations (Fig. 5: RP-HPLC and Fig. 6: protein bioanalyzer). These plots show few differences in average expression between ANF and ONF populations, and the peaks that are more variable are weakly expressed. In comparison to ANF and ONF, we found less agreement in average expression between ONF and TSO, and between ANF and TSO, including in proteins that are more highly expressed. When looking at average expression between males and females, including individuals from all populations, we found variation in expression that appears to be evenly spread between both highly and weakly expressed proteins.

3.2. Venom variation within populations

3.2.1. Intersexual and among-individual variation contributed to overall venom variation within populations

We performed a nonparametric MANOVA on the complete RP-HPLC ilr transformed data set (three samples per individual), using a nested model to account for the relationships among animal, sex and population, while constraining permutations within each population. We

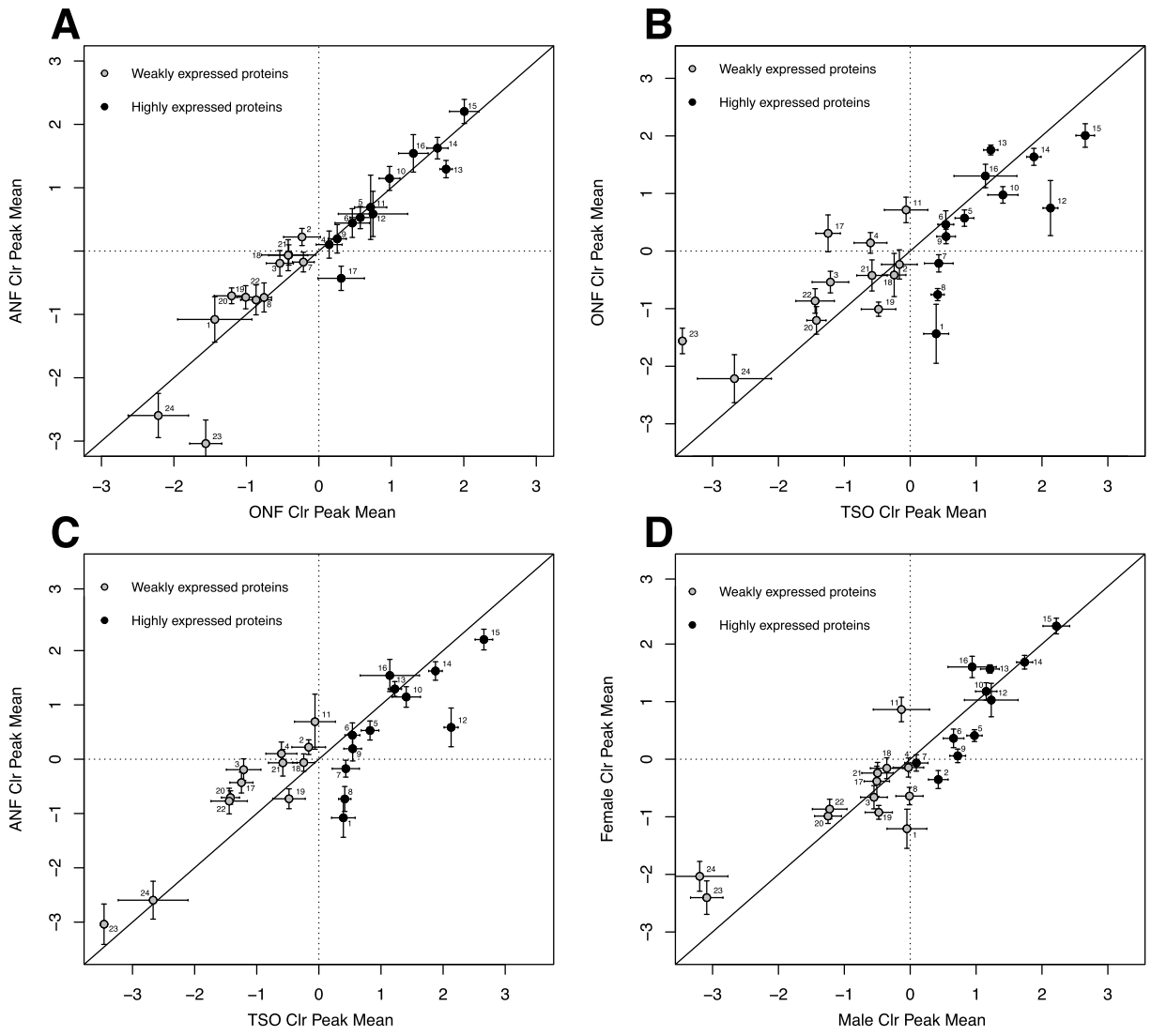

Fig. 5. Average expression between each of the three *C. hentzi* populations (A–C) and between the sexes (D), using the partial RP-HPLC profile data set. Individuals from all three populations are included in the sex comparison. Points along the diagonal line indicate strong agreement in expression and points deviating from the diagonal indicate more variable expression. Clr—centered logratio transformation.

found significant levels of variation within populations ($F = 13.88$, $R^2 = 0.17$, $P = 1.0 \times 10^{-6}$), between sexes within populations ($F = 5.07$, $R^2 = 0.09$, $P = 1.0 \times 10^{-6}$), and among individuals within sexes, within populations ($F = 2.01$, $R^2 = 0.33$, $P = 1.0 \times 10^{-6}$, Table 1). We then performed a series of *post-hoc* permutational MANOVA tests to determine the sources of variation within each population, using the Bonferroni adjusted *P*-values to correct for multiple comparisons, indicated by P^* (Table S4). Within the ANF population, we found a significant difference between the sexes ($F = 8.86$, $R^2 = 0.15$, $P^* = 9.0 \times 10^{-6}$) as well as among animals within the sexes ($F = 2.7$, $R^2 = 0.45$, $P^* = 9.0 \times 10^{-6}$). Within the ONF population, we did not find a significant difference between sexes ($F = 1.4$, $R^2 = 0.03$, $P^* = 1.0$), but we did find significant variation among animals within the sexes ($F = 1.86$, $R^2 = 0.42$, $P^* = 2.3 \times 10^{-3}$). Within the TSO population, we found significant variation between the sexes ($F = 6.23$, $R^2 = 0.17$, $P^* = 6.6 \times 10^{-4}$), but not among animals within the sexes ($F = 1.38$, $R^2 = 0.3$, $P^* = 0.72$).

3.2.2. Males contributed more to venom variation within populations than females

Using only female samples from the complete RP-HPLC data set, we found significant variation within females of the ANF population ($F = 1.82$, $R^2 = 0.44$, $P^* = 3.1 \times 10^{-3}$), but we did not find significant levels of variation within females of the ONF or TSO populations. Using only male samples from the complete RP-HPLC data set, we found significant levels of variation within males of the ANF population ($F = 5.66$, $R^2 = 0.68$, $P^* = 6.1 \times 10^{-4}$) and within males of the ONF population ($F = 3.00$, $R^2 = 0.53$, $P^* = 3.6 \times 10^{-3}$), but we did not find variation within males of the TSO population.

3.2.3. Within-individual variation did not impact the within-population results

To determine whether or not replicate samples from the same individual were variable, we ran a nonparametric MANOVA using a model with the calculated timing of the venom sample (see methods section 2.6.5) being nested within animal, and used the strata function to constrict permutations within each animal. The results revealed no significant differences among replicate samples for an individual due to the timing of the venom extraction ($F = 3.32$, $R^2 = 0.60$, $P = 0.21$).

3.2.4. Venom components contributing to venom variation within populations were consistent with those detected among populations

We used variance matrices to determine the percent variance of each RP-HPLC and protein bioanalyzer peak within each population

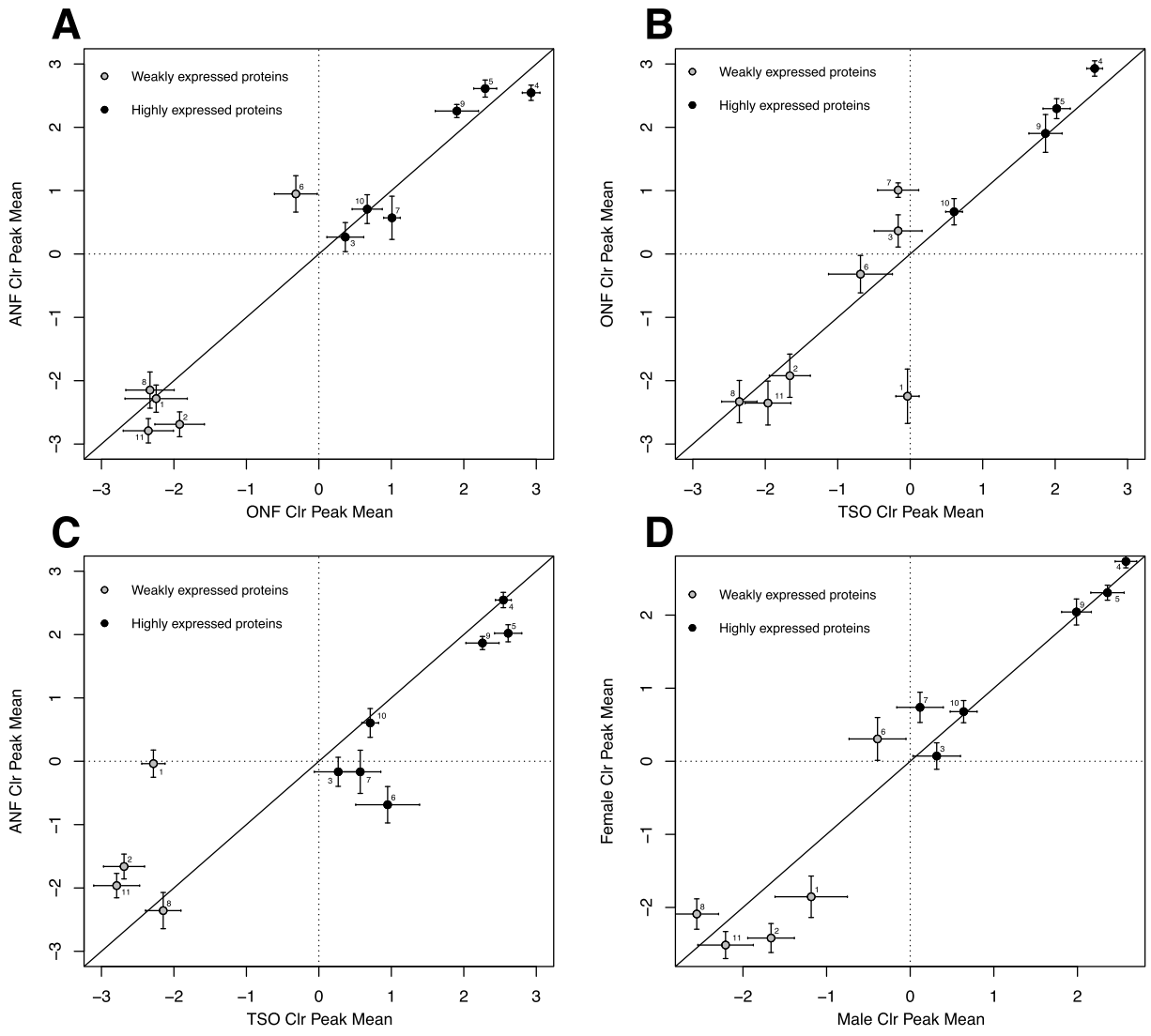

Fig. 6. Average expression between each of the three *C. hentzi* populations (A–C) and between the sexes (D), using the protein bioanalyzer profile data set. Individuals from all three populations are included in the sex comparison. Points along the diagonal line indicate strong agreement in expression and points deviating from the diagonal indicate more variable expression. Clr—centered logratio transformation.

and within each sex (Tables S5–S10). Within the complete RP-HPLC data set (Table S5), peak 11 was the most variable peak within ANF, representing 17.0% of the variance, peak 12 was the most variable peak within ONF, representing 14.7% of the variance, and peak 24 was the most variable peak within TSO, representing 20.6% of the variance. These peaks correspond with the most variable peaks that were identified among populations. Within the RP-HPLC data set using only male samples (Table S6), peak 11 was the most variable peak within the ANF population, representing 29.2% of the variance, peak 12 was the most variable peak within ONF males, representing 11.7% of the variance, and peak 16 was the most variable peak within TSO males, representing 18.3% of the variance. Within the RP-HPLC data set using only female samples (Table S7), peak 1 was the most variable peak within ANF females, representing 10.5% of the variance, peak 12 was the most variable peak within ONF females, representing 15.2% of the variance, and peak 24 was the most variable peak within TSO females, representing 23.8% of the variance. These results correspond to the patterns of variation that were detected by the within population MANOVA analyses (Table S4). Within the ONF population, the same peak was the most variable regardless of sex, suggesting less intersexual variation within the ONF population (MANOVA $P^* = 1.0$). Within the ANF and TSO populations, different peaks were identified in each sex as being the most variable, suggesting more intersexual variation within these two populations (ANF MANOVA $P^* = 9.0 \times 10^{-6}$, TSO MANOVA $P^* = 6.6 \times 10^{-4}$).

In the protein bioanalyzer data (Table S8), peak 7 was the most variable peak within the ANF population (21.2% of the variance), peak 1 was the most variable peak within the ONF population (21.2% of the variance), and peak 6 was the most variable peak within the TSO population (25.5% of the variance). Similar to the RP-HPLC results, these peaks correspond with the most variable peaks that were identified among populations. Within the protein bioanalyzer data set using only male samples (Table S9), peak 6 was the most variable within ANF males (24.2% of the variance), and peak 11 was the most variable among both ONF and TSO males, with 23.6% and 19.1% of the total variance, respectively. Within the protein bioanalyzer data set using only female samples (Table S10), peak 7 was the most variable within ANF females (21.3% of the variance), peak 1 was the most variable within ONF females (23.1% of the variance), and peak 6 was the most variable within TSO females (38.3% of the variance). The different peaks identified as being the most variable in each sex suggests possible intersexual variation within each population, but this was not detected as significant in the protein bioanalyzer MANOVA analyses (Table S3).

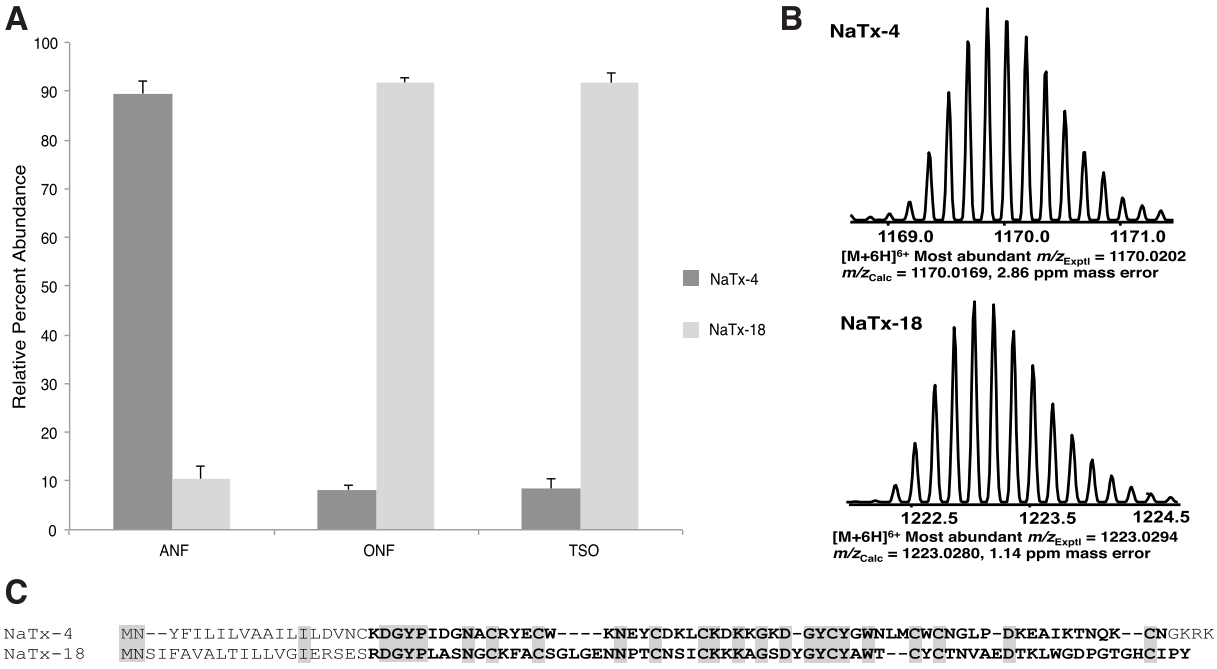

Fig. 7. Peak 12 protein identification of NaTx-4 and NaTx-18. (A) Differences in relative abundance of NaTx-4 and NaTx-18 in a female venom sample from each population. MS signals were averaged over the nanoLiquid Chromatography (nLC) chromatogram of eluted proteins and are reported as relative percent abundance. (B) Proteomic verification of NaTx-4 (7029 Da.) and NaTx-18 (7332 Da), with the most abundant mass-to-charge (m/z) ratio indicated. (C) Amino acid sequence alignments of NaTx-4 and NaTx-18 show evidence of post-translational processing and conserved cysteine residues required for the formation of four disulfide bonds. The precursor sequences for NaTx-4 and NaTx-18 were obtained from published data (Ward et al., 2018) and have 31% sequence identity to each other. Conserved amino acids between the two peptides are highlighted and include eight cysteine residues. Processed amino acid sequences are shown in bold.

3.3. Protein identification

The variance matrix of the complete RP-HPLC data revealed that peaks 12 and 24 contributed most to the variation among the three populations (Fig. 4A, Table S5). Peak 12, which was present in 93% of RP-HPLC venom samples, was identified as the most abundant variable peak among the three populations. This peak was isolated from a representative female sample from each of the three populations, and nLC/MS/MS, using both top-down and bottom-up approaches (see methods section 2.4), was used to identify the proteins present in this peak. The results of nLC/MS/MS revealed the same two Na^+-channel toxins: NaTx-4 (7029 Da) and NaTx-18 (7332 Da), although the relative abundance of each of these toxins was variable among populations (Fig. 7). In the ANF sample, NaTx-4 was the most abundant toxin identified in peak 12, with NaTx-18 representing only a small portion of this peak. In the ONF and TSO samples we see the opposite pattern, with NaTx-18 being the most abundant toxin identified and NaTx-4 representing only a small portion of peak 12. Sodium channel toxins are characterized by the presence of four disulfide bonds, are 60–76 amino acids in length, and act by slowing or blocking the inactivation of Na^+-channels (Possani et al., 1999; Bosmans and Tytgat, 2007; Gurevitz et al., 2007). Precursor and processed amino acid alignments for NaTx-4 and NaTx-18 are shown in Fig. 7. The precursor amino acid lengths for NaTx-4 and NaTx-18 were 83 and 89 amino acids, respectively (Ward et al., 2018), and show 31% sequence identity to each other. In NaTx-4, we found evidence of post-translational processing at both the N and C-terminus, resulting in a final processed length of 60 amino acids with eight cysteine residues, consistent with the formation of four disulfide bonds. In NaTx-18, we found evidence of post-translational processing at the N-terminus, resulting in a final length of 68 amino acids, including eight cysteine residues required for the formation of four disulfide bonds.

3.4. Functional venom variation

When testing venom from the three *C. hentzi* populations in *D. melanogaster*, we were able to use the 24 h time point to calculate LD_{50} and ED_{50} values. We found a significant functional difference between the ANF and TSO populations (ratio test result did not contain 1, see methods section 2.6.7), with ANF venom being nearly twice as toxic to *D. melanogaster* compared to TSO venom (Table 2). We did not, however, find a significant functional difference between ANF and ONF or between ONF and TSO venoms when injected into *D. melanogaster*.

When testing *C. hentzi* venom from the ANF population in *G. sigillatus*, one of the ten controls was recorded as dead at the 12 h time point. Therefore, all LD_{50} and ED_{50} values were calculated using recorded data from the 8 h time point of each assay. Consistent with our compositional results, the results of the ED_{50} ratio tests in *G. sigillatus* revealed significant functional differences between TSO and ANF as well as between TSO and ONF, and no significant difference in venom function between the ANF and ONF venoms (Table 2). In comparison to the TSO venom, both ANF and ONF venoms were more toxic to *G. sigillatus*, with ANF being slightly more so than ONF.

Because we were able to capture the same variation pattern as in the compositional analyses when injecting *C. hentzi* venom into *G. sigillatus*, and because the most significant compositional differences between males and females seemed to be present in the ANF *C. hentzi* population, we tested for functional differences between sexes using pooled venom samples (one all-male and one all-female) from the ANF population. We found a significant difference between male and female venoms of the ANF population when injected into *G. sigillatus*, with the female venom being almost three times more toxic relative to the male venom (Table 2).

4. Discussion

Intraspecific venom variation has been found among populations (Abdel-Rahman et al., 2009; Ruiming et al., 2010; Rodríguez-Ravelo

Table 2
LD_{50} and ED_{50} values.

G. sigillatus	LD_{50} (TSK)				ED_{50} (DRM)			
	LD_{50} (μg/mg)	Std. Dev.	Lower 95% CI	Upper 95% CI	ED_{50} (μg/mg)	Std. Error	Lower 95% CI	Upper 95% CI
ANF	0.512	0.072	0.368	0.653	0.429	0.0767	0.279	0.579
ONF	0.608	0.0554	0.499	0.716	0.560	0.0645	0.434	0.687
TSO	1.02	0.124	0.775	1.26	0.955	0.112	0.735	1.17
ANF (Male)	0.688	0.103	0.486	0.889	0.467	0.112	0.284	0.686
ANF (Female)	0.409	0.104	0.205	0.614	0.168	0.0733	0.0245	0.312
D. melanogaster	LD_{50} (ng/mg)	Std. Dev.	Lower 95% CI	Upper 95% CI	ED_{50} (ng/mg)	Std. Error	Lower 95% CI	Upper 95% CI
ANF	0.885	0.109	0.67	1.10	0.717	0.089	0.542	0.892
ONF	1.20	0.155	0.892	1.50	0.812	0.145	0.527	1.10
TSO	1.53	0.214	1.11	1.95	1.24	0.281	0.693	1.79

LD_{50} values were calculated using the Trimmed Spearman-Karber (TSK) method, with standard deviation and 95% confidence intervals shown. ED_{50} values were calculated by fitting the responses to a dose-response model (DRM), and provide standard errors in the measurement, which can be used in statistical testing using the ratio test. The lower and upper 95% confidence intervals for the calculated ED_{50} values are also shown.

et al., 2013; Estrada-Gómez et al., 2014; Carcamo-Noriega et al., 2017; Schaffrath et al., 2018) and between sexes (De Sousa et al., 2010; Rodríguez-Ravelo et al., 2015; Miller et al., 2016; Cid-Uribe et al., 2017) of multiple scorpion species, but none of these studies have examined venom expression variation within populations or within sexes. Focusing on three populations of *C. hentzi* in North and Central Florida, using both RP-HPLC and protein bioanalyzer venom profiles, we found significant venom variation among populations, with greater significance between populations with farther geographic separation. We also found significant differences in venom composition between males and females across all populations. Variation among populations, however, was only detected using female, not male, venom samples. This indicates a female-driven population divergence in this species and is suggestive of venom-specific, sex-biased gene expression.

We found significant levels of variation within populations, which was attributed to variation between sexes and/or among individuals within each population. Overall, we found greater variation within males of each population in comparison to females (aside from the TSO population). Although not specifically documented in *C. hentzi*, other *Centruroides* species have exhibited differences in dispersal and behavior between males and females that may explain, at least in part, differences in venom composition (Carlson and Rowe, 2009; Carlson et al., 2014). Our results support the theory that male *C. hentzi* are more widely dispersed as they search for mates, with possible migration between populations, and therefore are subject to different types of selective pressures than if they remained tethered to a specific location (Polis and Farley, 1979; Booncham et al., 2007; Carlson and Rowe, 2009; Kaltsas and Mylonas, 2010). Male venom variation may also be due to sexual selection, such as female mate choice (Polis and Sissom, 1990; Tallarovic et al., 2000; Contreras-Garduño et al., 2006) or male-male competition (Benton, 1992) that have been observed in some species, but, to our knowledge, these behaviors have not been specifically documented in *Centruroides*. Regardless, considering females were responsible for the variation observed among populations, this also supports that female *C. hentzi* are less mobile than males, and that their venom may be more suited to their specific microenvironment within their respective population.

The recently completed venom-gland transcriptome and proteome of a male and female *C. hentzi* showed differences in expression between the sexes (Ward et al., 2018). Although only one male and one female *C. hentzi* were analyzed, several toxins were detected in the female that were completely absent from the male, and vice versa. This genetic framework may indicate which venom genes are sex-specific and therefore contribute more to population divergence. The female *C. hentzi* venom showed a greater percentage of sodium channel toxins (NaTxs) in comparison to the male, with seven NaTxs expressed in high concentrations (relative to other proteomically detected toxins) in the female venom that were completely absent from the male venom. Although some scorpion NaTxs are known to cause severe pain (Rowe et al., 2013; Miller et al., 2016), many NaTxs are insect specific (Possani et al., 1999; Bosmans and Tytgat, 2007; Gurevitz et al., 2007) and ideal for prey capture. The seven NaTxs identified in the female *C. hentzi* venom are likely female-biased venom genes that could have arisen for more effective and specialized prey capture, defense, or both. The male *C. hentzi* venom has a greater abundance of potassium channel toxins (KTxs) and general venom proteins (VPs) that have yet to be characterized by function or homology to known toxins (Ward et al., 2018), suggesting that these toxins may play a role in mating, or in less-specialized prey capture or defensive strategies while on the move. However, additional venom-gland transcriptome and proteome characterization of multiple males and females from different populations would be necessary to determine which venom genes are truly sex-biased.

We identified and characterized the most abundant RP-HPLC peak responsible for variation among populations and found the same two NaTxs present in all populations, NaTx-4 and NaTx-18, although the relative amounts of each NaTx differed by population (Fig. 7). In comparing the average expression of each peak between populations and between sexes (Fig. 5), peak 12 does not seem to differ in expression between ANF and ONF, with greater expression differences observed between ANF and TSO. Despite the protein content being nearly identical in the ONF and TSO samples, we still see differences in average expression of peak 12 between these populations. In the average expression differences between the sexes (Fig. 5D), males and females from all three populations do not seem to differ in average expression of peak 12, indicating that although NaTx-4 and NaTx-18 may be candidates for sex-biased genes, these toxins are more involved in population venom variation rather than intersexual variation.

Variation in venom function as measured by toxicity assays was strongly correlated with the compositional variation determined by RP-HPLC and protein bioanalyzer data. In *D. melanogaster*, we were able to capture the more extreme differences between the ANF and TSO populations, with the ANF venom being more toxic to *D. melanogaster* relative to TSO. In *G. sigillatus*, the functional results corresponded to the molecular results exactly. We found no functional difference between the ANF and ONF populations, and we did find significant differences in toxicity between the ONF and TSO as well as the ANF and TSO populations, where TSO was the less toxic venom in both cases. We also found a significant difference in toxicity between male and female venoms of the ANF population, with the female venom being more toxic than the male venom. Higher female venom toxicity has been found in

other scorpion species (De Sousa et al., 2010). This may, however, be a prey-specific phenomenon as Miller et al. (2016) found male venom to be more of an irritant to mice in comparison to female venom from the closely related *Centruroides vittatus*. One possible explanation is that female venom may be more locally adapted for prey capture and defense of smaller predators (including other arachnids), whereas male venom may be better suited for defense against larger predators such as birds and mammals.

In our compositional and functional analyses of venom variation among three Florida populations of *C. hentzi*, we established female-biased intraspecific venom variation as well as venom variation within populations that is due to both differences among individuals and between sexes. Although the underlying mechanisms relating natural and sexual selection to sex-biased genes resulting in sexual dimorphism remain unknown, our results provide evidence that male and female scorpions experience different natural and sexual selective pressures that have led to the expression of sex-biased venom genes, and that these genes are more consequential in population divergence than previously considered. As our database of venom-gland transcriptome and proteome characterizations continues to grow, so does our ability to identify potential sex-biased venom genes, which, when partnered with behavior and ecology, will allow us to unravel the complex evolutionary phenomenon of intersexual venom variation.

Ethical statement

Reporting standards

The authors declare that our manuscript describes original research and every effort was made to ensure the accuracy of the results and the account.

Data access and retention

The authors will make the raw RP-HPLC and Protein Bioanalyzer data available upon request for those who wish to perform their own analyses using the raw profiles. The data will be retained among the authors indefinitely.

Originality and plagiarism

The authors declare that our manuscript is an original work with proper citations as needed.

Multiple, redundant or concurrent publication

The authors declare that the data and work described in our manuscript has not and will not be submitted for consideration to another journal.

Acknowledgment of sources

The authors have provided proper acknowledgment of sources to the best of their abilities.

Authorship of paper

All authors of the manuscript made significant contributions, and no one making significant contributions was excluded from authorship. All authors have seen and read the submitted version of the manuscript and have approved submission.

Hazards and human or animal subjects

This review study did not involve the use of vertebrate animals.

Disclosure and conflicts of interest

The authors declare no conflicts of interest.

Fundamental errors in published works

If a fundamental error or inaccuracy is discovered in the results described by our literature review, the authors will immediately notify the editor or publisher.

Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1449440. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Funding for this work was provided by the National Science Foundation (NSF DEB-1145978 and NSF DEB-1638902) and the Florida State University Council on Research and Creativity (CRC PG-036698). We thank Margaret Seavy of the Florida State University Analytical Lab and Research facility for advice and assistance with RP-HPLC analyses, Kimberly Hughes for providing lab space for the *D. melanogaster* used in this study, and Mark Margres and David Houle for their guidance in statistical analyses. We thank Pierson Hill and Andrew Mason for their assistance in acquiring *C. hentzi* specimens, and we also thank Cynthia Wang and Alie Akins for their contributions to this project through the Young Scholars Program at Florida State University.

Transparency document

Transparency document related to this article can be found online at https://doi.org/10.1016/j.toxicon.2018.07.026.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.toxicon.2018.07.026.

References

Abdel-Rahman, M.A., Omran, M.A.A., Abdel-Nabi, I.M., Ueda, H., McVean, A., 2009. Intraspecific variation in the Egyptian scorpion *Scorpio maurus palmatus* venom collected from different biotopes. Toxicon 53, 349–359.
Aitchison, J., 1982. The statistical analysis of compositional data. J. Roy. Stat. Soc. B 139–177.
Akaike, H., 2011. Akaike's Information Criterion, International Encyclopedia of Statistical Science. Springer, pp. 25.
Benton, T., 1992. Determinants of male mating success in a scorpion. Anim. Behav. 43, 125–135.
Benton, T., 2001. Reproductive Ecology. Scorpion Biology and Research. Oxford University Press, New York, pp. 278–301.
Biardi, J.E., Chien, D.C., Coss, R.G., 2005. California ground squirrel (*Spermophilus beecheyi*) defenses against rattlesnake venom digestive and hemostatic toxins. J. Chem. Ecol. 31, 2501–2518.
Biardi, J.E., Nguyen, K.T., Lander, S., Whitley, M., Nambiar, K.P., 2011. A rapid and sensitive fluorometric method for the quantitative analysis of snake venom metalloproteases and their inhibitors. Toxicon 57, 342–347.
Booncham, U., Sitthicharoenchai, D., Pradatsundarasar, A., Prasarnpun, S., Thirakhupt, K., 2007. Sexual dimorphism in the Asian giant forest scorpion, *Heterometrus laoticus* Couzijn, 1981. NU. Int. J. Sci. 4, 42–52.
Bosmans, F., Tytgat, J., 2007. Voltage-gated sodium channel modulation by scorpion α-toxins. Toxicon 49, 142–158.
Carcamo-Noriega, E.N., Olamendi-Portugal, T., Restano-Cassulini, R., Rowe, A., Uribe-Romero, S.J., Becerril, B., Possania, L.D., 2017. Intraspecific variation of *Centruroides sculpturatus* scorpion venom from two regions of Arizona. Arch. Biochem. Biophys. 638, 52–57.
Carlson, B.E., Rowe, M.P., 2009. Temperature and desiccation effects on the antipredator behavior of *Centruroides vittatus* (Scorpiones: Buthidae). J. Arachnol. 37, 321–330.
Carlson, B.E., McGinley, S., Rowe, M.P., 2014. Meek males and fighting females: sexually-dimorphic antipredator behavior and locomotor performance is explained by morphology in bark scorpions (*Centruroides vittatus*). PLoS One 9, e97648.
Chenoweth, S.F., Rundle, H.D., Blows, M.W., 2007. Genetic constraints and the evolution

of display trait sexual dimorphism by natural and sexual selection. Am. Nat. 171, 22–34.
Cid-Uribe, J.I., Vargas, J.M.J., Batista, C.V.F., Zuñiga, F.Z., Possani, L.D., 2017. Comparative proteomic analysis of female and male venoms from the Mexican scorpion *Centruroides limpidus*: novel components found. Toxicon 125, 91–98.
Connallon, T., Knowles, L.L., 2005. Intergenomic conflict revealed by patterns of sex-biased gene expression. Trends Genet. 21, 495–499.
Contreras-Garduño, J., Peretti, A.V., Córdoba-Aguilar, A., 2006. Evidence that mating plug is related to null female mating activity in the scorpion *Vaejovis punctatus*. Ethology 112, 152–163.
De Sousa, L., Borges, A., Vásquez-Suárez, A., den Camp, H.J.O., Chadee-Burgos, R.I., Romero-Bellorín, M., Espinoza, J., De Sousa-Insana, L., Pino-García, O., 2010. Differences in venom toxicity and antigenicity between females and males *Tityus nororientalis* (Buthidae) scorpions. J. Venom Res. 1, 61–70.
D'suze, G., Sandoval, M., Sevcik, C., 2015. Characterizing *Tityus discrepans* scorpion venom from a fractal perspective: venom complexity, effects of captivity, sexual dimorphism, differences among species. Toxicon 108, 62–72.
Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barceló-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. Math. Geol. 35, 279–300.
Ellegren, H., Parsch, J., 2007. The evolution of sex-biased genes and sex-biased gene expression. Nat. Rev. Genet. 8, 689–698.
Estrada-Gómez, S., Cupitra, N.I., Arango, W.M., Muñoz, L.J.V., 2014. Intraspecific variation of *Centruroides edwardsii* venom from two regions of Colombia. Toxins 6, 2082–2096.
Filzmoser, P., Hron, K., Reimann, C., 2009. Principal component analysis for compositional data with outliers. Environmetrics 20, 621–632.
Gangur, A.N., Smout, M., Liddell, M.J., Seymour, J.E., Wilson, D., Northfield, T.D., 2017. Changes in predator exposure, but not in diet, induce phenotypic plasticity in scorpion venom. Proc. R. Soc. B 284, 20171364. https://doi.org/10.1098/rspb.2017.1364.
Gibbs, H.L., Sanz, L., Calvete, J.J., 2009. Snake population venomics: proteomics-based analyses of individual variation reveals significant gene regulation effects on venom protein expression in *Sistrurus* rattlesnakes. J. Mol. Evol. 68, 113–125.
Gold, R.V., Loening, K., McNaught, A., Sehmi, P., 1987. Compendium of Chemical Terminology. IUPAC. JSTOR.
Gurevitz, M., Karbat, I., Cohen, L., Ilan, N., Kahn, R., Turkov, M., Stankiewicz, M., Stühmer, W., Dong, K., Gordon, D., 2007. The insecticidal potential of scorpion β-toxins. Toxicon 49, 473–489.
Hamilton, M.A., Russo, R.C., Thurston, R.V., 1977. Trimmed Spearman-Karber method for estimating median lethal concentrations in toxicity bioassays. Environ. Sci. Technol. 11, 714–719.
Herzig, V., Hodgson, W.C., 2009. Intersexual variations in the pharmacological properties of *Coremiocnemis tropix* (Araneae, Theraphosidae) spider venom. Toxicon 53, 196–205.
Herzig, V., Ward, R.J., dos Santos, W.F., 2002. Intersexual variations in the venom of the Brazilian 'armed' spider *Phoneutria nigriventer* (Keyserling, 1891). Toxicon 40, 1399–1406.
Herzig, V., Khalife, A.A., Chong, Y., Isbister, G.K., Currie, B.J., Churchill, T.B., Horner, S., Escoubas, P., Nicholson, G.M., Hodgson, W.C., 2008. Intersexual variations in Northern (*Missulena pruinosa*) and Eastern (*M. bradleyi*) mouse spider venom. Toxicon 51, 1167–1177.
Hurlbert, S.H., 1984. Pseudoreplication and the design of ecological field experiments. Ecol. Monogr. 54, 187–211.
Kaltsas, D., Mylonas, M., 2010. Locomotory activity and orientation of *Mesobuthus gibbosus* (Scorpiones: Buthidae) in central aegean archipelago. J. Nat. Hist. 44, 1445–1459.
Khaitovich, P., Hellmann, I., Enard, W., Nowick, K., Leinweber, M., Franz, H., Weiss, G., Lachmann, M., Pääbo, S., 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. Science 309, 1850–1854.
Lourenço, W.R., 2000. Reproduction in scorpions, with special reference to parthenogenesis. Eur. Arachnol. 2000, 71–85.
Mank, J.E., 2008. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. Am. Nat. 173, 141–150.
Mank, J.E., Ellegren, H., 2009. Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. Evolution 63, 1464–1472.
Margres, M.J., McGivern, J.J., Seavy, M., Wray, K.P., Facente, J., Rokyta, D.R., 2015a. Contrasting modes and tempos of venom expression evolution in two snake species. Genetics 199, 165–176.
Margres, M.J., Wray, K.P., Seavy, M., McGivern, J.J., Sanader, D., Rokyta, D.R., 2015b. Phenotypic integration in the feeding system of the eastern diamondback rattlesnake (*Crotalus adamanteus*). Mol. Ecol. 24, 3405–3420.
Margres, M.J., Wray, K.P., Hassinger, A.T., Ward, M.J., McGivern, J.J., Moriarty Lemmon, E., Lemmon, A.R., Rokyta, D.R., 2017. Quantity, not quality: rapid adaptation in a polygenic trait proceeded exclusively through expression differentiation. Mol. Biol. Evol. 34, 3099–3110.
Miller, D.W., Jones, A.D., Goldston, J.S., Rowe, M.P., Rowe, A.H., 2016. Sex differences in defensive behavior and venom of the striped bark scorpion *Centruroides vitattus* (Scorpiones: Buthidae). Integr. Comp. Biol. 56, 1022–1031.
Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civello, D., White, T.J., et al., 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol. 3, e170.
Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Simpson, G., Solymos, P., Stevens, M., Wagner, H., 2013. Vegan: Community Ecology Package. R Package. ver. 2.0–10.
Palarea-Albaladejo, J., Martín-Fernández, J.A., 2015. zCompositions–R package for multivariate imputation of left-censored data under a compositional approach. Chemometr. Intell. Lab. Syst. 143, 85–96.
Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J., Eastman, S., Oliver, B., 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. Science 299, 697–700.
Polis, G.A., Farley, R.D., 1979. Behavior and ecology of mating in the cannibalistic scorpion, *Paruroctonus mesaensis* Stahnke (Scorpionida: vaejovidae). J. Arachnol. 33–46.
Polis, G., Sissom, W., 1990. Life History. The Biology of Scorpions. pp. 161–223.
Possani, L.D., Becerril, B., Delepierre, M., Tytgat, J., 1999. Scorpion toxins specific for Na^+-channels. FEBS J. 264, 287–300.
Pröschel, M., Zhang, Z., Parsch, J., 2006. Widespread adaptive evolution of *Drosophila* genes with sex-biased expression. Genetics 174, 893–900.
Ranz, J.M., Castillo-Davis, C.I., Meiklejohn, C.D., Hartl, D.L., 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. Science 300, 1742–1745.
Ritz, C., Baty, F., Streibig, J.C., Gerhard, D., 2015. Dose-response analysis using R. PLoS One 10, e0146021.
Rodríguez-Ravelo, R., Coronas, F.I., Zamudio, F.Z., González-Morales, L., López, G.E., Urquiola, A.R., Possani, L.D., 2013. The Cuban scorpion *Rhopalurus junceus* (Scorpiones, Buthidae): component variations in venom samples collected in different geographical areas. J. Venom. Anim. Toxins Incl. Trop. Dis. 19, 13.
Rodríguez-Ravelo, R., Batista, C.V., Coronas, F.I., Zamudio, F.Z., Hernández-Orihuela, L., Espinosa-López, G., Ruiz-Urquiola, A., Possani, L.D., 2015. Comparative proteomic analysis of male and female venoms from the Cuban scorpion *Rhopalurus junceus*. Toxicon 107, 327–334.
Rokyta, D.R., Ward, M.J., 2017. Venom-gland transcriptomics and venom proteomics of the blackback scorpion (*Hadrurus spadix*) reveal detectability challenges and an unexplored realm of animal toxin diversity. Toxicon 128, 23–37.
Rowe, A.H., Xiao, Y., Rowe, M.P., Cummins, T.R., Zakon, H.H., 2013. Voltage-gated sodium channel in grasshopper mice defends against bark scorpion toxin. Science 342, 441–446.
Ruiming, Z., Yibao, M., Yawen, H., Zhiyong, D., Yingliang, W., Zhijian, C., Wenxin, L., 2010. Comparative venom gland transcriptome analysis of the scorpion *Lychas mucronatus* reveals intraspecific toxic gene diversity and new venomous components. BMC Genom. 11, 452.
Schaffrath, S., Predel, R., 2014. A simple protocol for venom peptide barcoding in scorpions. EuPA Open Protenomics 3, 239–245.
Schaffrath, S., Prendini, L., Predel, R., 2018. Intraspecific venom variation in southern African scorpion species of the genera *Parabuthus, Uroplectes* and *Opistophthalmus* (Scorpiones: Buthidae, Scorpionidae). Toxicon 144, 83–90.
Schenker, N., Gentleman, J.F., 2001. On judging the significance of differences by examining the overlap between confidence intervals. Am. Statistician 55, 182–186.
Shaffer, L.R., Formanowicz Jr., D.R., 1996. A cost of viviparity and parental care in scorpions: reduced sprint speed and behavioural compensation. Anim. Behav. 51, 1017–1024.
Shine, R., 1989. Ecological causes for the evolution of sexual dimorphism: a review of the evidence. Q. Rev. Biol. 64, 419–461.
Stone, B.R., 2015. Tsk: Trimmed Spearman-Karber Method. URL. https://github.com/brsr/tsk.
Tallarovic, S.K., Melville, J.M., Brownell, P.H., 2000. Courtship and mating in the giant hairy desert scorpion, *Hadrurus arizonensis* (Scorpionida, Iuridae). J. Insect Behav. 13, 827–838.
Templ, M., Hron, K., Filzmoser, P., 2011. robCompositions: Robust Estimation for Compositional Data. Manual and Package. version 1.
Torgerson, D.G., Kulathinal, R.J., Singh, R.S., 2002. Mammalian sperm proteins are rapidly evolving: evidence of positive selection in functionally diverse genes. Mol. Biol. Evol. 19, 1973–1980.
Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, fourth ed. Springer, New York0-387-95457-0 URL. http://www.stats.ox.ac.uk/pub/MASS4.
Ward, M.J., Ellsworth, S.A., Rokyta, D.R., 2018. 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 142, 14–29.
Wheeler, M.W., Park, R.M., Bailer, A.J., 2006. Comparing median lethal concentration values using confidence interval overlap or ratio tests. Environ. Toxicol. Chem. 25, 1441–1444.
Williams, S.C., 1987. Scorpion bionomics. Annu. Rev. Entomol. 32, 275–295.
Yang, X., Schadt, E.E., Wang, S., Wang, H., Arnold, A.P., Ingram-Drake, L., Drake, T.A., Lusis, A.J., 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Res. 16, 995–1004.
Zhang, Z., Hambuch, T.M., Parsch, J., 2004. Molecular evolution of sex-biased genes in *Drosophila*. Mol. Biol. Evol. 21, 2130–2139.
Zobel-Thropp, P.A., Bulger, E.A., Cordes, M.H., Binford, G.J., Gillespie, R.G., Brewer, M.S., 2018. Sexually dimorphic venom proteins in long-jawed orb-weaving spiders (*Tetragnatha*) comprise novel gene families. PeerJ 6, e4691.