MOLECULAR ECOLOGY

Molecular Ecology (2015) doi: 10.1111/mec.13240

Phenotypic integration in the feeding system of the eastern diamondback rattlesnake (*Crotalus adamanteus*)

MARK J. MARGRES, KENNETH P. WRAY, MARGARET SEAVY, JAMES J. MCGIVERN, DRAGANA SANADER and DARIN R. ROKYTA

Department of Biological Science, Florida State University, 319 Stadium Drive, Tallahassee, FL 32306, USA

Abstract

Selection can vary geographically across environments and temporally over the lifetime of an individual. Unlike geographic contexts, where different selective regimes can act on different alleles, age-specific selection is constrained to act on the same genome by altering age-specific expression. Snake venoms are exceptional traits for studying ontogeny because toxin expression variation directly changes the phenotype; relative amounts of venom components determine, in part, venom efficacy. Phenotypic integration is the dependent relationship between different traits that collectively produce a complex phenotype and, in venomous snakes, may include traits as diverse as venom, head shape and fang length. We examined the feeding system of the eastern diamondback rattlesnake (*Crotalus adamanteus*) across environments and over the lifetime of individuals and used a genotype–phenotype map approach, protein expression data and morphological data to demonstrate that: (i) ontogenetic effects explained more of the variation in toxin expression variation than geographic effects, (ii) both juveniles and adults varied geographically, (iii) toxin expression variation was a result of directional selection and (iv) different venom phenotypes covaried with morphological traits also associated with feeding in temporal (ontogenetic) and geographic (functional) contexts. These data are the first to demonstrate, to our knowledge, phenotypic integration between multiple morphological characters and a biochemical phenotype across populations and age classes. We identified copy number variation as the mechanism driving the difference in the venom phenotype associated with these morphological differences, and the parallel mitochondrial, venom and morphological divergence between northern and southern clades suggests that each clade may warrant classification as a separate evolutionarily significant unit.

Keywords: gene copy number, gene expression, ontogeny, phenotypic integration, snake venom

Received 12 December 2014; revision received 4 May 2015; accepted 8 May 2015

Introduction

Natural selection is the mechanism of adaptive evolution (Darwin 1859) and can vary geographically as a result of genotype-by-genotype-by-environment interactions (Thompson 2005), resulting in extensive phenotypic variation (Núñez *et al*. 2009; Pavelka *et al*. 2010; Kingsolver & Diamond 2011). Factors that shape the phenotype can also vary temporally over the lifetime of an individual, resulting in age-specific selective pressures (Le Corre & Kremer 2012). These selective pressures can produce ontogenetic phenotypic variation if adults and juveniles differ in phenotypic optima (Kawajiri *et al*. 2014) because age classes, like populations, can evolve divergent phenotypes to adapt to their respective environments (Le Corre & Kremer 2012). Unlike geographic contexts where selection can act on different allele frequencies, however, temporally varying selection (i.e. within the lifetime of a single individual) is constrained by acting on the same genome, even if specific alleles have very different fitness effects in

Correspondence: Darin R. Rokyta, Fax: (850) 645-8447; E-mail: drokyta@bio.fsu.edu

different age classes (Kawajiri *et al*. 2014). Therefore, ontogenetic variation in any trait is limited to age-specific expression variation.

Ontogenetic and geographic variation in the location of fitness optima should be a product of niche variation across time and space (Kingsolver & Diamond 2011; Amano *et al*. 2014). Ontogenetic niche shifts in resource use have been documented (Werner & Gilliam 1984; Ramos-Jiliberto *et al*. 2011; Jensen *et al*. 2012) and may increase foraging efficiency and growth rates and ultimately produce size-structured populations (Jensen *et al*. 2012). Qualitative and quantitative changes in prey availability can occur over an individual's lifetime due to factors such as foraging ability and gape limitation (Ramos-Jiliberto *et al*. 2011). For fishes, squamates and other gape-limited predators, the latter is critical to their feeding ecology because these predators must reach a size sufficient to consume a particular prey item whole (Adriaens *et al*. 2001; Jensen *et al*. 2012). Therefore, the size relationship between predator and prey dictates the foraging ecology of gape-limited predators.

Ontogenetic shifts in diet have been documented in venomous snakes (Shine & Sun 2003). Juveniles, constrained by gape, consume only small prey, while adults specialize on larger food items and actively refuse small prey (Shine & Sun 2003; Saviola *et al*. 2012). These differences in resource use can be coupled with ontogenetic changes in the venom phenotype (Mackessy 1988; Madrigal *et al*. 2012; Durban *et al*. 2013). Snake venoms are complex traits comprised of proteinaceous toxins that are the products of many loci and collectively function in predation and defence (Boldrini-França *et al*. 2010; Calvete *et al*. 2010; Pavlicev *et al*. 2011; Rokyta *et al*. 2011 2012 2013; Durban *et al*. 2013; Margres *et al*. 2013, 2014b). Venoms are exceptional traits for studying ontogeny because toxin expression variation directly changes the phenotype; relative amounts of venom components determine, in part, venom efficacy. Particular toxins may be present in one age class, while absent in the other, or age-specific variation in the expression of a toxin can produce different relative abundances. Expression is often measured at the transcript level even though the proteome represents the phenotype (Diz *et al*. 2012). Because venom is secreted, we can directly measure protein expression by reversed-phase high-performance liquid chromatography (RP-HPLC).

Although ontogenetic changes (Mackessy 1988; Madrigal *et al*. 2012; Durban *et al*. 2013) and intraspecific variation (Daltry *et al*. 1996; Creer *et al*. 2003; Mackessy 2010; Margres *et al*. 2014a; Sunagar *et al*. 2014) in venoms have been documented, the timing of such shifts in a life history context and a comparison of the extent of ontogenetic and geographic variation in natural populations have not been investigated. Venom is a trophic adaptation, and although venom is ultimately responsible for prey incapacitation, a suite of morphological traits such as gape and fang length should be equally important to the feeding ecology of venomous species (Adriaens *et al*. 2001). Phenotypic integration is the dependent relationship between different traits that collectively produce a complex phenotype (Santos & Cannatella 2011; Klingenberg 2014). In the case of venomous snakes, phenotypic integration may include traits as diverse as venom, head shape and fang length. The optimal depth of venom injection (i.e. fang length) may depend on venom composition which, along with head shape, may covary with prey size. Morphological differences associated with variation in venom composition, and therefore phenotypic integration of the complete feeding system, have not been investigated at any level.

Crotalus adamanteus is the largest rattlesnake species and exclusively consumes endothermic prey with rats, squirrels and rabbits comprising the majority of the diet (Klauber 1997). Although historically native to seven states in the south-eastern Coastal Plain, this species has been extirpated from Louisiana, is listed as endangered in North Carolina (Palmer & Braswell 1995) and is currently under consideration for listing as threatened under the Endangered Species Act (United States Fish and Wildlife Service 2012). Margres *et al*. (2014a) used a joint transcriptomic and proteomic approach to create a genotype–phenotype map for the venom system of *C. adamanteus* and then used this map to identify extensive protein expression variation across seven adult populations, identifying five unique phenotypes as well as the specific loci being differentially expressed across populations. In this study, we included juveniles from these same seven populations in addition to the adults analysed previously to examine the spatial–temporal dynamics of venom evolution. We used the approach of Margres *et al*. (2014a) to determine whether: (i) ontogenetic or geographic effects explain more of the variation in toxin expression, (ii) geographic variation is restricted to adults, (iii) toxin expression variation is adaptive, plastic or a product of neutral processes, and (iv) different venom phenotypes covary with morphological traits also associated with feeding in both a temporal (ontogenetic integration) and geographic (functional integration) context.

Materials and methods

Sampling

We collected venom and blood samples from 123 *Crotalus adamanteus* from seven putative populations (Fig. 1). Sixty-five of these animals were used in the analyses of

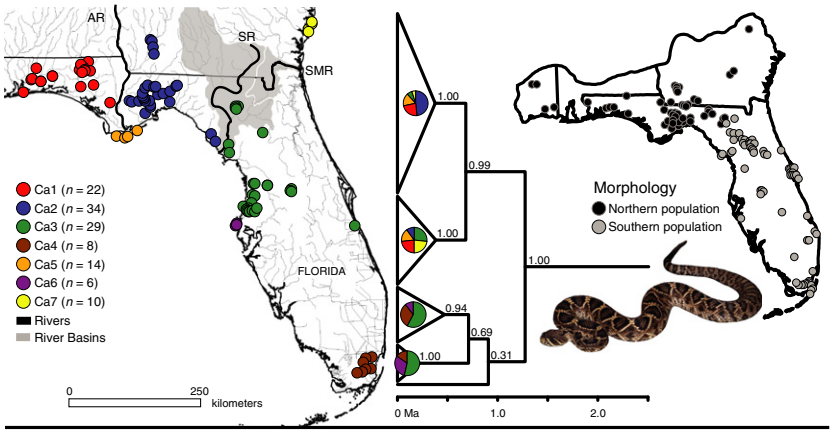

Fig. 1 Sampling for *Crotalus adamanteus*. We collected venom and blood samples from 123 *C. adamanteus* from seven putative populations; 127 preserved *C. adamanteus* specimens were used for morphological analyses. Phylogenetic analyses identified two distinct clades, one north of the Suwannee River and one south of the Suwannee River, with dating estimates placing the split at approximately 1.27 Ma. AR, Apalachicola River; Ca, *Crotalus adamanteus*; SMR, Saint Mary's River; SR, Suwannee River.

Margres *et al*. (2014a). We recorded snout–vent length (SVL) and total length (TL) for each individual. Maturity was assessed based on SVL with individuals ≥102 cm classified as adults and individuals <102 cm classified as juveniles (Waldron *et al*. 2013). Population designations were taken from Margres *et al*. (2014a) and were as follows: Ca1, Florida panhandle west of the Apalachicola River; Ca2, Florida panhandle and southern Georgia east of the Apalachicola River and west of the Suwannee River; Ca3, northern peninsular Florida; Ca4, Everglades National Park; Ca5, Little St. George Island; Ca6, Caladesi Island; and Ca7, Sapelo Island (Fig. 1). Samples were collected under the following permits: Florida Fish and Wildlife Conservation Commission (FWC) LSSC-13-00004 and LSSC-09-0399, Eglin Air Force Base 96 SFS/SFOIRP, Everglades National Park - EVER-2012-SCI-0053, Florida Department of Environmental Protection - Permit #04101310, and St. Vincent National Wildlife Refuge - Permit #41650-2012-08. The above procedures were approved by the Florida State University Institutional Animal Care and Use Committee (IACUC) under protocols #0924 and #1333.

Reversed-phase high-performance liquid chromatography

Reversed-phase high-performance liquid chromatography was performed on a Beckman System Gold HPLC (Beckman Coulter, Fullerton, CA, USA) equipped with Beckman 32 Karat Software version 8.0 for peak quantification as described by Margres *et al*. (2014a,b). Briefly, 100 μg of total protein was injected onto a Jupiter C18 column, 250 × 4.6 mm (Phenomenex, Torrence, CA, USA) using the solvent system of A = 0.1% trifluoroacetic acid (TFA) in water and B = 0.075% TFA in acetonitrile. After 5 min at 5% B, a 1% per min linear gradient of A and B was run to 25% B, followed by a 0.25% per min gradient from 25% to 65% B at a flow rate of 1 mL per min. Column effluent was monitored at 220 and 280 nm.

Venom statistical analyses

Twenty-five RP-HPLC peaks per *C. adamanteus* venom sample were quantified as previously described (Margres *et al*. 2014a,b; Wray *et al*. 2015). As discussed by Margres *et al*. (2014a), this approach produces compositional data (i.e. each individual peak represents a component of the whole) that are subject to constant-sum constraints and are inherently biased towards negative correlation among components (Aitchison 1986). Therefore, we used centred log ratio (clr) and isometric log ratio (ilr) transformations (Egozcue *et al*. 2003), when appropriate, to transform the data using the robCompositions package (Templ *et al*. 2011) in R prior to statistical analyses (Filzmoser *et al*. 2009). Because the clr transformation retains the individual identities of the peaks, this transformation was used for visualization purposes. The clr-transformed data, however, still suffer from a sum constraint because the components must add to zero. Therefore, we used the ilr transformation for statistical analyses testing for differences among venom samples because this transformation does not suffer from the same sum constraint. We treated zeros as trace values (Aitchison 1986). Margres *et al*. (2014a) performed sensitivity analyses with multiple trace values, found no differences and used a trace value of 10^{-4} in all statistical analyses. We followed this approach and performed all statistical analyses using a trace value of 10^{-4}. We also conducted an additional test for robustness using the multiplicative replacement strategy (Martin-Fernandez *et al*. 2003) implemented in the R package zCompositions assuming a detection threshold of 0.01% (the smallest measured value) and a fraction of 0.5. Results (not shown) were identical. All reported results were based on analyses using a trace value of 10^{-4}.

We used the adonis function from the vegan package (Oksanen *et al*. 2013) in R and Euclidean distances to perform a permutational or nonparametric MANOVA (McArdle & Anderson 2001) on the ilr-transformed data to test for significant venom protein expression variation as previously described (Margres *et al*. 2014a). We used a data set containing 114 *C. adamanteus* individuals, removing 26 venom samples from nine specimens to exclude duplicate samples from an individual as well as animals with unknown SVLs. These additional samples were used, however, in the principal components analysis (PCA) and a hierarchical cluster analysis. Principal components analysis was used to determine whether geographic or ontogenetic venom differences explained more of the variation in the complete *C. adamanteus* data set (e.g. 140 venom samples) as well as identify the specific RP-HPLC peaks driving this pattern. We performed a robust PCA specific to compositional data using the R package robCompositions (Templ *et al*. 2011) and the method described by Filzmoser *et al*. (2009). To determine the timing of the ontogenetic shift, we used the clr-transformed data to perform a hierarchical cluster analysis based on Euclidean distances and the average agglomeration method in the R package 'compositions' (van den Boogaart & Tolosana-Delgado 2008). Support values for the cluster analysis dendrogram were generated by multiscale bootstrap resampling (nboot = 10 000) using the R package pvclust (Suzuki & Shimodaira 2011).

We used the sample function in R to randomly permute *C. adamanteus* population designations and identify which juvenile populations exhibited unique expression patterns as described by Margres *et al*. (2014a). We used the ilr-transformed data and limited the pairwise comparisons to the four mainland populations (Ca1, Ca2, Ca3 and Ca4; Fig. 1) for juveniles ($n = 43$) because of the small sample sizes of Ca5, Ca6 and Ca7 ($n \leq 3$). We performed a linear discriminant function analysis using the lda function in R on the *C. adamanteus* ilr-transformed data to assess group membership placement probabilities across age classes and populations as previously described (Margres *et al*. 2014a). We used the predict function following the discriminant function analysis to assign age-class and population membership to the 14 samples collected from laboratory-raised *C. adamanteus* to determine the relative importance of environmental effects on toxin gene expression. All raw RP-HPLC data are in Table S1.

DNA sequencing and phylogenetics

Crotalus damanteus and *C. horridus* (outgroup) DNA was extracted from whole blood samples drawn from the caudal vein using the Omega bio-tek E.Z.N.A Tissue DNA Kit according to the manufacturer's protocol. A 1,018-bp fragment of cytochrome *b* was amplified in 25 μL PCR runs using the H16064 and L14910 primers and thermal cycling protocol described by Burbrink *et al*. (2000) for both species. A 986-bp fragment of NADH dehydrogenase subunit 5 (ND5) was amplified from *C. adamanteus* DNA in 25 μL PCR runs using sense (5′–GGT-GCA-AGT-CCA-AGT-GAT-A–3′) and antisense (5′–GGT-CTT-GTT-TTC-TGT-TTT-AGT-TA–3′) primers under the following thermal cycling protocol: 95 °C for 2 min, 35 cycles of 95 °C for 30 s, 61 °C for 30 s and 72 °C for 1 min 30 s, followed by a final extension time of 4 min at 72 °C. The same thermal cycling protocol was used for *C. horridus* using sense (5′–GGT-GCA-AAT-CCA-AGT-GAT-A–3′) and antisense (5′–GGT-TTT-GTT-TTC-TGT-CTT-AGT-AA–3′) primers. PCR products were purified using the QIagen QIAquick PCR Purification Kit, and sequencing was performed using amplification primers on the Applied Biosystems 3730 Genetic Analyzer using the Big Dye Terminator v3.1 reaction kit.

All sequences were aligned and edited using Geneious v. 5.5.7. We used the Geneious alignment algorithm for initial alignment and then adjusted manually. We translated all six sequence-reading frames into amino acids to check for stop codons and to verify the alignment. Models of nucleotide evolution were chosen using jModeltest v. 2.1.1 (Guindon & Gascuel 2003; Darriba *et al*. 2012), with the Akaike information criterion (AIC) used to determine the most appropriate model for each locus (Akaike 1974). Phylogenetic reconstruction was carried out using maximum likelihood (ML) and Bayesian inference (BI). The combined data set ML analysis was run in PAUP* 4.0b10 (Swofford 1998) using a heuristic search with 100 stepwise random-addition sequence replicates using the tree-bisection–reconnection method. To assess support for the ML tree, we also performed a nonparametric bootstrap analysis using 1000 pseudoreplicates with 10 stepwise random-addition sequence replicates. Base frequencies, rate matrix, proportion of invariable sites and shape were estimated from the data.

A Bayesian analysis and divergence time estimation were conducted using BEAST v. 1.7.4 (Drummond & Rambaut 2007). We used a random starting tree for the partitioned BEAST analysis. The divergence time estimations utilized an uncorrelated lognormal, relaxed molecular clock (Drummond *et al*. 2006), a constant size coalescent tree prior (Gernhard 2008) and two fossil-based node calibrations. A lognormal prior distribution was used on both fossil calibrations (Ho 2007) with a mean equal to 0 and a SD equal to 1.0. For the offset of each prior, we used the lower boundary of the time period the fossil was confirmed from, making this a

conservative estimate of the node age. We constrained the outgroup (*C. horridus*) to be monophyletic and calibrated the basal node of this group at 0.126 Ma using a fossil *C. horridus* from the Middle Pleistocene [Irvingtonian II of the North American land mammal age (NALMA); Holman 2000]. A fossil sample from the Early Pleistocene (Irvingtonian I NALMA) was used to calibrate the ancestral node of *C. adamanteus* at 0.781 Ma (Holman 2000). Three independent runs were conducted for 10^8 generations each, sampling every 2000 generations, and the resulting data sets were combined using LogCombiner v. 1.7.4 (Drummond & Rambaut 2007). We checked for convergence using AWTY (Nylander *et al*. 2008) and Tracer (Rambaut & Drummond 2007). Tracer was also used to check for stationarity and to determine burn-in.

Myotoxin gene copy number estimation

We followed the method of Oguiura *et al*. (2009) to estimate myotoxin gene copy number by means of qPCR for 17 *C. adamanteus*. Genomic DNA was extracted from whole blood samples as described above. Exon two was amplified from 13 ng of DNA in 20 μL qPCR runs using sense (5′-CGG-TGT-CAT-AAG-AAA-GGA-GG-3′) and antisense (5′-CAT-CTC-CAT-CGA-CAG-TCC-AT-3′) primers at 0.5 μM and the Invitrogen SYBR Green PCR Master Mix. qPCR was performed on the Applied Biosystems 7500 Fast Real-Time PCR System under the following thermal cycling protocol: 2 min at 50 °C, 10 min at 95 °C, followed by 40 cycles of 15 s at 95 °C, 30 s at 58 °C and 30 s at 70 °C. All samples were run in duplicate, and melting curve analysis was performed with a temperature gradient from 60 to 95 °C at default settings. The cycle threshold (CT), the number of cycles required for the fluorescent signal to exceed 0.05 ΔRn, was averaged across both runs for each sample. The CT mean is inversely proportional to the amount of target sequence within each sample and was compared to the proteome percentage for each sample. Raw data are provided in Table S2.

Morphological analysis

We collected data for five morphological characters on 127 preserved *C. adamanteus* specimens (Fig. 1). All specimens used in the morphological analyses were different than those used in the molecular and venom analyses. Measurements consisted of (i) snout–vent length (SVL, distance from tip of snout to posterior edge of cloaca), (ii) head length (HL, distance from tip of snout to articular-quadrate joint), (iii) head width (HW, distance across widest point of head behind eyes), (iv) interfang distance (IF, distance between fangs at maxillae) and (v) fang length (FL, distance from anterior end of maxilla to the tip of fang as folded against roof of mouth). Snout–vent length was measured to the nearest 0.5 cm using a tailor's tape, while HL, HW, IF and FL were each measured to the nearest 0.01 mm using 150 mm digital calipers (Table S3). Maturity was assessed based on SVL ≥102 cm (Waldron *et al*. 2013).

All analyses were conducted using the R statistical package v. 2.15.3 (R Development Core Team 2013). We generated histograms and conducted a Shapiro–Wilk test on the raw data to test for normality. We also conducted a Levene's test on each variable to assess the equality of variances, and then natural-log-transformed each variable.

We performed a PCA on the correlation matrix of the four, size-corrected traits in order to obtain standardized principal components (PC) and to explore the effect of shape during ontogeny. When variable loadings of PC1 are of similar magnitude, this is often interpreted as being a strong correlation between size and other morphometric variables (Jolicoeur 1963). However, as we had already removed the effect of size, we interpreted PC1 as the correlation between shape and the morphometric variables. We checked for outliers using box plots and removed these samples from all subsequent analyses. Kruskal–Wallis tests (using maturity as a factor) were then run on each variable, as well as on PC1 and PC2, to test for significant differences between juveniles and adults. The total data set was then split into two age classes (juveniles and adults) as described above (Waldron *et al*. 2013). Principal component analyses were performed on each of these age classes to explore the effect of shape among populations north and south of the Suwannee River drainage. We then used box plots to check for, and remove, outliers in each data set. Kruskal–Wallis tests were run on all four variables, as well as on the first two principal component axes, using population as the factor.

Results and discussion

Ontogenetic vs. geographic expression variation

We conducted a nonparametric MANOVA on the ilr-transformed RP-HPLC data and detected significant expression variation with respect to maturity ($P < 0.0001$) and among populations ($P < 0.0001$). The interaction between age class and population was also significant ($P = 0.013$), indicating that variation across populations differs in degree between juveniles and adults and that selection may act differently on mature and immature individuals. The R^2 values from the nonparametric MANOVA demonstrated that ontogenetic effects ($R^2 = 0.147$) explained more of the variation in

toxin expression than geographic effects (R^2 = 0.116), suggesting age-specific selective pressures may be stronger than geographically varying selection (Le Corre & Kremer 2012).

To visualize this pattern, we conducted a PCA (Fig. 2; Tables S4 and S5). The first four components explained 37.8%, 22.1%, 14.3% and 11.6% of the variance, respectively, accounting for nearly 86.0% of the variance (Tables S4 and S5). The plot of PC1 vs. PC2 (Fig. 2) revealed a clear ontogenetic separation of samples along the first PC axis, but not the second, demonstrating that ontogenetic variation in gene regulation explained the majority of the variation in our data set. The genotype–phenotype map allowed us to identify the loci underlying the phenotypic differences between age classes. The five RP-HPLC peaks with the greatest loadings along PC1 (20c, 20b, 21, 6 and 15a) correspond to ten loci (C-type lectins 9, 10, 11, 13; L-amino acid oxidase; snake venom type II metalloproteinase 1, 2; snake venom type III metalloproteinase 2; snake venom serine proteinase 4; unidentified protein(s) in peak 6), representing at least four gene families (Margres *et al*. 2014a). Snake venom type II metalloproteinase 1 (peak 21) and the unidentified protein(s) in peak 6 were upregulated in juveniles (Figs 2 and 3A; Tables S4 and S5). Peaks 15a, 20b and 20c contained multiple proteins.

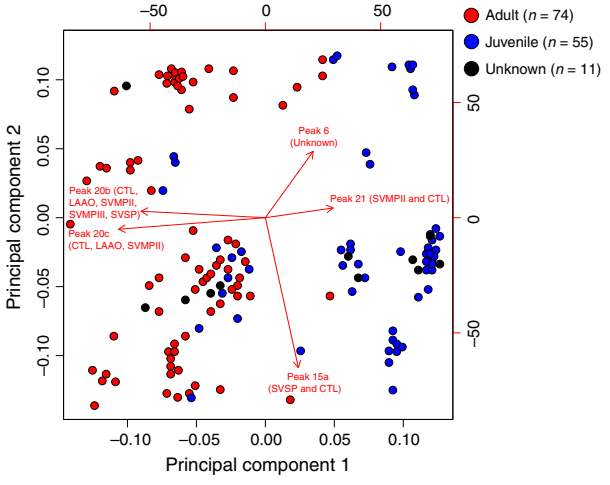

Fig. 2 Principal components analysis revealed that the differential expression of 10 loci belonging to four toxin gene families drove the ontogenetic shift in venom composition. We performed a robust PCA specific to compositional data and identified ten loci along principal component axis one that drove the ontogenetic shift in venom composition. Individual venom samples are colour-coded by maturity, with adults ≥102 cm snout–vent length. Eleven samples with known localities but unknown size were included in this exploratory analysis. CTL, C-type lectin; LAAO, L-amino acid oxidase; SVMP, snake venom metalloproteinase (types II and III); SVSP, snake venom serine proteinase.

We could not resolve which particular protein or proteins exhibited the significant variation detected across age classes within these peaks, and any number (but not necessarily all) of these seven loci were much more abundant in adults (C-type lectins 10, 11, 13; L-amino acid oxidase; snake venom type II metalloproteinase 2; snake venom type III metalloproteinase 2; snake venom serine proteinase 4).

Juvenile venoms, therefore, expressed the majority of the larger enzymes (e.g. snake venom metalloproteinases) that may aid in digestion (Mackessy 2010) at much lower levels than adults (but see McCue 2007). These toxins may be an important component in adult venoms because adults are eating large prey items and the addition of these enzymes may ensure more effective digestion (Mackessy 1988, 2010). However, these large proteins are metabolically costly to produce and may not be required for juveniles to effectively digest smaller prey. This transition may increase fitness by optimizing resource allocation, and it has been demonstrated that securing prey at low energy costs is critical to juvenile fitness (Adriaens *et al*. 2001) because fast growth rates early in life reduce the amount of time spent in smaller, more vulnerable size classes (Werner & Gilliam 1984).

The timing and mode of ontogenetic expression variation

Although fast growth rates early in life reduce the amount of time spent in more vulnerable size classes (Werner & Gilliam 1984), the expression of a specific toxin locus may not be able to be simultaneously optimal for rapid growth and feeding efficiency. As a result of this putative inherent trade-off, selection is predicted to produce weighted averages of the optimal growth phenotype and optimal feeding phenotype (Shoval *et al*. 2012). In the juvenile age class, the phenotype may be more influenced by optimal growth expression, while the opposite may be true in adults. Over time, we would expect an individual's expression pattern to move from near the growth optimum to near the feeding optimum (Shoval *et al*. 2012). These ontogenetic shifts in expression can be discrete or gradual (Jensen *et al*. 2012), and neither the type nor timing of such shifts in the context of the life history of the animal have been investigated in snake venoms.

Waldron *et al*. (2013) used SVL data to estimate maturity in *Crotalus adamanteus* and determined that 102 cm SVL was the smallest size of reproductively active individuals. We used this measurement as a threshold, classifying individuals ≥SVL 102 cm as adults and individuals <SVL 102 cm as juveniles. In addition to our nonparametric MANOVA where we

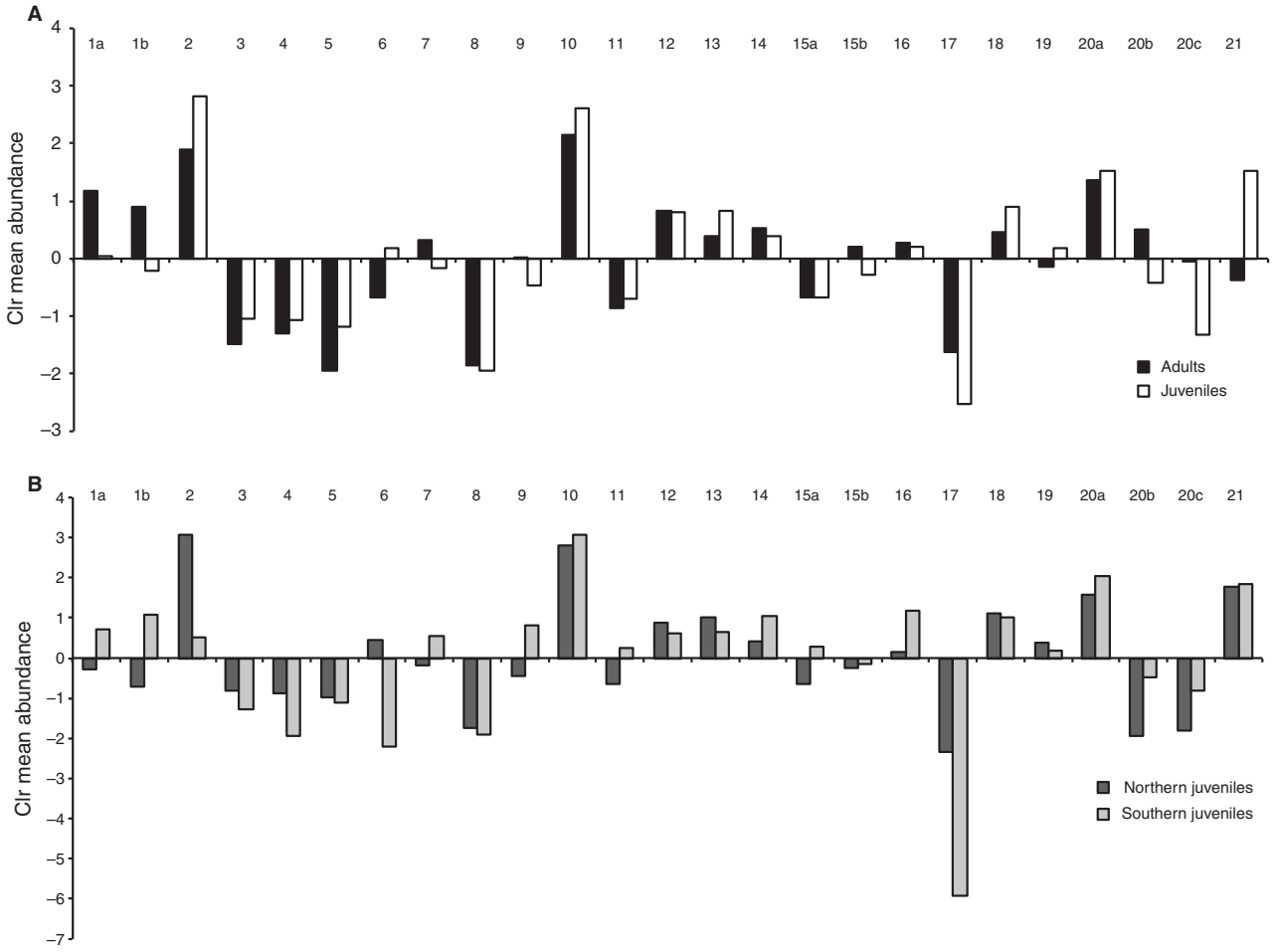

Fig. 3 Expression variation of individual toxins in *Crotalus adamanteus*. We plotted the centred log ratio mean abundance for each reversed-phase high-performance liquid chromatography peak. X-axis labels correspond to RP-HPLC peak numbers. (A) Mean abundances for juvenile and adult *C. adamanteus*. Snake venom type II metalloproteinase-1 (peak 21) was upregulated in juveniles, while C-type lectin-10, 11, L-amino acid oxidase, snake venom type II metalloproteinase-2 and type III metalloproteinase-2 (peak 20b) expression were much higher in adults. (B) Mean abundances for the two distinct juveniles populations of *C. adamanteus*; the northern population comprised populations Ca1, Ca2 and Ca3, while the southern population was Ca4. Myotoxin and peak 6 (unidentified) were much more abundant in northern populations.

used RP-HPLC data to detect significant venom protein expression variation with respect to maturity using this threshold (above), we also performed a linear discriminant function analysis to assess age-class membership placement probabilities using the ilr-transformed data. The analysis placed 89.20% of coded adults and 76.36% of coded juveniles into the correct age class, suggesting that this threshold accurately partitioned the data based on venom phenotype. This threshold, however, was based on data from a single population in South Carolina (Waldron *et al*. 2013) and may not be fixed within the species.

To determine whether there was any bias in using this threshold, we performed a hierarchical cluster analysis using all 140 venom samples from 123 individuals (Fig. 4). This method clusters samples based on compositional similarity and does not require a priori classifications as did the nonparametric MANOVA and discriminant function analysis. To first gauge the effectiveness of the clustering algorithm, we compared multiple venom samples collected from an individual animal during a single milking. Reversed-phase high-performance liquid chromatography and peak quantification for each sample were performed independently, providing replicates which allowed us to test the robustness of the algorithm. Six venom samples from KW0780, an adult male (122 cm SVL) from Ca1 (Fig. 1), and two samples from KW1517, an adult male (137.5 cm SVL) from Ca7 (Fig. 1), clustered with their respective sister samples with high support, indicating the hierarchical cluster analysis effectively identifies and clusters venoms with similar compositions (Fig. 4).

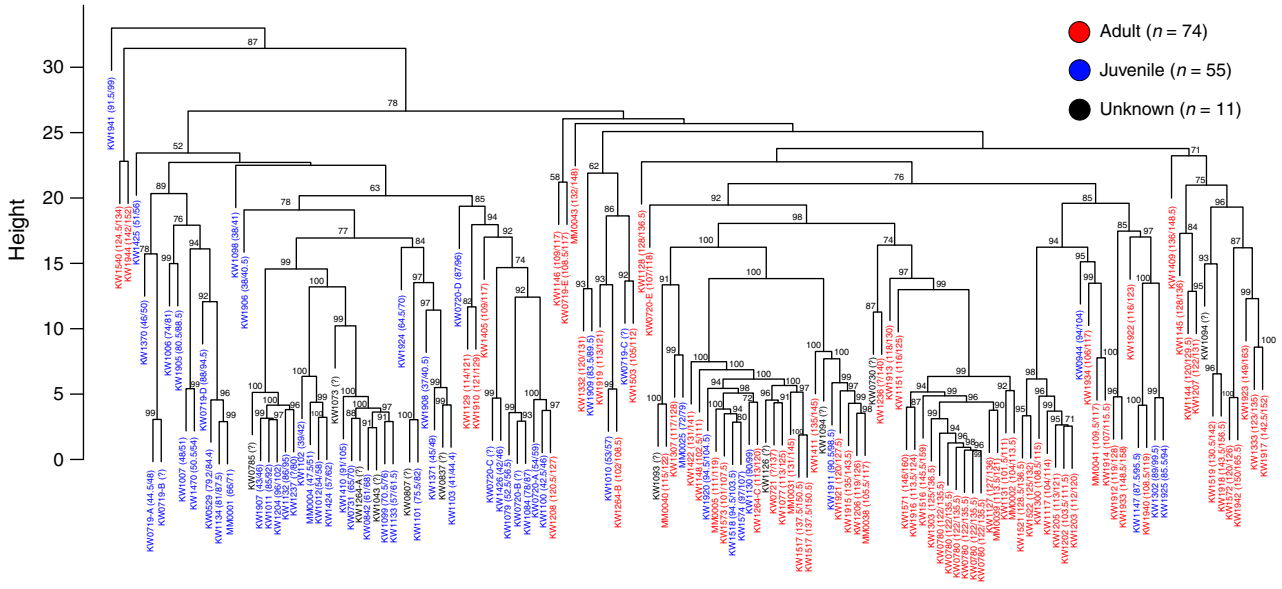

Fig. 4 A hierarchical cluster analysis of venoms revealed a distinct ontogenetic shift. We clustered 140 venom samples from 123 individuals based on compositional similarity. The cluster analysis revealed a distinct ontogenetic shift in venom composition at approximately 1 m in length, coinciding with sexual maturation. Individual venom samples are colour-coded by maturity, with adults ≥102 cm snout–vent length. Support values are approximate unbiased *P*-values calculated by multiscale bootstrap resampling. Only support values >50 are shown. Numbers in parentheses following specimen identification are snout–vent length and total length listed in cm.

The cluster analysis identified two major clusters. One cluster was dominated by individuals ≥102 cm SVL and the other was comprised mainly of snakes <102 cm SVL, indicating that the ontogenetic shift in venom composition occurred at approximately 1 m SVL and our a priori age-class data partitioning did not bias our previous analyses. Only four samples from four individuals belonging to three different populations ≥102 cm SVL were placed in the juvenile cluster, and the majority of the 13 samples from specimens <102 cm SVL placed in the adult cluster were close to this SVL cut-off (μ = 85.9 cm).

To determine whether this ontogenetic shift in venom composition was discrete or gradual, we plotted the PC1 scores (Fig. 2) against SVL for 119 individuals (Fig. 5), removing individuals with an unknown SVL as well as the duplicate samples from KW0780 and KW1517 discussed above. A linear relationship would indicate a gradual change in venom composition, and a sigmoidal curve would indicate a discrete shift. We used a nonparametric locally weighted scatterplot smoothing approach to fit a curve to our data because this method relaxes certain assumptions (e.g. linearity) of traditional regression approaches. This approach fit a sigmoidal curve to our data (Fig. 5), indicating that the ontogenetic shift in venom composition was discrete rather than gradual.

Collectively, the nonparametric MANOVA, PCA, discriminant function analysis, hierarchical cluster analysis,

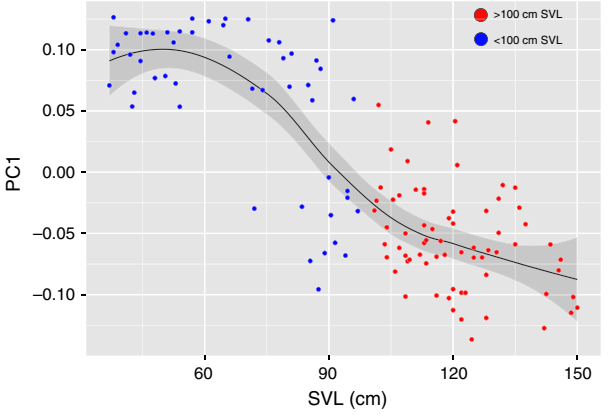

Fig. 5 The ontogenetic shift in venom composition was discrete rather than gradual. We plotted PC1 scores against SVL for 119 individuals, removing individuals with an unknown SVL as well as the duplicate samples from KW0780 and KW1517. We used the nonparametric locally weighted scatterplot smoothing approach in *R* to fit a curve to our data. This approach produced a sigmoidal curve, indicating a discrete shift in venom phenotype. The grey shading indicates confidence limits.

PC1/SVL plot and laboratory-observed ontogenetic shifts showed that the shift in toxin gene expression occurred at approximately 1 m SVL, the same size at which *C. adamanteus* reaches sexual maturity (Waldron *et al.* 2013). This shift may be a result of steroid-dependent gene regulation because the hormonal changes

associated with sexual maturation provide a mechanism that could have been co-opted to change the expression of a specific toxin locus (Kawajiri *et al*. 2014), and the shift in phenotype appears to be discrete rather than gradual (Jensen *et al*. 2012). Ontogenetic dietary shifts have been hypothesized to not be fixed within species due to variation in food availability and habitat heterogeneity (Takimoto 2003). However, we found that ontogenetic phenotypic variation in snake venom was fixed within *C. adamanteus*, suggesting that prey size may be a stronger selective pressure than prey-specific spatial effects (i.e. geographic variation in available prey species and/or population-specific co-evolutionary interactions).

Geographic expression variation in juveniles

Although our results demonstrated that ontogenetic effects explained more of the variation than geographic effects in *C. adamanteus*, pervasive geographic variation in toxin expression among adult populations has been documented (Margres *et al*. 2014a). Little if anything is known, however, about expression variation among juveniles in any species. Because the fixation of the ontogenetic shift in venom composition suggested that prey size, rather than prey-specific geographic effects, dictated venom evolution, we expected a lack of geographic variation in the juvenile phenotype because prey size for juvenile snakes is not expected to vary geographically.

We conducted a nonparametric MANOVA on the juvenile data set ($n = 49$) and detected significant geographic expression variation ($P = 0.001$), indicating that geographic variation in venom was not restricted to adults. Because of the small sample sizes in populations Ca5, Ca6 and Ca7, the randomization tests for juveniles were restricted to mainland populations Ca1, Ca2, Ca3 and Ca4 (Fig. 1; $n = 43$). The three northern, mainland juvenile populations (Ca1, Ca2 and Ca3; Fig. 1) were combined into a single population, while population Ca4 possessed a unique expression profile (Fig. 3B). We used a linear discriminant function analysis to assess group membership placement probabilities, and the analysis accurately assigned 92.31% of Ca1 individuals (formerly populations Ca1, Ca2 and Ca3; $n = 39$) and 50.00% of Ca4 individuals ($n = 4$). The low placement probability percentage for Ca4 may have been a function of unequal prior probabilities as a result of the large difference in sample sizes.

We again used the genotype–phenotype map to identify the loci underlying the phenotypic differences between the northern and southern juvenile populations. We followed the approach of Margres *et al*. (2014a) and compared the mean clr-transformed venom protein expression levels for individual peaks for both populations and identified the most variable (difference > 2) and conserved (difference < 1) loci. A difference >2 indicates that the expression of the gene in one population is more than seven times greater than the expression of the gene in the other population relative to the geometric mean, and a difference <1 indicates that the expression of the gene in one population is approximately 2.7 times greater than the expression of the gene in the other population relative to the geometric mean. Myotoxin (peak 2), an unidentified protein (peak 6), and peak 17 containing four proteins (snake venom serine proteinase-4, 7, nucleotidase and hyaluronidase) were upregulated in the northern juvenile population (Fig. 3B). Snake venom type III metalloproteinase-4 (peak 18) and snake venom type II metalloproteinase-1 (peak 21) were some of the most conserved across both juvenile populations. We also looked at the variance–covariance matrix of the clr-transformed juvenile data set to identify the most variable peaks without establishing arbitrary thresholds (Table S6). The most variable peak in juveniles was peak 6 (unidentified), accounting for 14.7% of the variance. Peaks 20b (C-type lectin 10, 11, 13, L-amino acid oxidase, snake venom type II metalloproteinase-2 and type III metalloproteinase-2, snake venom serine proteinase-4), 20c (C-type lectin-10, 11, L-amino acid oxidase, snake venom type II metalloproteinase-1, 2), 15a (C-type lectin-9, 10, snake venom serine proteinase-4), 8 (Cysteine-rich secretory protein, snake venom serine proteinase-5), 17 (hyaluronidase, nucleotidase, snake venom serine proteinase-4, 7) and 15b (C-type lectin-10, 11, 13, L-amino acid oxidase, snake venom serine proteinase-4, 7) accounted for 13.8%, 12.5%, 12.0%, 11.4%, 10.0% and 6.7% of the variance, respectively. Again, it should be noted that these peaks contained multiple proteins, and we could not resolve which particular protein or proteins exhibited the detected expression variation. We, therefore, have listed all proteins present in that peak.

Contrary to our expectations, we found significant geographic expression variation in juveniles despite the ontogenetic shift being fixed within the species. The most divergent expression patterns between northern and southern juvenile populations were in myotoxin (peak 2) and the unidentified toxin(s) in peak 6. Margres *et al*. (2014a) also documented a north–south difference in the expression of these loci in adults from the same populations, demonstrating that these differences in expression were fixed across age classes within populations. However, we cannot determine the age class upon which selection acted (i.e. the juvenile pattern may be a result of selection in adults, or the adult pattern may be a result of selection in juveniles).

Population structure

To determine whether north–south differences in toxin gene expression reflected population structure based on neutral markers, we sequenced a 1018-bp fragment of cytochrome *b* and a 986-bp fragment of NADH dehydrogenase subunit 5 (ND5) for all 123 *C. adamanteus*. A total of 125 sequences (123 *C. adamanteus* and 2 *C. horridus*) were used in the alignments. No premature codons were discovered in either alignment, strongly suggesting that these sequences were the target mitochondrial genes and not duplicated pseudogenes (Zhang & Hewitt 1996). The resulting alignment was unambiguous.

The Bayesian tree with posterior probabilities and bootstrap support values from the maximum-likelihood (ML) analysis are reported in Figs 1 and 6. The two analyses recovered nearly identical tree topologies, with a monophyletic *C. adamanteus* consisting of two clades: one made up entirely of individuals from south of the Suwannee River (Clade C, Fig. 6) and a second clade with mostly animals from north of the river (Clade B, Fig. 6). Eleven animals from south of the Suwannee River were placed into the northern clade. The southern clade and northern clade were further split into two subclades each (Figs 1 and 6). The two best trees (ln likelihood score = −3905.75) reconstructed in the ML analyses differed only in their placement of an enigmatic sample from Suwannee County, Florida (MM0031), with one tree placing this individual sister to the northern clade and the other tree placing it in a basal polytomy with the northern and southern clades. Similarly, the Bayesian analysis had poor support for

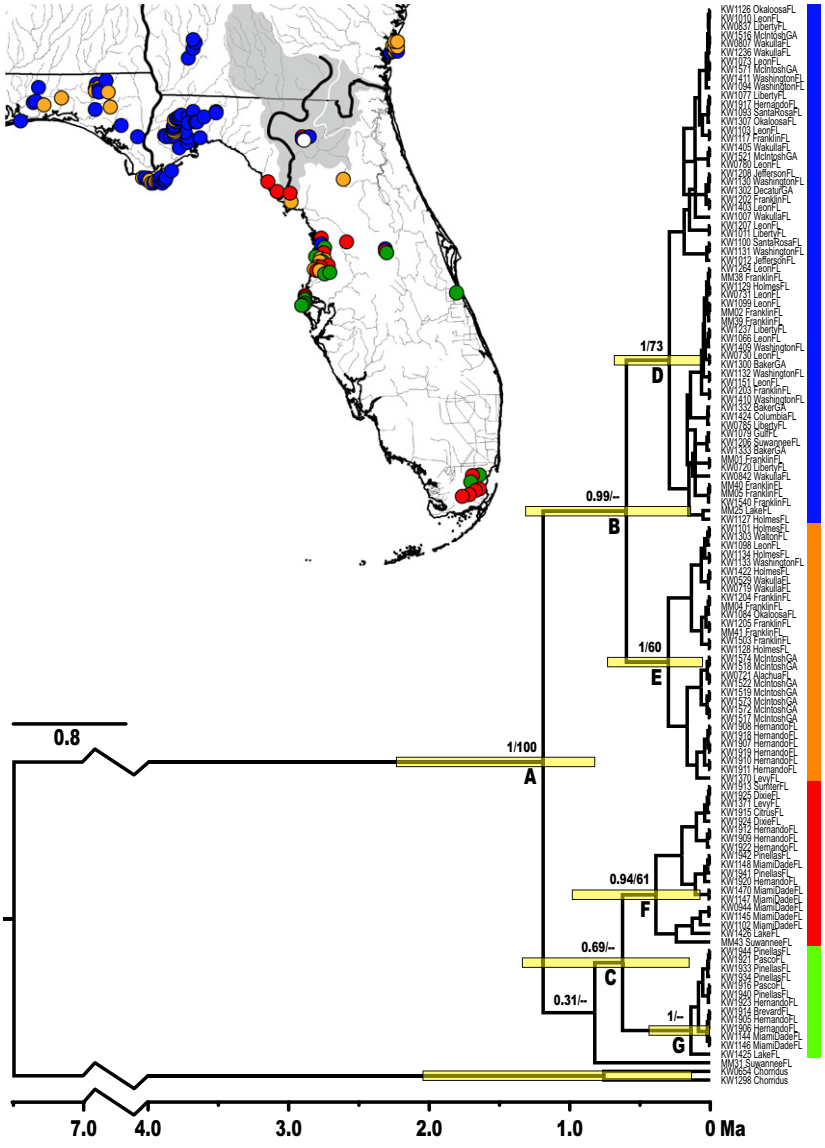

Fig. 6 Phylogenetic analysis of two mitochondrial loci identified genetic divergence between animals north and south of the Suwannee River. The Bayesian tree consisted of two clades: one made up entirely of individuals from south of the Suwannee River and a second clade with mostly animals from north of the river. The southern clade was further split into two subclades (represented by red and green bars), while the northern clade was further divided into two subclades (represented by the blue and the orange bars). One enigmatic sample (MM0031) from along the Suwannee River had poor support and is indicated by a white bar. The 95% highest posterior density credible intervals are reported as gold bars. Support values are Bayesian posterior probabilities/maximum-likelihood bootstrap values, with dashes representing bootstrap support <50. Inset map shows the geographic distribution of major clades with coloured circles corresponding to coloured bars at tips of tree. Tree scale represents substitutions/site/million years.

the placement of this sample. The multiple haplotypes within northern and southern clades were present in all populations within that clade (e.g. Ca5, a small island population with a unique venom phenotype, contained multiple northern haplotypes), indicating a lack of population structure at these neutral markers within clades (Figs 1 and 6).

The analysis of divergence time estimation yielded a median of 1.27 Ma with a 95% highest posterior density confidence interval of 0.81–2.22 Ma (herein represented as 1.27 Ma [0.81–2.22]) for the split between the northern and southern clades of *C. adamanteus* (Clade A, Fig. 6). This initial split was then followed by subsequent divergences within the southern clade (C, Fig. 6) at approximately 0.68 Ma [0.15–1.33] and the northern clade (B, Fig. 6) at approximately 0.71 Ma [0.15–1.32]. Other major divergences are reported in Table 1.

Our phylogenetic analysis identified genetic divergence between animals north and south of the Suwannee River, consistent with the differences in the expression of myotoxin and the unidentified protein(s) in both age classes. The Suwannee Straits are a known suture zone for continental and peninsular organisms (Bert 1986), and this body of water ran from the Gulf of Mexico to the Atlantic Ocean through the Okefenokee Trough, separating peninsular Florida from the mainland (Randazzo & Jones 1997). The Suwannee Straits are traditionally thought to have closed 3.8–4.2 mya, but it has been hypothesized that the Straits may have opened briefly 1.75 mya (Bert 1986), and this estimate coincides with our data and has significant conservation implications. *C. adamanteus*, due to habitat fragmentation and human persecution, is currently under consideration for listing as threatened under the Endangered Species Act (United States Fish and Wildlife Service 2012), and the mitochondrial and venom divergence between the northern and southern clade suggests that each may warrant classification as a separate evolutionarily significant unit.

Table 1 Analysis of divergence time estimation based on two mitochondrial loci

Node	Median age (Ma)	Lower age (Ma)	Upper age (Ma)
A	1.2745	0.8086	2.2201
B	0.7097	0.1503	1.3369
C	0.6822	0.1483	1.3223
D	0.2231	0.0149	0.4457
E	0.4723	0.0826	0.9915
F	0.3767	0.0739	0.6859
G	0.3834	0.0625	0.7371

Node letters correspond to labels in Fig. 6

Expression variation driven by selection

Expression variation can be a result of selection, but demonstrating local adaptation requires fitness comparisons across populations. In the absence of fitness data, comparing the differentiation of traits under putative selection to that of neutral markers across populations may allow the identification of adaptive variation (Savolainen *et al*. 2013). Expression variation among populations that can be accounted for by genetic divergence at neutral markers may be a result of genetic drift, but variation that exceeds neutral divergence is indicative of directional selection (Whitehead & Crawford 2006; Richter-Boix *et al*. 2010). To determine whether the identified geographic expression variation was a result of selection or neutral processes, we re-analysed the *C. adamanteus* expression data using the northern clade (B, Fig. 6) and southern clade (C, Fig. 6) as populations. We detected significant venom variation when comparing the northern clade to the southern clade ($P_{\text{Adults}} = 0.014, P_{\text{Juveniles}} = 0.040$), consistent with the patterns found in adults (Margres *et al*. 2014a) and juveniles (our analyses). However, we detected no venom variation among northern ($P_{\text{Adults}} = 0.349$, $P_{\text{Juveniles}} = 0.815$) or southern ($P_{\text{Adults}} = 0.673, P_{\text{Juveniles}} = 0.112$) subclades despite previous work demonstrating that multiple, unique venom phenotypes were present in each (Margres *et al*. 2014a). Failing to detect the significant expression variation previously documented among northern and southern subclades (Margres *et al*. 2014a) along with the lack of mitochondrial divergence (i.e. < 1% at either locus) indicated that the distinct venom phenotypes detected in adults by Margres *et al*. (2014a) and juveniles in our analyses were a result of positive selection over very short time periods. Venom divergence preceded neutral divergence among populations within the northern and southern clades, demonstrating that, in *C. adamanteus* venom, geography was more informative than phylogeny.

Venom expression differences were not environmentally induced

Distinguishing between environmental and genetic effects on gene expression in natural populations is difficult (Romero *et al*. 2012), and genotype-by-environment interactions have been linked to significant expression variation (Gerke *et al*. 2010; Richter-Boix *et al*. 2010). To confirm that age-related and geographic variation in *C. adamanteus* expression were a result of genetic rather than environmental effects as first discussed by Gibbs *et al*. (2009), we raised three *C. adamanteus* juveniles under identical laboratory conditions over a 3-year period.

We collected time series data from three *C. adamanteus* from population Ca2 (Fig. 1): three samples from KW1264, the specimen used in the proteomic analyses of Margres *et al*. (2014b), five samples from KW0720 and five samples from KW0719. The first four samples for KW0719 and KW0720 as well as the initial sample for KW1264 were collected when the individual animals were <102 cm SVL. All nine of these samples were assigned juvenile membership in the discriminant function analysis, while the four samples collected after each animal reached ≥102 cm SVL were assigned adult membership, revealing the ontogenetic shift in expression is discrete and under genetic control (Daltry *et al*. 1996).

We next collected venom samples from four captive, adult *C. adamanteus*: the three individuals from population Ca2 discussed above and KW0944, a specimen from population Ca4 (Fig. 1). We collected two samples from KW1264 and a single sample from all other snakes. The discriminant function analysis accurately placed three of the five adult venom samples in the correct population. Both venom samples from KW1264 and the single sample for KW0719 were correctly assigned to the large northern population comprising populations Ca1, Ca2 and Ca7 (Fig. 1) as described by Margres *et al*. (2014a). The adult sample from KW0720 was placed in Ca3 rather than the northern population described above, demonstrating intrapopulation variation. A previous discriminant function analysis accurately assigned 79.31% of the northern population adults (Margres *et al*. 2014a), and our laboratory samples are consistent with these estimates (75%). The single sample from KW0944 was predicted to belong to population Ca6 rather than population Ca4 (Fig. 1), its true population. However, both of these populations are distinct from all others in lacking myotoxin, making this misassignment unsurprising. Our results indicate that the pervasive ontogenetic and geographic expression variation identified in *C. adamanteus* was not environmentally induced and toxin gene expression is under genetic control.

Copy number variation can lead to rapid expression evolution

Relatively few proteomic studies have attempted to identify the genetic architecture underlying protein expression variation (Diz *et al*. 2012). Our genotype–phenotype map approach, however, identified candidate genes for which to test the mechanism producing the identified expression variation (Margres *et al*. 2014a). Myotoxin, a small basic peptide that causes skeletal muscle spasms following envenomation (Peigneur *et al*. 2012), was identified as the major difference between northern and southern adults (Margres *et al*. 2014a) and juveniles. This peptide was often the most abundant toxin in northern venoms and absent in southern venoms (Straight *et al*. 1991; Margres *et al*. 2014a), comprising 0.88–37.94% of adult venoms ($\mu = 15.19\%$, $\sigma = 9.70$) and 0.72–56.50% of juvenile venoms ($\mu = 28.99\%$, $\sigma = 15.20$). We found a significant correlation between myotoxin concentration and gene copy number in *C. adamanteus* ($R^2 = 0.605$, $P < 0.0001$; Fig. 7; Table S2), and a similar correlation was previously found in *C. durissus* (Oguiura *et al*. 2009). Our results suggest the variation in myotoxin abundance in the venom of *C. adamanteus* is a result of variation in gene copy number, and gene duplication/loss is one mechanism of modulating protein abundances. Despite myotoxin comprising 30.35% of the total venom in the transcriptome animal, an individual from Ca2 (Fig. 1), only a single myotoxin single nucleotide polymorphism

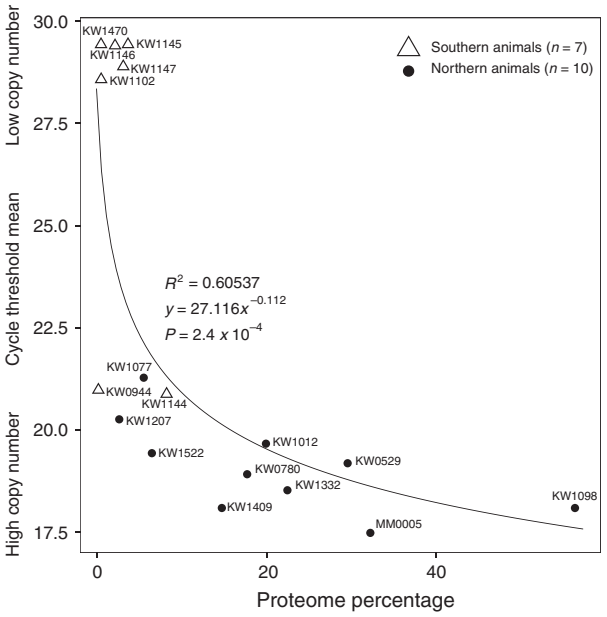

Fig. 7 Myotoxin abundance was significantly influenced by gene copy number. We used qPCR to test for a relationship between myotoxin proteome abundance and gene copy number in 17 *C. adamanteus*. We found a significant correlation between myotoxin protein concentration and myotoxin gene copy number, suggesting the abundance of myotoxin in the venom was significantly influenced by gene copy number. Proteome abundance was calculated as the area under reversed-phase high-performance liquid chromatography peak two relative to the area under all peaks. The cycle threshold mean was calculated as the average number of cycles required for the fluorescent signal to exceed a threshold of 0.05 ΔRn across both qPCR runs per sample. This measure is inversely proportional to the amount of target sequence within each sample. A power trendline was used to fit the data. Dots represent animals in the northern clade, and triangles represent animals in the southern clade in Figs 1 and 6.

was detected in the transcriptome (Margres *et al*. 2014a). Selection may have originally acted on protein abundance rather than diversification, and selection to increase the expression of a locus may be an exaptation to gene family origin and expansion.

Phenotypic integration: morphology covaries with expression

To establish whether the two largest expression differences (i.e. ontogeny and northern vs. southern animals) were integrated with morphological characters also associated with feeding, we measured head length, head width, interfang distance, fang length and snout–vent length for 127 preserved *C. adamanteus* and divided these samples into age classes and two populations based on the northern and southern clades identified in our phylogenetic analyses (Figs 1 and 6; Table S3). Using SVL as a proxy for size, we subtracted each of the four natural-log-transformed head and fang measurements from the natural-log-transformed SVL to remove the effect of size. These four variables, herein referred to as lnSVL.FL, lnSVL.HL, lnSVL.HW and lnSVL.IF, were used in all subsequent morphological analyses.

Histograms and Shapiro–Wilk tests revealed that most of the variables were not normally distributed, although natural log transformations did improve the normality of some variables. Levene's test revealed that nearly all variables (raw and transformed) were homoscedastic. Therefore, we utilized the conservative Kruskal–Wallis test because it does not assume normality. Outliers removed from analyses are highlighted in bold in Table S3.

The first two principal components of the PCA on the full data set accounted for nearly 80.0% of the variance. PC1 component loadings (Table S3) were all in the same direction and nearly identical in magnitude. As we previously accounted for the effect of body size, we interpreted PC1 to represent the effect of head shape (explaining nearly 64.0% of the variance). Component loadings for PC2 indicate an inverse relationship between lnSVL.HL and lnSVL.HW with that of lnSVL.FL and lnSVL.IF. A plot of PC1 vs. PC2 revealed a clear ontogenetic separation of samples along the first PC axis, but not the second. Using maturity as a factor, the Kruskal–Wallis tests confirmed this pattern with PC1 ($H = 54.296$, d.f. = 1, $P < 0.0001$) being significantly different between adults and juveniles, but not PC2 ($H = 0.172$, d.f. = 1, $P = 0.678$). Further Kruskal–Wallis tests using maturity as the factor revealed significant differences between adults and juveniles in lnSVL.FL ($H = 47.694$, d.f. = 1, $P < 0.0001$), lnSVL.HL ($H = 50.752$, d.f. = 1, $P < 0.0001$), lnSVL.HW ($H = 23.612$, d.f. = 1, $P < 0.0001$) and lnSVL.IF ($H = 23.716$, d.f. = 1, $P < 0.0001$), with juveniles having relatively wider and longer heads as well as longer fangs than adults (Table S3). The differences between age classes in venom phenotype, head length, head width, interfang distance and fang length demonstrate that the entire feeding system of *C. adamanteus* is ontogenetically integrated (Klingenberg 2014).

PCAs for juvenile and adult data sets showed similar results when compared to the total data set. The first two principal component analyses explained 67.19% of the variance in the juvenile data set and 77.62% of the variance in the adult data set. The component loadings for PC1 in the juvenile and adult data sets loaded in the same direction and had similar magnitudes (Table S3), and we interpreted this as the effect of head shape on the respective variances. The same pattern showing an inverse relationship between lnSVL.HL and lnSVL.HW with that of ln.SVL.FL and lnSVL.IF was found in PC2 for both data sets. In a plot of PC1 vs. PC2, we found no obvious relationship between animals north and south of the Suwannee River drainage for juveniles or adults. However, PC1 was significantly different between northern and southern juveniles ($H = 3.720$, d.f. = 1, $P = 0.050$), and PC2 was significantly different between northern and southern adults ($H = 4.143$, d.f. = 1, $P = 0.040$). Kruskal–Wallis tests using population as the factor found significant differences in lnSVL.FL in juveniles ($H = 6.482$, d.f. = 1, $P = 0.010$) and adults ($H = 7.624$, d.f. = 1, $P = 5.7 \times 10^{-3}$), but no significant differences in the other three variables (Table S3).

In both data sets, juveniles and adults in the northern clade, which generally possess high levels of myotoxin, had significantly longer fangs than juveniles and adults in the southern clade, which generally lack myotoxin (Straight *et al*. 1991; Margres *et al*. 2014a), suggesting myotoxin expression is positively correlated with fang size in *C. adamanteus*. Myotoxin causes extensive myonecrosis and hind limb paralysis (Oguiura *et al*. 2009), and increasing the depth of injection for myotoxin-rich venoms may increase the efficacy of this basic peptide by more rapidly incapacitating prey. The differences between northern and southern clades in venom phenotype and fang length demonstrate that the feeding system of *C. adamanteus* is functionally integrated (Klingenberg 2014) within both age classes, highlighting the dependent relationship between myotoxin concentration and fang length.

Conclusion

We used the genotype–phenotype map approach of Margres *et al*. (2014a,b), protein expression data and morphological data to identify ecological diversification in a complex trait both geographically and ontogenetically. Our results demonstrated that: (i) ontogenetic

effects explained more of the variation in toxin expression variation than geographic effects, (ii) both juveniles and adults vary geographically, (iii) toxin expression variation was not random but rather was a result of directional selection, and (iv) different venom phenotypes covaried with morphological traits also associated with feeding in temporal (ontogenetic integration) and geographic (functional integration) contexts. The latter are the first data, to our knowledge, demonstrating phenotypic integration between multiple morphological characters and a biochemical phenotype across populations and age classes. We identified copy number variation as a mechanism driving the difference in the venom phenotype (i.e. myotoxin concentration) associated with these geographic and ontogenetic morphological differences. The parallel differences in venom and morphology between northern and southern populations within both age classes in conjunction with the lack of genetic divergence (i.e. <1%) highlight the functional significance of these differences as well as the rapidity in which they fixed within each population. Mitochondrial, venom and morphological divergence between the northern and southern clade, however, suggests that each clade may warrant classification as a separate evolutionarily significant unit and next-gen approaches (e.g. RAD-seq) are needed to test for gene flow across this boundary.

Acknowledgements

This work was supported by the National Science Foundation (DEB 1145987 to D.R.R.). The authors thank Karalyn Aronow, Richard Bartlett, Jeffrey Fobb, Nathanael Herrera, Pierson Hill, Chip Leavine, Charles Looney, Jacob Loyacano, John Malone, Mark S. Margres, Jim Mendenhall, Moses Michelson, Flavio Morrissiey, Joe Pfaller, Michael Rochford, Nathan Schuessler, Jordan Sirosky, and Lora Smith and Jennifer Howze at the Joseph W. Jones Ecological Research Center at Ichuaway for help in acquiring venom samples. The authors thank Megan Lamb, Danielle Jones, and Jennifer Wanat with the Florida DEP and Apalachicola River NERR, Bradley Smith and Shelley Stiaes with US Fish and Wildlife Service and St. Vincent NWR, Dorset Hurley and the Sapelo Island NERR, Brett Williams and the Department of the Air Force, and Peter Krulder, Carl Calhoun, and Rick Coosey at Caladesi Island State Park for access to field sites. We thank Brian Washburn for his assistance with qPCR.

References

Adriaens D, Aerts P, Verraes W (2001) Ontogenetic shift in mouth opening mechanisms in a catfish (Clariidae, Siluriformes): a response to increasing functional demands. *European Journal of Morphology*, **247**, 197–216.

Aitchison J (1986) *The Statistical Analysis of Compositional Data*. Chapman and Hall, London, UK.

Akaike H (1974) A new look at statistical model identification. *IEEE Transactions*, **19**, 716–723.

Amano T, Freckleton R, Queenborough S *et al*. (2014) Links between plant species' spatial and temporal responses to a warming climate. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, 20133017.

Bert T (1986) Speciation in western Atlantic stone crabs (genus *Menippe*): the role of geological processes and climatic events in the formation and distribution of species. *Marine Biology*, **93**, 157–170.

Boldrini-França J, Corrêa-Netto C, Silva MMS *et al*. (2010) Snake venomics and antivenomics of *Crotalus durissus* subspecies from Brazil: assessment of geographic variation and its implication on snakebite management. *Journal of Proteomics*, **73**, 1758–1776.

van den Boogaart KG, Tolosana-Delgado R (2008) Compositions: a unified R package to analyze compositional data. *Computers & Geosciences*, **34**, 320–338.

Burbrink FT, Lawson R, Slowinski J (2000) Mitochondrial DNA phylogeography of the polytypic North American rat snake (*Elaphe obsoleta*): a critique of the subspecies concept.*Evolution*, **54**, 2107–2118.

Calvete JJ, Sanz L, Cid P *et al*. (2010) Snake venomics of the Central American rattlesnake *Crotalus simus* and the South American *Crotalus durissus* complex points to neurotoxicity as an adaptive paedomorphic trend along *Crotalus* dispersal in South America. *Journal of Proteome Research*, **9**, 528–544.

Creer S, Malhotra A, Thorpe RS, Stöcklin R, Favreau P, Chou WH (2003) Genetic and ecological correlates of intraspecific variation in pitviper venom composition detected using matrix-assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF-MS) and isoelectric focusing. *Journal of Molecular Evolution*, **56**, 317–329.

Daltry JC, Wüster W, Thorpe RS (1996) Diet and snake venom evolution. *Nature*, **379**, 537–540.

Darriba D, Taboada G, Doallo R, Posada D (2012) jmodeltest2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection*. Murray, London, UK.

Diz A, Martinez-Fernandez M, Rolan-Alvarez E (2012) Proteomics in evolutionary ecology: linking the genotype with the phenotype. *Molecular Ecology*, **21**, 1060–1080.

Drummond A, Rambaut A (2007) Beast: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Drummond A, Ho S, Phillips M, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.

Durban J, Perez A, Sanz L *et al*. (2013) Integrated "omics" profiling indicates that miRNAs are modulators of the ontogenetic venom composition shift in the Central American rattlesnake, *Crotalus simus simus*. *BMC Genomics*, **14**, 234.

Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C (2003) Isometric logratio transformations for compositional data analysis. *Journal of the International Association for Mathematical Geology*, **35**, 279–300.

Filzmoser P, Hron K, Reimann C (2009) Principal component analysis of compositional data with outliers. *Environmetrics*, **20**, 621–632.

Gerke J, Lorenz K, Ramnarine S, Cohen B (2010) Gene-environment interactions at nucleotide resolution. *PLoS Genetics*, **6**, e1001144.

Gernhard T (2008) The conditioned reconstructed process. *Journal of Theoretical Biology*, **253**, 769–778.

Gibbs HL, Sanz L, Calvete JJ (2009) Snake population venomics: proteomics-based analyses of individual variation reveals significant gene regulation effects on venom protein expression in *Sistrurus* rattlesnakes. *Journal of Molecular Evolution*, **68**, 113–125.

Guindon S, Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology*, **52**, 696–704.

Ho S (2007) Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology*, **38**, 409–414.

Holman J (2000) *Fossil Snakes of North America: Origin, Evolution, Distribution, Paleoecology*. Indiana University Press, Bloomington, Indiana.

Jensen H, Kiljunen M, Amundsen P (2012) Dietary ontogeny and niche shift to piscivory in lacustrine brown trout *Aalmo trutta* revealed by stomach content and stable isotope analyses. *Journal of Fish Biology*, **80**, 2448–2462.

Jolicoeur P (1963) The multivariate generalization of the allometry equation. *Biometrics*, **19**, 497–499.

Kawajiri M, Yoshida K, Fujimoto S *et al*. (2014) Ontogenetic stage-specific quantitative trait loci contribute to divergence in developmental trajectories of sexually dimorphic fins between medaka populations. *Molecular Ecology*, **23**, 5258–5275.

Kingsolver J, Diamond S (2011) Phenotypic selection in natural populations: what limits directional selection? *The American Naturalist*, **177,** 346–357.

Klauber LM (1997) *Rattlesnakes: Their Habits, Life Histories, and Influence on Mankind*. University of California Press, Berkeley, California.

Klingenberg C (2014) Studying morphological integration and modularity at multiple levels: concepts and analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **369**, 20130249.

Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, **21**, 1548–1566.

Mackessy SP (1988) Venom ontogeny in the Pacific rattlesnakes *Crotalus viridis helleri* and *C. v. oreganus*. *Copeia*, **1988**, 92–101.

Mackessy SP (2010) Evolutionary trends in venom composition in the Western Rattlesnakes (*Crotalus viridis* sensu lato): toxicity vs. tenderizers. *Toxicon*, **55**, 1463–1474.

Madrigal M, Sanz L, Flores-Diaz M, Sasa M, Nunez V, Alape-Giron A (2012) Snake venomics across genus *Lachesis*. Ontogenetic changes in the venom composition of *Lachesis stenophrys* and comparative proteomics of the venoms of adult *Lachesis melanocephala* and *Lachesis acrochorda*. *Journal of Proteomics*, **77**, 280–297.

Margres MJ, Aronow K, Loyacano J, Rokyta DR (2013) The venom-gland transcriptome of the eastern coral snake (*Micrurus fulvius*) reveals high venom complexity in the intragenomic evolution of venoms. *BMC Genomics*, **14**, 531.

Margres MJ, McGivern JJ, Seavy M, Wray KP, Facente J, Rokyta DR (2014a) Contrasting modes and tempos of venom expression evolution in two snake species. *Genetics*, **199**, 1, 165–176.

Margres MJ, McGivern JJ, Wray KP, Seavy M, Calvin K, Rokyta DR (2014b) Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *Journal of Proteomics*, **96**, 145–158.

Martin-Fernandez J, Barcelo-Vidal C, Pawlowsky-Glahn V (2003) Dealing with zeros and missing values in compositional data sets using nonparametric imputation. *Mathematical Geology*, **35**, 253–278.

McArdle B, Anderson M (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology*, **82**, 290–297.

McCue MD (2007) Prey envenomation does not improve digestive performance in western diamondback rattlesnakes (*Crotalus atrox*). *Journal of Experimental Zoology*, **307A**, 568–577.

Núñez V, Cid P, Sanz L *et al*. (2009) Snake venomics and antivenomics of *Bothrops atrox* venoms from Colombia and the Amazon regions of Brazil, Perú and Ecuador suggest the occurrence of geographic variation of venom phenotype by a trend towards paedomorphism. *Journal of Proteomics*, **73**, 57–78.

Nylander J, Wilgenbusch J, Warren D, Swofford D (2008) Awty (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. *Bioinformatics*, **24**, 581–583.

Oguiura N, Collares MA, Furtado MFD, Ferrarezzi H, Suzuki H (2009) Intraspecific variation of the crotamine and crotasin genes in *Crotalus durissus* rattlesnakes. *Gene*, **446**, 35–40.

Oksanen J, Blanchet F, Kindt R *et al*. (2013) vegan: Community ecology package. *R package version 2.0-7*.

Palmer WM, Braswell AL (1995) *Reptiles of North Carolina*. University of North Carolina Press, Chapel Hill, North Carolina.

Pavelka N, Rancati G, Zhu J *et al*. (2010) Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature*, **468**, 321–325.

Pavlicev M, Norgard E, Fawcett G, Cheverud J (2011) Evolution of pleiotropy: epistatic interaction pattern supports a mechanistic model underlying variation in genotype-phenotype map. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, **316**, 371–385.

Peigneur S, Orts DJB, Preito da Silva AR *et al*. (2012) Crotamine pharmacology revisited: novel insights based on the inhibition of k_v channels. *European Journal of Pharmacology: Molecular Pharmacology*, **82**, 90–96.

R Development Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Development Core Team, Vienna, Austria.

Rambaut A, Drummond A (2007) *Tracer v1.4*.

Ramos-Jiliberto R, Valdovinos F, Arias J, Alcaraz C, Garcia-Berthou E (2011) A network-based approach to the analysis of ontogenetic diet shifts: an example with an endangered, small-sized fish. *Ecological Complexity*, **8**, 123–129.

Randazzo A, Jones D (1997) *The Geology of Florida*. University Press of Florida, Gainesville, Florida.

Richter-Boix A, Teplitsky C, Rogell B, Laurila A (2010) Local selection modifies phenotypic divergence among *Rana temporaria* populations in the presence of gene flow. *Molecular Ecology*, **19**, 716–731.

Rokyta DR, Wray KP, Lemmon AR, Lemmon EM, Caudle SB (2011) A high-throughput venom-gland transcriptome for the eastern diamondback rattlesnake (*Crotalus adamanteus*) and evidence for pervasive positive selection across toxin classes. *Toxicon*, **57**, 657–671.

Rokyta DR, Lemmon AR, Margres MJ, Aronow K (2012) The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genomics*, **13**, 312.

Rokyta DR, Wray KP, Margres MJ (2013) The genesis of an exceptionally deadly venom in the timber rattlesnake (*Crotalus horridus*) revealed through comparative venom–gland transcriptomics. *BMC Genomics*, **14**, 394.

Romero I, Ruvinsky I, Gilad Y (2012) Comparative studies of gene expression and the evolution of gene regulation. *Nature Review Genetics*, **13**, 505–516.

Santos J, Cannatella D (2011) Phenotypic integration emerges from aposematism and scale in poison frogs. *Proceedings of the National Academy of Sciences USA*, **108**, 6175–6180.

Saviola A, Chiszar D, Mackessy S (2012) Ontogenetic shift in response to prey-derived chemical cues in prairie rattlesnakes *Crotalus viridis viridis*. *Current Zoology*, **58**, 549–555.

Savolainen O, Lascoux M, Merila J (2013) Ecological genomics of local adaptation. *Nature Review Genetics*, **14**, 807–820.

Shine R, Sun L (2003) Attack strategy of an ambush predator: which attributes of the prey trigger a pit-viper's strike? *Functional Ecology*, **17**, 340–348.

Shoval O, Sheftel H, Shinar G *et al*. (2012) Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. *Science*, **336**, 1157–1160.

Straight RC, Glenn JL, Wolt TB, Wolfe MC (1991) Regional differences in content of small basic peptide toxins in the venoms of *Crotalus adamanteus* and *Crotalus horridus*. *Comparative Biochemistry and Physiology Part B*, **100**, 51–58.

Sunagar K, Undheim E, Scheib H *et al*. (2014) Intraspecific venom variation in the medically significant southern pacific rattlesnake (*Crotalus oreganus helleri*): biodiscovery, clinical and evolutionary implications. *Journal of Proteomics*, **99**, 68–83.

Suzuki R, Shimodaira H (2011) *pvclust: Hierarchical Clustering with p–values via Multiscale Bootstrap Resampling*. R package version 1.2–2, http://CRAN.R-project.org/package=pvclust.

Swofford DL (1998) *Phylogenetic Analysis Using Parsimony* (PAUP*), version 4.0*. Sinauer Associates, Sunderland, Massachusetts.

Takimoto G (2003) Adaptive plasticity in ontogenetic niche shifts stabilizes consumer-resource dynamics. *The American Naturalist*, **162**, 93–109.

Templ M, Hron K, Filzmoser P (2011) *Compositional Data Analysis. Theory and Applications.*, chap. robCompositions: an R–package for robust statistical analysis of compositional data, pp. 341–355. John Wiley & Sons, Chichester, UK.

Thompson JN (2005) *The Geographic Mosaic of Coevolution*. University of Chicago Press, Chicago, Illinois.

United States Fish and Wildlife Service (2012) Endangered and threatened wildlife and plants: 90-day finding on a petition to list the eastern diamondback rattlesnake as threatened. *Federal Register*, **77**, 27403–27411.

Waldron JL, Welch SM, Bennett SH, Kalinowsky WG, Mousseau TA (2013) Life history constraints contribute to the vulnerability of a declining North American rattlesnake. *Biological Conservation*, **159**, 530–538.

Werner E, Gilliam J (1984) The ontogenetic niche and species interactions in size-structured populations. *Annual Review of Ecology and Systematics*, **15**, 319–341.

Whitehead A, Crawford D (2006) Neutral and adaptive variation in gene expression. *Proceedings of the National Academy of Sciences USA*, **103**, 5425–5430.

Wray KP, Margres MJ, Seavy M, Rokyta DR (2015) Early significant ontogenetic changes in snake venoms. *Toxicon*, **96**, 74–81.

Zhang D, Hewitt G (1996) Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology & Evolution*, **11**, 247–251.

D.R.R. and K.P.W. conceived the study. D.R.R., K.P.W. and M.J.M. designed the study. M.J.M., K.P.W., J.J.M., M.S. and D.S. performed all laboratory work. M.J.M., K.P.W. and D.R.R. analysed the data. M.J.M. and K.P.W. wrote the manuscript.

Data accessibility

Sanger sequences were submitted to the National Center for Biotechnology Information (NCBI) Trace Archive under Accession nos KJ730274-KJ730523. All data (Tables S1–S6, alignments, and the .tre file) has been deposited on Dryad under doi: 10.5061/dryad.4mt34.

Supporting information

Additional supporting information may be found in the online version of this article.

Tables S1–S6. Supporting information.