Adaptive evolution of distinct prey-specific toxin genes in rear-fanged snake venom

Cassandra M. Modahl1,†, Mrinalini2,†, Seth Frietze1,‡ and Stephen P. Mackessy1

1School of Biological Sciences, University of Northern Colorado, 501 20th Street, Greeley, CO 80639-0017, USA
2Department of Biological Sciences, National University of Singapore, Singapore 117543, Republic of Singapore

Venom proteins evolve rapidly, and as a trophic adaptation are excellent models for predator–prey evolutionary studies. The key to a deeper understanding of venom evolution is an integrated approach, combining prey assays with analysis of venom gene expression and venom phenotype. Here, we use such an approach to study venom evolution in the Amazon puffing snake, Spilotes sulphureus, a generalist feeder. We identify two novel three-finger toxins: sulditoxin and sulmotoxin 1. These new toxins are not only two of the most abundant venom proteins, but are also functionally intriguing, displaying distinct prey-specific toxicities. Sulditoxin is highly toxic towards lizard prey, but is non-toxic towards mammalian prey, even at greater than 22-fold higher dosage. By contrast, sulmotoxin 1 exhibits the reverse trend. Furthermore, evolutionary analysis and structural modelling show highest sequence variability in the central loop of these proteins, probably driving taxon-specific toxicity. This is, to our knowledge, the first case in which a bimodal and contrasting pattern of toxicity has been shown for proteins in the venom of a single snake in relation to diet. Our study is an example of how toxin gene neofunctionalization can result in a venom system dominated by one protein superfamily and still exhibit flexibility in prey capture efficacy.

1. Introduction

Snake venoms are composed of numerous proteins and peptides as a result of extensive gene duplication and neofunctionalization events [1–4]. The evolutionary advantage of this can be twofold: individual toxin genes can evolve alternate functionalities through sequence substitutions without incurring negative selection [5], and duplicate gene copies may evolve and adapt transcriptionally, via differential expression of each copy [6,7]. However, venom proteins are also rapidly evolving on a molecular scale, with many toxin superfamilies experiencing positive selection [8,9] and commonly exhibiting nucleotide substitution rates at a higher frequency in the exons rather than the intron regions, especially in those exons that code for interactive surface residues [4,10–13]. This evolutionary process results in large venom protein superfamilies with toxins that could potentially and actually perform very different biological activities.

Three-finger toxins (3FTxs), a non-enzymatic superfamily found in many snake venoms, are a prime example of toxin functional diversification. Although 3FTxs share a common structural scaffold consisting of three β-stranded ‘fingers’ or loops stabilized by four conserved disulfide bridges, they exhibit a large diversity of biological activities, including neurotoxic, cardiotoxic, cytotoxic, haemorrhagic or necrotic effects [14]. In front-fanged Elapidae (cobras, mambas, kraits, etc.) venoms, prominent 3FTxs cause neurotoxicity and respiratory paralysis that can lead to rapid death [15]. 3FTxs have also been identified in several rear-fanged venomous snakes, but they have typically not been found to impart lethal mammalian toxicity [16–19]. For example, two Boiga sp. (cat snakes) and Oxybelis fulgidus (green vinesnake) secrete 3FTxs that are highly lethal to lizards and/or birds, their primary prey, but non-lethal towards mammals [20–22]. By contrast, the most abundant
3FTx (α-cobratoxin) in the venom of the front-fanged monocoloned cobra (Naja kaouthia) is equally toxic towards both endothermic and ectothermic prey [23]. To date, only the venoms of rear-fanged snakes are known to contain 3FTxs with prey-specific effects, but it is unknown how widely-occurring this phenomenon is among venom-producing snakes within Colubridae. However, individual toxins that are prey-specific towards lizards/birds or towards mammals are not known to co-occur in the venom of a single species of rear-fanged or front-fanged snake. It is hypothesized that prey-specific toxins may co-occur, as some species use many prey types, commonly showing an ontogenetic shift in prey preference, so a two-pronged arsenal should be an adaptive trait in these instances.

Despite the fact that rear-fanged snake venoms may contain novel toxins, front-fanged snake families Elapidae and Viperidae have remained the primary focus of venom research, as they cause large-scale human morbidity and mortality [24,25]. While snakebite from several species of rear-fanged snakes (e.g. Dispholidus typus and Thelotornis sp.) have caused human fatalities [19,26,27], most rear-fanged venomous snakes rarely produce systemic envenomation in humans. This lack of serious effects explains to a large extent why venoms from most rear-fanged snake species remain unstudied [28–30]. Further, the presence of novel venom protein families in these snakes, coupled with the lack of relevant toxin sequences in public databases, present major hurdles for characterizing rear-fanged snake venoms using proteomic methods [30–33]. However, the combined use of venom gland transcriptomics and venom proteomic analysis now allows researchers to explore rear-fanged snake venom by generating a custom transcriptome database for venom peptide identification [27,34,35].

To determine if multiple novel prey-specific toxins exist within the venom of a single rear-fanged snake species, a combined transcriptomic and proteomic approach was used to characterize the currently unstudied venom from Spilotes sulphureus (formerly Pseustes sulphureus; Amazon puffing snake), a neotropical rear-fanged species [36]. Though considered by Taub (1967) to lack a Duvennoy’s venom gland [37], it does in fact possess an elongate maxilla with three enlarged rear teeth, and a large Duvennoy’s venom gland, distinct from the supralabial salivary gland (see the electronic supplementary material, figure S7a–e), with a size and location very similar to that of the well-characterized brown tree snake, Boiga irregularis ([38]; see also the electronic supplementary material, figure S7f). These snakes are semi-arboreal and feed on mammals, birds and lizards, and they do not appear to use constriction, either in the laboratory or the field [39]. They are among the largest of the advanced colubroid snakes, reaching lengths of up to 2.7 m [40]. This species has become popular in the herpetoculture trade, but nothing is known about their venom composition or production; because of their large size, they have the potential for producing high venom yields that could result in serious envenomations. We describe here the results of transcriptomic, proteomic and functional assays of S. sulphureus venom, and we uncover a venom rich in 3FTxs, several of which show differential toxicity towards specific prey types. Our study shows that S. sulphureus venom production represents a venom system of toxin gene duplication and neofunctionalization which, while dominated by one venom protein superfamily (3FTxs), exhibits an unparalleled capacity to target different prey specifically using taxon-specific chemical weapons. As a trophic adaptation, venom has the potential to provide insights into unique predator and prey evolutionary dynamics, and the venom of S. sulphureus provides a prime example of minimizing overall compositional complexity while retaining broad predatory efficacy.

2. Material and methods

(a) Rear-fanged snake venom extraction
Venom was manually extracted from all rear-fanged snakes used in this study. Snakes were maintained in the University of Northem Colorado Animal Resource Facility in accordance with UNC-IACUC protocol no. 9204 (refer to the electronic supplementary material, table S1 for snake localities). Three Amazon puffing snakes (S. sulphureus) originating from Suriname, South America were purchased from Bushmaster Reptiles. All specimens were adults, weighing over 500 g and measuring over 1500 mm snout-to-vent length. Venom was manually extracted from each individual after subcutaneous injections of ketamine-HCl (20–30 mg kg\(^{-1}\)) followed by pilocarpine-HCl (6 mg kg\(^{-1}\)) [41]. Venom was centrifuged at 9500 x g for 5 min, frozen at −80 °C, lyophilized, and stored at −20 °C until use.

(b) Toxin gene expression
Venom glands were dissected from one S. sulphureus within several hours of death from natural causes and added to TRIzol reagent (Thermo Fisher Scientific, USA) for immediate RNA isolation. To characterize toxin gene expression in the venom gland of S. sulphureus, RNA-seq was performed and the S. sulphureus venom gland transcriptome assembled de novo. Toxin sequences were identified and individual toxin transcript abundances determined (see the electronic supplementary material for RNA-seq, transcriptome assembly, annotation and quantification details).

(c) Venom phenotype
A complete venom compositional profile was generated for S. sulphureus with one-dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), size exclusion high performance liquid chromatography (SEC HPLC), and matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS). Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was performed on trypsin digested venom and individual purified toxins for protein identification using the translated venom gland transcriptome dataset as a custom reference, in addition to vertebrate protein databases. Proteomic analyses were completed to both characterize S. sulphureus venom and individual toxins, including determining protein subunit composition (see the electronic supplementary material for complete proteomic methodology). Biochemical assays for common venom enzymes were also performed (see the electronic supplementary material).

(d) Characterization of prey-specific toxins
Crude S. sulphureus venom toxicity and the toxicity of the three most abundant toxins observed within S. sulphureus venom were evaluated with LD\(_{50}\) assays using a non-model species, house gecko (Hemidactylus frenatus), and a standard model species, NSA mouse (Mus musculus). House geckos were obtained from Bushmaster Reptiles and NSA mice were bred in the UNC Animal Resource Facility (UNC IACUC protocol no. 9401). All doses (in sterile 0.9% saline) were injected intraperitoneally and three animals per dose were used. Three mice and
three lizards were injected with only 0.9% saline as controls. Doses were adjusted to individual animal body masses, with lethality expressed as μg venom or toxin per gram body mass producing 50% mortality after 24 h [42]. The nonparametric Spearman-Kärber method was used for LD₅₀ value estimations and determination of 95% confidence intervals [43]. All procedures were reviewed and approved by the UNC IACUC (protocol 1504D-SM-SMLBirds-18).

(e) Prey-specific toxin evolution in rear-fanged venomous snakes

Prey-specific monomeric 3FTxs from O. fulgidus (C0HJD3), Boiga dendrophila (Q06ZW0) and sulmotoxin 1 from S. sulphureus (this study) were submitted to the Phytre2 server [44] for modelling based on homology detection methods using server default parameters. To determine 3FTx residues available for binding, modelling was restricted to monomeric 3FTxs because the dimerization between the sulidotoxin 3FTx subunits might result in different modes of receptor binding. These three models were overlain in UCSF Chimera (http://www.rbvi.ucsf.edu/chimera), and amino acids forming the central structural loop were visualized.

Nucleotide and amino acid sequences belonging to non-conventional 3FTxs were aligned, and model selection performed for tree construction (see the electronic supplementary material for additional details). The K80 model [45] was selected by Akaiake information criterion (AIC). A maximum-likelihood tree incorporating the model was constructed in PrynML 3.1 [46], with 1000 bootstrap replicates for node support evaluation.

3. Results

(a) Toxin gene expression

Among the top 1000 expressed contigs, many highly expressed contigs were found to be venom toxins (figure 1a). The venom gland transcriptome assemblies produced 37 full-length coding DNA sequences (CDS) that were identified as venom proteins (figure 1b). These were from thirteen toxin superfamilies: 3FTxs, a cysteine rich secretory protein (CRiSP), C-type lectins (CTls), α-amino acid oxidases (LAAOs), ficolins, a snake venom metalloproteinase (SVMP), a cobra venom factor (VF), a phospholipase B (PLB), a 50 snake venom metalloproteinase (SVM), an anti-oxidant (AOX), a Kunitz-type protease inhibitor (KUN), phospholipase inhibitors, and metalloproteinase inhibitors.

Of the full-length venom protein sequences, the most abundantly expressed toxin superfamily in the S. sulphureus venom gland was 3FTx (60% of total toxin reads), followed by CTls (11%) and ficolins (8%) (figure 1c). In addition to being highly expressed, these three superfamilies also displayed the greatest diversity of isoforms, with 16 isoforms found among 3FTxs, 5 CTL isoforms and four ficolin-like isoforms (electronic supplementary material, table S2). The 3FTx isoforms were found to be highly diverse, with sequence similarities ranging from 20.7% to 91.9%.

All toxin transcript sequences were submitted to GenBank and are available at NCBI under Bioproject PRJNA448354 and accession numbers MH232964–MH233012.

(b) Venom phenotype

A venom proteome profile was constructed using a combination of SDS-PAGE, MALDI-TOF MS, SEC HPLC, and LC-MS/MS in order to determine the actively secreted venom phenotype of S. sulphureus. Reduced SDS-PAGE and MALDI-TOF MS demonstrated a high abundance of low molecular weight proteins (6000–9000 Da), within the mass range of 3FTxs observed in rear-fanged venomous snakes [20–22,34,35] (see the electronic supplementary material, results and figures S1 and S2). When using the S. sulphureus venom gland toxin translated transcriptome as a custom reference, seven venom protein families were identified with LC-MS/MS: the most abundant included 3FTxs (92%), CRiSPs (6%) and IIIF SVMs (1%), and all other toxins (cobra venom factor, LAAO, phospholipase A₂ (PLA₂) inhibitor and PLB) comprised less than 1% of the total toxin normalized spectra (figure 1d). Of the 37 full-length venom proteins that were derived from the venom gland transcriptome, peptides from LC-MS/MS mapped to 15 of these toxins (figure 1b, noted by an asterisk).

Size exclusion chromatography was used to profile crude S. sulphureus venom and to purify venom proteins for identification and biological assays, resulting in the resolution of six peaks (figure 2a). A size exclusion profile for B. irregularis venom, known to contain an abundance of 3FTxs and reduced amounts of larger toxins [47], is shown for comparison (figure 2b), and reduced and non-reduced SDS-PAGE profiles are shown for S. sulphureus peaks (figure 2c). The most abundant peaks (3–6) comprised approximately 91% of the total peak area and only contained proteins of 6000–8000 Da when reduced (figure 2c).

Peaks 3, 4 and 6 (figure 2a) were subjected to reversed phase high pressure liquid chromatography (RP-HPLC), and each produced a single purified protein product (electronic supplementary material, figure S3). Peak 3 had a MALDI-TOF mass of 17 800 Da, peak 4 a mass of 1871.9, and peak 6a, the most abundant isoform of peak 6, a mass of 7579.9 (electronic supplementary material, figure S4). Purified peaks were also digested with trypsin and subjected to LC-MS/MS analysis. Transcripts from the S. sulphureus venom gland transcriptome were used as a reference to identify these venom proteins, and peptide spectra from peaks 3, 4, and 6 all matched with high probability to 3FTx sequences (greater than 99.9%). Using LC-MS/MS spectra in combination with the S. sulphureus transcriptome reference and MALDI-TOF masses of each purified 3FTx (electronic supplementary material, figure S4), transcripts corresponding to each 3FTx were positively identified, although two transcripts (3FTx11 and 14) were possible for peak 6 (differing by one residue) (figure 3).

Peaks 3 and 4 were reduced, alkylated and then subjected to RP-HPLC to evaluate potential subunit composition (figure 2d,e). Peak 3 showed two peaks for the reduced and alkylated protein and peak 4 showed only one, demonstrating that peak 3 is a 3FTx heterodimer complex and peak 4 is a monomer. For peak 3, the 3FTx complex, reduced and alkylated subunits A and B had masses of 9338.7 Da and 9222.6 Da, respectively (electronic supplementary material, figure S5). The protein from peak 3 is only the second heterodimeric 3FTx complex that has been described from a rear-fanged snake venom and we have named it suldotoxin, derived from S. sulphureus dimeric toxin. The monomers (peaks 4 and 6) are named sulmotoxin 1 and 2, from S. sulphureus monomeric toxins. Two transcripts have been labelled sulmotoxin 2 because both are equally likely to belong to the peak 6 protein; they differ by only one amino acid residue in loop 3.
Figure 1. Venom gland transcriptome and venom proteome abundance profiles for Spilotes sulphureus. Toxin contigs are expressed at high levels (a), with three-finger toxins exhibiting the highest levels of toxin gene expression (b). Toxin transcripts for venom proteins detected within the venom via LC-MS/MS are indicated with an asterisk (*). Abundances for transcripts belonging to each venom protein superfamily are shown (c) with superfamilies making up less than 1% combined to form the ‘other’ category. Relative abundances of each superfamily within the venom proteome are also shown (d). Abbreviations: 3FTx = three-finger toxin, CRiSP = cysteine-rich secretory protein, CTL = C-type lectin, KUN = Kunitz-type protease inhibitor, LAAO = L-amino acid oxidase, PLA2_inhib = phospholipase A2 inhibitor, PLB = phospholipase B, SVMP = snake venom metalloproteinase (PIII), SVMP_inhib = snake venom metalloproteinase inhibitor, NUC = 5’ nucleotidase, VEGF = venon endothelial growth factor, and VF = cobra venom factor. (Online version in colour.)
Figure 2. Size exclusion venom profiles for Spilotes sulphureus and Boiga irregularis, with SDS-PAGE results of S. sulphureus peaks. Crude venoms (3 mg) from S. sulphureus (a) and B. irregularis (b) were fractionated using size exclusion HPLC; protein identities are based on SDS-PAGE results, N-terminal sequencing, LC-MS/MS, and enzymatic assays. Reduced and non-reduced SDS-PAGE gels are shown for the S. sulphureus chromatogram (c); crude venom (V) was also run under non-reduced conditions. HPLC chromatograms for reduced sulditoxin (d, arrows) and reduced sulmotoxin 1 (e, arrow) show the presence of two and one peaks, respectively; other peaks (approx. 5.4 & 12 min, unlabelled) are reduction chemistry peaks. Abbreviations: Mr = relative molecular masses, 3FTx = three-finger toxin, CRiSP = cysteine-rich secretory protein, and SVMP = snake venom metalloproteinase. (Online version in colour.)
For crude *S. sulphureus* venom, no detectable enzyme activity was present for many common venom enzymatic families: LAAO, acetylcholinesterase, phosphodiesterase (PDE), PLA2, thrombin-like or kallikrein-like snake venom serine proteinases. SVM activity was detectable, but at relatively low levels (0.150 ± 0.014 units; \( U = \Delta A_{342\text{nm}} \text{min}^{-1}\text{mg}^{-1}\text{ venom protein} \)). These results demonstrate that *S. sulphureus* crude venom largely lacks enzymatic activity and larger proteolytic proteins that are observed in many other snake venoms [38]. Instead this venom is heavily dominated by 3FTxs.

Despite the low levels of enzymatic activity, crude *S. sulphureus* venom was lethally toxic towards both *H. frenatus* (house geckos) and *M. musculus* (NSA mice), with considerably greater toxicity towards geckos than mice (table 1). The murine LD\(_{50}\) value for *S. sulphureus* venom (2.56 \( \mu \text{g} \text{g}^{-1} \)) was similar to values reported for moderately toxic rear-fanged venomous snakes, such as *Alsophis portoricensis* [48], although *A. portoricensis* venom was not nearly as toxic towards lizards as *S. sulphureus* venom. The higher toxicity of *S. sulphureus* venom towards geckos (1.01 \( \mu \text{g} \text{g}^{-1} \)) strongly suggests the presence of prey-specific toxins, as observed in several other rear-fanged snake venoms [20–22].

**Figure 3.** Amino acid sequence alignment of all *Spilotes sulphureus* three-finger toxin (3FTx) isoforms and other known prey-specific 3FTxs. Characterized lizard-specific 3FTxs are shown in green, and sulmotoxin 1, a mammal-specific 3FTx, is in red. All prey-specific 3FTxs sequences from this study are bolded. Sulmotoxin 2 is labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops. Amino acid variation labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops. Amino acid variation labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops. Amino acid variation labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops. Amino acid variation labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops. Amino acid variation labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops. Amino acid variation labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops. Amino acid variation labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops. Amino acid variation labelled twice because of the presence of either of these isoforms within RP-HPLC peak 6. Cysteines are highlighted in pink. The five disulfides characteristic of non-conventional 3FTxs are shown with dotted lines. The numbers above the alignments note the arrangement of the three 3FTx structural loops.

(d) Prey-specific toxin evolution in rear-fanged venomous snakes

Sulditoxin and sulmotoxins exhibit similar 3FTx structural scaffolds, and the majority of *S. sulphureus* 3FTx isoforms also contain 10 conserved cysteines, a characteristic feature of both non-conventional and other prey-specific 3FTxs (figure 3) [4,20–22,49]. The greatest amino acid residue variability between 3FTxs occurs in the central/second loop (figure 3 and electronic supplementary material, figure S6). However, sulditoxin A shares only 66% sequence identity with irditoxin A, and sulditoxin B shares only 64% identity with irditoxin B. Interestingly, sulditoxin shares the same additional cysteine residues in the first and central loops (figure 3) that are involved in the intersubunit disulphide bond for irditoxin, and these conserved cysteines probably also link the sulditoxin subunits, forming the heterodimeric complex (electronic supplementary material, figure S6A).

A maximum-likelihood tree constructed with non-conventional 3FTx nucleotide sequences from rear-fanged venomous snakes and 3FTxs from the front-fanged families Viperidae and Elapidae revealed extensive diversification of 3FTxs within rear-fanged venomous snakes (figure 4). The *S. sulphureus* 3FTx genes demonstrated species-level neofunctionalization, with both lizard- and mammal-specific 3FTxs clustering together, separate from the lizard-specific 3FTxs of *Boiga*, in spite of shared features such as the conserved intersubunit disulphide linkage forming the heterodimeric complex for both irditoxin and sulditoxin.

**Figure 2.** Predicted three-dimensional models of the two sulmotoxin molecules, including mature toxin and modelled N- and C-terminal extensions. The models were generated using the already existent 3FTx structure of *Spilotes pullatus* as a template. Sulmotoxin A is shown in green, and sulmotoxin B is shown in purple.

### (c) Characterization of prey-specific toxins

Lethal toxicity (LD\(_{50}\)) assays were performed with sulditoxin and with sulmotoxins 1 and 2 using house geckos and NSA mice (table 1). Sulditoxin was highly toxic to lizards (LD\(_{50} = 0.22\ \mu \text{g} \text{g}^{-1}\) in house geckos) and was non-toxic in mammals (LD\(_{50} >> 5 \mu \text{g} \text{g}^{-1}\) for mice). Sulmotoxin 1 exhibited the opposite trend, with a higher LD\(_{50}\) for mice (approx. 4 \( \mu \text{g} \text{g}^{-1}\)) and no toxicity towards house geckos (greater than 5 \( \mu \text{g} \text{g}^{-1}\)) (table 1). Sulmotoxin 2 was non-toxic (greater than 5 \( \mu \text{g} \text{g}^{-1}\)) towards both geckos and mice. These results demonstrate that the two most abundant 3FTxs in *S. sulphureus* venom display prey-specific toxicity, with sulditoxin exhibiting lizard-specific toxicity and sulmotoxin 1 showing mammal-specific toxicity.

### 4. Discussion

Venoms are a complex trophic adaptation used by many snakes to facilitate feeding by causing prey quiescence, immobilization, death and/or tissue degradation [38]. *Spilotes sulphureus*, like many rear-fanged snakes, produces a venom with homologues of the toxins from venoms of front-fanged snakes, and its venom is clearly lethal to both lizards and...
rodents. Unexplored venoms and venom gland transcriptomes from rear-fanged snakes are probably rich sources in which to discover functional diversification of toxin proteins. We find that the venom of *S. sulphureus* is unique, as two distinct toxins that target very different prey taxa have evolved in this species. This is, to our knowledge, the first instance in which this dual phenomenon has been discovered in any snake species.

Venom gene expression within the venom gland of *S. sulphureus* is similar to the 3FTx-dominated elapid-like venom gland transcriptome of the rear-fanged brown tree-nake *B. irregularis* [34,35]. However, the presence of many fewer 3FTx transcript isoforms in *S. sulphureus* (16) compared to *B. irregularis* (65) [35] indicates that gene duplication is much less extensive and that *S. sulphureus* venom composition is considerably less complex, and this is also corroborated by the presence of only five 3FTxs highly expressed in the venom. Fifteen of the 37 full-length venom proteins in *S. sulphureus* were detected via LC-MS/MS as translated proteins in crude venom, demonstrating that post-transcriptional regulation of specific isoforms within a single venom protein superfamily is occurring, especially in the case of 3FTxs.

Of these 3FTxs, only two showed moderate to high protein abundance: sulditoxin, a lizard-specific heterodimeric 3FTx complex, and sulmotoxin 1, a mammal-specific monomeric 3FTx. Sulditoxin is a lizard-specific toxin that is highly toxic towards house geckos, with a low LD$_{50}$ (0.22 µg g$^{-1}$). Strikingly however, we found that even at a greater than 22-fold increase in dosage, it is non-toxic towards inbred NSA mice. On the other hand, sulmotoxin 1 exhibits the opposite trend, showing higher toxicity towards mice, whereas it is non-toxic towards house geckos. These results represent a classic example of adaptation via venom gene duplication and neofunctionalization within a toxin superfamily, allowing for broader prey use.

In a previous study focused on the monomeric lizard-specific 3FTx fulgimotoxin, the amino acid sequence motifs WAVK and CYTLY in the 3FTx loop II were found to be conserved in all lizard/bird-specific 3FTxs and were identified as possible receptor binding sites involved in prey-specific toxicity [22]. Sulditoxin, which we found to be lizard-specific in this study, also shares these WAVK and CYTLY motifs in the 3FTx loop II of both subunits, supporting this hypothesis. In contrast to this, the mammal-specific sulmotoxin 1 has the sequence WTVK (A → T substitution) and CYNLY (T → N substitution). These findings are a classic illustration of sequence-based adaptive evolution of a venom gene family (3FTxs) which has occurred directly in the genome of the snake; previously, adaptive venom evolution has been reported to occur via post-genomic mechanisms such as differential regulation of expression of venom genes and venom proteins [6,7,50]. Here it is observed that amino acid substitutions, products from variations in gene sequences, have resulted in the evolution of 3FTxs that display distinct toxicities. It is interesting to note that all lizard-specific 3FTxs also have a longer second loop owing to the presence of at least two (fulgimotoxin) or four (denmotoxin, irditoxin, and sulditoxin) additional amino acids. This longer central loop, combined with an additional proline followed by a negatively charged amino acid (either aspartic acid or glutamic acid, or both), could also result in an altered structure that facilitates binding with lizard/bird nicotinic acetylcholine receptors (electronic supplementary material, figure S6b).

It appears that the mammalian-specific toxicity is a more recent state, based on gene tree topology (figure 4) and multiple sequence indels. Deletions in 3FTx second loop result in a reduced number of amino acids in this region (14 amino acids), and this is only seen for sulmotoxin 1 (mammalian specific), sulmotoxin 2A and 2b, and several other isoforms (transcript-based). Having at least 16 amino acids, and usually 18, in this central loop, as is seen in the lizard-specific 3FTxs, is a shared feature with all other nonconventional 3FTxs from rear-fanged snake species, and probably it is more similar to the ancestral state. In addition, the mammal-specific toxin is not as potent as lizard-specific 3FTxs, supporting the idea that perhaps a 3FTx with more moderate toxicity has not been under selection as long as the development of lizard-specific toxicity, which also appears to have evolved convergently in *Boiga* [21]. However, it is very difficult to perform an ancestral state reconstruction or determine time scales confidently because critical data are missing for rear-fanged snakes. It is also difficult to perform proper selection analyses because with multiple gene duplications in species, it is problematic to identify orthologous genes.

The gene tree composed of non-conventional 3FTxs from Viperidae, Elapidae and ‘Colubridae’ demonstrated diversification of 3FTxs within rear-fanged venomous snakes and neofunctionalization of these genes within species in which prey-specific toxins have evolved. It is likely that in the genus *Boiga*, lizard/bird-specific heterodimeric 3FTxs evolved once, and these novel toxins were then conserved.

| Table 1. Intraperitoneal lethal toxicity (LD$_{50}$) of *Spilotes sulphureus* venom and purified three-finger toxins. (95% confidence intervals are listed in parentheses (lower, upper), determined by Spearman-Kärber methodology.) |
|-------------------------------|------------------|------------------|
|                              | mice (*Mus musculus*) | house geckos (*Hemidactylus frenatus*) |
| crude venom (*Spilotes sulphureus*) | 2.56 (2.51, 2.61) µg g$^{-1}$ | 1.01 (0.78, 1.24) µg g$^{-1}$ |
| sulditoxin                    | nontoxic          | 0.22 (0.09, 0.53) µg g$^{-1}$ |
|                              | >5 µg g$^{-1}$    | >5 µg g$^{-1}$    |
| sulmotoxin 1                  | 4.0 (3.7, 4.3) µg g$^{-1}$ | nontoxic          |
|                              | >5 µg g$^{-1}$    | >5 µg g$^{-1}$    |
| sulmotoxin 2                  | nontoxic          | nontoxic          |
across numerous species of Boiga. The lack of toxic venom components specifically targeting mammalian prey, as demonstrated by previous studies [47], does not restrict the trophic ecology of B. irregularis owing to the use of differing predatory strategies when subduing prey: envenomation without constriction to subdue bird and lizard prey (primarily via irditoxin), and constriction alone to subdue mammalian prey [47]. Spilotes sulphureus, on the other hand, produces venom containing both lizard- and mammal-specific toxins, and adult S. sulphureus do not constrict mammalian prey, but use venom and body pinning to subdue both lizard and mammal prey (C.M. Modahl 2014, personal observation). The efficacy of this mode of feeding is greatly facilitated by the presence of different taxon-specific toxins that affect both prey types. The occurrence of prey-specific 3FTxs in the venom of S. sulphureus, a snake species native to the western hemisphere and evolutionarily distinct from Asian Boiga, emphasizes the significance of diet and predatory strategies in snake toxin evolution: both species preferentially use bird and lizard prey, but both can also take mammalian prey using different strategies (chemical versus mechanical subjugation). Sulditoxin and irditoxin therefore probably represent convergence upon a common structural solution (heterodimerism) to produce a lizard/bird-specific toxin, and the lack of close sequence identity between them supports the interpretation of functional convergence.

Relatively simple venoms with a small number of highly abundant toxins have been suggested to be characteristic of snakes with simple diets, such as sea snakes feeding on teleost fishes [51]. The sea snake Hydrophis cyanocinctus has only six major venom components, five 3FTxs and one PLA2 [51], and another sea snake species, Aipysurus laevis, has only one primary 3FTx and five PLA2s that make up the large majority of its venom [52]. Other elapids exhibit more complex venoms; even though they are often dominated by one or two venom protein superfamilies (typically 3FTxs and PLA2s), they have a greater number of individual protein isoforms within each of these families, with at least 12 3FTx isoforms, and usually more, present in venoms. This includes the diverse genera Micrurus, Naja, Ophiophagus, Dendroaspis and Bungarus [53–58]. Even the rear-fanged snake B. irregularis produces a venom with at least 25 different 3FTx isoforms, 10 of which are abundant [36]. The venom of S. sulphureus is therefore simple, similar to a sea snake, in that only five 3FTx isoforms make up 92% of the venom (based upon the HPLC chromatogram). The identification of prey-specific 3FTxs targeting lizards or mammals within the venom of S. sulphureus indicates that snakes with relatively simple venoms, such as many rear-fanged snake species, can still target a diversity of prey types. Among the 3FTx protein superfamily, a wide variety of disparate pharmacologies has evolved on a highly conserved structural scaffold [14], and within this one colubrid snake species, the venom has capitalized upon this conserved yet versatile platform to produce a simple but effective trophic adaptation. This diversified activity is in contrast to observations for the front-fanged N. kaouthia (monocled cobra), where a single abundant 3FTx (α-cobratoxin) was the only 3FTx lethal towards both lizards and mammals [23]. This study of S. sulphureus venom toxins is, to our knowledge, the first time that 3FTxs targeting different prey items have been characterized in a single venom and demonstrates the significance of venom protein gene neofunctionalization as an important and versatile trophic adaptation.

Figure 4. Non-conventional three-finger toxin (3FTx) maximum-likelihood tree showing transcript sequence divergence between the three venomous snake families (Viperidae, Elapidae and ‘Colubridae’). Prey-specific 3FTxs are coloured green for lizard-specific and red for mammal-specific. Only bootstrap node support ≥60 are displayed. (Online version in colour.)
Ethics. All experiments involving animals were subjected to prior review by the University of Northern Colorado Institutional Animal Care and Use Committee (UNC-IACUC), and protocols 9204 and 1504D-SML. Birds-18 were approved before any work was initiated. All work conforms to the ethics guidelines of the University of Northern Colorado, the International Society on Toxirology and the Society for the Study of Amphibians and Reptiles.

Data accessibility. Additional details of the analyses can be found in the electronic supplementary material, appendices S1 and S2. All toxin transcript sequences were submitted to GenBank and are available at NCBI under Bioproject PRJNA448354 and accession nos MH232864-MH233012. Correct nos.: MH233083-MH233118.

Authors’ contributions. C.M.M. and S.P.M. jointly conceived of and designed the study. C.M.M. and S.F. conducted all transcriptome experiments and C.M.M. and S.P.M. jointly conducted proteome and biochemical/biological characterization experiments. C.M.M., M. and S.F. conducted all bioinformatic analyses. C.M.M. and S.P.M. jointly wrote the original draft of the manuscript, and all authors edited the final manuscript and gave final approval for publication.

Competing interests. We have no competing interests.

Funding. The University of Northern Colorado Graduate Student Association provided a grant (C.M.M.) for research materials and the UNC Office of Research provided additional financial support (S.P.M., S.F.). The National University of Singapore (South East Asian Biodiversity Genomics grant no. R-154-000-648-646) also provided resources and support (M.). Bioinformatic research was conducted using equipment provided by NSF grant no. MRI-1429003 to S.F.

Acknowledgements. We thank Colorado HistoPrep (P. Collins, CO, USA) for providing histological preparations used in this work.

References

9. Roktya DR, Wray KP, Lemmon AR, Lemmon EM, Caudle BS. 2011 A high-throughput venom-gland transcriptome for the eastern diamondback rattlesnake (Crotalus adamanteus) and evidence for pervasive positive selection across toxin classes.


33. Ching AT et al. 2012 Venomics profiling of Thamnodynastes striatus unveils matrix metalloproteinases and other novel proteins recruited to the toxin arsenal of rear-fanged snakes. J. Proteome Res. 11, 1152 – 1162. (doi:10.1021/pr200876c)


39. Boos HEA. 2001 The snakes of Trinidad and Tobago. College station, TX: Texas A&M University Press.


