



# Three-Finger Toxin Diversification in the Venoms of Cat-Eye Snakes (Colubridae: *Boiga*)

Daniel Dashevsky<sup>1</sup> · Jordan Debono<sup>1</sup> · Darin Rokyta<sup>2</sup> · Amanda Nouwens<sup>3</sup> · Peter Josh<sup>3</sup> · Bryan G. Fry<sup>1</sup>

Received: 27 April 2018 / Accepted: 6 September 2018 / Published online: 12 September 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

The Asian genus *Boiga* (Colubridae) is among the better studied non-front-fanged snake lineages, because their bites have minor, but noticeable, effects on humans. Furthermore, *B. irregularis* has gained worldwide notoriety for successfully invading Guam and other nearby islands with drastic impacts on the local bird populations. One of the factors thought to allow *B. irregularis* to become such a noxious pest is irditoxin, a dimeric neurotoxin composed of two three-finger toxins (3FTx) joined by a covalent bond between two newly evolved cysteines. Irditoxin is highly toxic to diapsid (birds and reptiles) prey, but roughly 1000× less potent to synapsids (mammals). Venom plays an important role in the ecology of all species of *Boiga*, but it remains unknown if any species besides *B. irregularis* produce irditoxin-like dimeric toxins. In this study, we use transcriptomic analyses of venom glands from five species [*B. cynodon*, *B. dendrophila dendrophila*, *B. d. gemmicincta*, *B. irregularis* (Brisbane population), *B. irregularis* (Sulawesi population), *B. nigriceps*, *B. trigonata*] and proteomic analyses of *B. d. dendrophila* and a representative of the sister genus *Toxicodryas blandingii* to investigate the evolutionary history of 3FTx within *Boiga* and its close relative. We found that 92.5% of *Boiga* 3FTx belong to a single clade which we refer to as denmotoxin-like because of the close relation between these toxins and the monomeric denmotoxin according to phylogenetic, sequence clustering, and protein similarity network analyses. We show for the first time that species beyond *B. irregularis* secrete 3FTx with additional cysteines in the same position as both the A and B subunits of irditoxin. Transcripts with the characteristic mutations are found in *B. d. dendrophila*, *B. d. gemmicincta*, *B. irregularis* (Brisbane population), *B. irregularis* (Sulawesi population), and *B. nigriceps*. These results are confirmed by proteomic analyses that show direct evidence of dimerization within the venom of *B. d. dendrophila*, but not *T. blandingii*. Our results also suggest the possibility of novel dimeric toxins in other genera such as *Telescopus* and *Trimorphodon*. All together, this suggests that the origin of these peculiar 3FTx is far earlier than was appreciated and their evolutionary history has been complex.

**Keywords** *Boiga* · Snake · Venom · Toxins · Dimers · Guam

Daniel Dashevsky and Jordan Debono have contributed equally to this work.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00239-018-9864-6>) contains supplementary material, which is available to authorized users.

✉ Bryan G. Fry  
bgfry@uq.edu.au

<sup>1</sup> Venom Evolution Lab, School of Biological Sciences, University of Queensland, St Lucia, QLD 4072, Australia

<sup>2</sup> Department of Biological Science, Florida State University, 319 Stadium Drive, Tallahassee, FL 32306, USA

<sup>3</sup> School of Chemistry and Molecular Biosciences, University of Queensland, St Lucia, QLD 4072, Australia

## Introduction

Venom is an adaptation that is central to the ecology and evolution of the front-fanged snakes from the families Elapidae, Viperidae, and Atractaspididae (Mitchell 1861; Boquet 1979; Jackson 2003; Fry et al. 2015). Snakes lacking these prominent fangs were thought, until recently, to be non-venomous as—with a few notable exceptions—they posed little medical threat to humans (Minton 1990; Fry et al. 2003a, b, 2006). A combination of factors, including the recognition that many non-front-fanged snakes (NFFS) possess enlarged, grooved, teeth in the back part of the mouth (Fritts et al. 1994; Jackson 2003; Deufel and Cundall 2006), studies showing the homology between the venom glands of colubroid snakes (Fry et al. 2008; Jackson et al. 2016),

adoption of a more inclusive definition of venom among the research community (Fry and Wüster 2004; Fry et al. 2009a, b, 2015; Nelsen et al. 2014; Jackson and Fry 2016), and actual extraction and analysis of secretions from the oral glands of these snakes (Fry et al. 2003a, b; Lumsden et al. 2005; Pawlak et al. 2006), has led to the realization that many NFFS should be considered venomous. The species that were the subjects of these latter studies—*Coelognathus radiatus* and *Boiga dendrophila*—belong to the Colubridae subfamily Colubrinae which is thought to have originated in Africa and radiated in a wide diversity of species across Africa, Europe, Asia, Oceania, and the Americas (Burbrink and Lawson 2007; Pyron et al. 2013a; Figueroa et al. 2016). The venoms of these snakes likely influenced their evolution as they migrated across continents and exploited new prey populations.

Venom genes are found to be subject to positive selection more frequently than housekeeping genes or the physiological homologs of toxin proteins (Fry 2005; Jackson et al. 2013; Casewell et al. 2013; Sunagar et al. 2013; Vonk et al. 2013; Fry et al. 2015; Dashevsky and Fry 2018). This positive selection drives toxins to evolve rapidly through mechanisms such as the rapid accumulation of variation at exposed residues or extensive duplication and neofunctionalization (Sunagar et al. 2013; Margres et al. 2017). One of the most significant sources of these selective pressures is the highly varied niches that snakes occupy, especially the correspondingly broad range of prey species (Sunagar et al. 2013, 2014; Fry et al. 2013). While snake venom may be secondarily used in defence—such as that found in spitting cobras (Cascardi et al. 1999; Hayes et al. 2008; Panagides et al. 2017)—its primary function is to subdue prey (Fry et al. 2008, 2009b, 2013; Jackson et al. 2013; Debono et al. 2017). Coupled with the rapid rate of venom evolution, this leads to many snakes evolving toxins that are particularly potent against their specific prey items (Daltry et al. 1996; Jackson et al. 2013, 2016; Cipriani et al. 2017). Studies of prey specificity across a broad selection of snakes and prey taxa have led to the rule of thumb small non-enzymatic neurotoxins are particularly useful against diapsid prey (birds and reptiles), while synapsids (mammals) tend to be more susceptible to enzymatic coagulotoxins (Li et al. 2005; Pawlak et al. 2006; Fry et al. 2008; Jackson et al. 2016; Cipriani et al. 2017). In fact, boigatoxin-A, denmotoxin, and irditoxin, the earliest characterized *Boiga* toxins, were found to be particularly potent in diapsid (avian) tissue assays (Lumsden et al. 2005; Pawlak et al. 2006, 2008).

Boigatoxin-A, denmotoxin, and irditoxin are all three-finger toxins (3FTx), which is one of the most widespread families of snake toxins (Fry 2005; Utkin et al. 2015). The 3FTx family originated early in snake venom evolution: over 100 million years ago (Fry et al. 2003a, b, 2006, 2013; Utkin et al. 2015). While present in the venoms of

multiple snake families, this toxin type is at its highest relative abundance in the families Colubridae and Elapidae (Fry et al. 2003b, 2008). This broad trend holds true in the genus *Boiga*, with various studies finding that 3FTx are the predominant toxins in their venoms (Fry et al. 2003b; McGivern et al. 2014; Modahl and Mackessy 2016; Pla et al. 2018). The 3FTx derive their name from their shared structure: three  $\beta$ -stranded loops, or ‘fingers’, emerging from a stable, cysteine-bonded core (Utkin et al. 2015). Despite the conservation of this structure, 3FTx have evolved numerous toxic activities including  $\alpha$ -neurotoxicity,  $\kappa$ -neurotoxicity, cytotoxicity, and platelet inhibition (Fry et al. 2003b; Fry and Wüster 2004; Sunagar et al. 2013; Kessler et al. 2017). Plesiotypic 3FTx are characterized by 10 conserved cysteine residues,  $\alpha$ -neurotoxicity, and display greater affinity towards the post-synaptic nicotinic acetylcholine receptors (nAChRs) of birds and reptiles than those of rodents (Fry et al. 2003a, b; Pawlak et al. 2006, 2008; Heyborne and Mackessy 2013). Virtually all 3FTx from non-elapid snakes retain these 10 cysteines (Fry et al. 2003a, 2008; Utkin et al. 2015). Because early studies of the activity of these plesiotypic toxins used synapsid animal models (rodents) rather than diapsid models—which would have been more similar to the avian and reptilian natural prey of the snakes being studied—the 10-cysteine 3FTx were initially mistakenly referred to as “weak neurotoxins” (Utkin et al. 2001). The prevalence of such anthropocentric experimental design in investigations of the toxicity of the venoms of NFFS led to some authors concluding *Boiga* species were non-venomous (Rochelle and Kardong 1993).

Irditoxin, from the venom of *B. irregularis*, is particularly notable among 3FTx, because the mature toxin is a covalently linked dimer of two distinct 3FTx which exhibits a tenfold higher potency than the monomeric denmotoxin (Pawlak et al. 2006, 2008). Due in part to its effective venom, *B. irregularis* has become a noxious pest that is responsible for the extirpation of bird populations on several Pacific islands that were previously snake-free (Savidge 1986; Rodda and Fritts 1992). Since bird populations have crashed after their introductions, lizards formed the majority of the diet for *B. irregularis* on Guam, especially smaller individuals (Savidge 1988). Irditoxin may have played a key role in the invasion, its diapsid-specific toxicity enabling the snakes to efficiently exploit the local bird populations and transition to eating lizards once birds became scarce, but it remains unknown if irditoxin is unique to *B. irregularis* or if closely related Asian or African species also produce irditoxin-like dimers (Pyron et al. 2013a; Figueroa et al. 2016). Modahl and Mackessy (2016) suggested, based on sequence similarity, that the presence of such dimers in *B. cynodon* venom is likely. However, this study did not explore whether the newly evolved cysteines characteristic of irditoxin were present in the sequences of species outside of *B. irregularis*.

To investigate these questions, we sequenced the transcriptomes of seven populations from five species of *Boiga* ranging from the semi-terrestrial *B. trigonata* which is found in the Middle East and South Asia (Baig et al. 2011), to highly arboreal Southeast Asian species including two subspecies of *B. dendrophila*—*B. d. dendrophila* and *B. d. gemmicincta*—as well as *B. cynodon* and *B. nigriceps* (Das 2012), and populations of *B. irregularis* from Sulawesi in western Indonesia and Brisbane in eastern Australia. On separate occasions, we collected venom samples from *B. d. dendrophila* and *Toxicodryas blandingii* for proteomic investigation. These species encompass almost all the genetic, geographical, ecological, and morphological diversification within this clade of colubrine snakes. *T. blandingii* is a sister taxon to the genus *Boiga* and is prevalent within the tree canopies of sub-Saharan Africa, where this clade is suspected to have originated (Pyron et al. 2013a; Figueroa et al. 2016). *T. blandingii* is one of the largest of species in this niche, exceeding 3 m in length, and predates largely upon amphibians, birds, and lizards (Broaders and Ryan 1997). *B. d. dendrophila* is widespread in Southeast Asia and also feeds on amphibians, birds, and lizards (Pawlak et al. 2006).

## Materials and Methods

### Venom and Tissue Supplies

Tissue samples in the form of venom glands and venom samples were collected from localities, as outlined in Table 1.

### Transcriptomics

A multi-step process was used to sequence and align *Boiga* 3FTx from the SE Asian clade.

### Species Examined

To determine the sequence variability of *Boiga* 3FTx, venom glands were obtained from *Boiga cynodon*, *B. d. dendrophila*, *B. d. gemmicincta*, *B. irregularis* (Sulawesi), *B. irregularis* (Brisbane), *B. nigriceps*, and *B. trigonata*. All glands were dissected from freshly euthanized specimens 4–5 day post venom extraction, as this is the time when toxin mRNA transcriptions are being transcribed at the highest rate (Pawlak et al. 2006, 2008; Rokyta et al. 2012). *Toxicodryas blandingii* venom glands were not available and, therefore, were not included (Table 1).

### RNA Extraction and mRNA Purification

Venom-gland tissue (20 mg) was homogenized using a rotor homogenizer and total ribonucleic acid (RNA) extracted using the standard TRIzol Plus methodology (Invitrogen). RNA quality was assessed using a Nanodrop (Nanodrop 2000/2000c, v1.4.2 Spectrophotometer, Thermo Scientific, USA). mRNA was extracted and isolated using and following standard Dynabeads mRNA DIRECT Kit (Life Technologies Ambion, 1443431, Thermo Fisher Scientific, USA).

### Sequencing

Total RNA was extracted from venom glands using the standard TRIzol Plus method (Invitrogen). Extracts were enriched for mRNA using standard RNeasy mRNA mini kit (Qiagen) protocol. mRNA was reverse transcribed, fragmented, and ligated to a unique 10-base multiplex identifier (MID) tag prepared using standard protocols and sequenced on a MiSeq platform (Australian Genome Research Facility). MID reads informatically separated sequences from the other transcriptomes on the plates, which were then post-processed to remove low-quality sequences before de novo assembly into contiguous sequences (contigs).

**Table 1** Species collection locations and analyses performed

Species	Collection location	Analysis
<i>Toxicodryas blandingii</i>	Tanzania	Proteomics
<i>Boiga cynodon</i>	West Java, Indonesia	Transcriptomics
<i>Boiga dendrophila dendrophila</i>	West Java, Indonesia	Proteomics & transcriptomics
<i>Boiga dendrophila gemmicincta</i>	South Sulawesi, Indonesia	Transcriptomics
<i>Boiga irregularis</i>	Sulawesi, Indonesia	Transcriptomics
<i>Boiga irregularis</i>	Brisbane, Australia	Transcriptomics
<i>Boiga nigriceps</i>	West Java, Indonesia	Transcriptomics
<i>Boiga trigonata</i>	Pakistan	Transcriptomics

## Assembly

Illumina reads that were likely to be cross contamination between multiplexed samples were removed from our read files by identifying 57 nucleotide k-mers in our focal read set that were present in another read set from the same lane at a 1000-fold or higher level. Reads with 25% or more of their sequence represented by such k-mers were filtered from the data set. This was accomplished using Jellyfish 2.2.6 (Marçais and Kingsford 2011) and K-mer Analysis Toolkit (KAT) 2.3.4 (Mapleson et al. 2017). We then removed adaptors and low-quality bases from the reads and removed any reads shorter than 75 base pairs using Trim Galore version 0.4.3 (Krueger 2015). We then used PEAR 0.9.10 (Zhang et al. 2013) to combine pairs of reads whose ends overlapped into one, longer, merged read. We then carried out several independent de novo assemblies of these reads using the programs Extender version 1.04 (Rokyta et al. 2012), Trinity version 2.4.0 (Grabherr et al. 2011), and SOAPdenovo version 2.04 (Xie et al. 2014). SOAPdenovo was run repeatedly with k-mer sizes of 31, 75, 97, and 127. The raw reads may be found in the NCBI Sequence Read Archive under the accession number SRP155444.

## Annotation

The de novo assemblies were concatenated and searched against reference toxin sequences obtained from UniProt using BLAST version 2.7.1 (Altschul et al. 1990; Consortium 2017). We then removed all remaining contigs that did not contain complete coding sequences. Those that did were screened by visualizing the read coverage for all contigs, whose highest and lowest coverage differed by a at least factor of 10 using the Burrows–Wheeler Aligner (BWA) 0.7.16a (Li and Durbin 2009), the Genome Analysis ToolKit (GATK) 3.8 (McKenna et al. 2010), and BEDTools 2.26.0 (Quinlan and Hall 2010). Those that showed sharp discontinuities indicative of chimerical assembly were removed from the data set. In cases, where two samples contained transcripts for identical amino acid sequences, we aligned the raw reads to all the 3FTx from each sample using BWA to check if either of them was expressed at unusually low levels which might indicate contamination. In every case, the number of reads aligned to these contigs was within an order of magnitude of the highest number aligned to any toxin from that sample. Finally, we used CD-HIT version 4.7 to cluster the remaining sequences and remove duplicates (Li and Godzik 2006; Fu et al. 2012). The sequences of these contigs are available in Supplementary File 1 or at the NCBI Transcriptome Shotgun Assembly Sequence Database under the following accession numbers: GGUA00000001 (*B. cynodon*), GGUB00000001–GGUB00000043 (*B. d. dendrophila*), GGUC00000001–GGUC00000012 (*B. d. gemmicincta*),

GGUD00000001–GGUD00000019 (*B. irregularis*, Brisbane), GGUE00000001–GGUE00000014 (*B. irregularis*, Sulawesi), GGUF00000001–GGUF00000007 (*B. nigriceps*), and GGUG00000001–GGUG00000007 (*B. trigonata*).

## Proteomics

Our proteomic investigations included using a combined approach of 1D and 2D SDS-PAGE, excision of gel bands and spots, and LC–MS/MS identification of the proteins therein. Proteomic methods were performed as previously described by us (Ali et al. 2013; Debono et al. 2016, 2017). Figure 5a shows bands picked for each species for MSMS processing.

## Analysis

### Phylogenetic Reconstruction

Protein sequences for all 10-cysteine 3FTx that were available from the UniProt database were combined with the translations of our 3FTx transcripts (Consortium 2017). The sequences were aligned using a combination of manual alignment of the conserved cysteine positions and alignment using the MULTiple Sequence Comparison by Log-Expectation (MUSCLE) algorithm implemented in AliView for the blocks of sequence in between these sites (Edgar 2004; Larsson 2014). This alignment contained 282 total sequences of which 159 sequences were from *Boiga*. We reconstructed the phylogeny of these sequences using MrBayes 3.2 for 15,000,000 generations and 1,000,000 generations of burnin with lset rates = invgamma (allows rate to vary with some sites invariant and other drawn from a  $\gamma$  distribution) and prset aamodelpr = mixed (allows MrBayes to generate an appropriate amino acid substitution model by sampling from 10 predefined models) (Ronquist et al. 2012). The run was stopped when convergence values stabilized at approximate 0.013. Two replicate runs recovered virtually identical topologies. A nexus file containing the full alignment and MrBayes settings as well as the output tree can be found in the Supplementary File 1.

### Similarity Network

An all-vs-all Basic Local Alignment Search Tool (BLAST) search was conducted on the same data set of protein sequences as was used for the phylogeny with -outfmt “10 qacc sacc qcovs eval” (Altschul et al. 1990). The results of this search were filtered using a custom R script (see Supplementary File 2) to remove self-to-self results, collapse bidirectional results into one entry, and create a similarity score defined as  $-\log_{10} e$  value for each entry. Edges with coverage < 70% or  $e$  value >  $1 \times 10^{-17}$  were excluded from

the analysis and the network was created in Cytoscape 3.5.1 using the Prefuse Force Directed OpenCL Layout on the similarity scores (Shannon et al. 2003).

### Protein Clustering

Clustering was carried out using the CD-HIT 4.7 algorithm with options `-c 0.45 -n 2 -d 0 -sc 1 -g 1` (Li and Godzik 2006; Fu et al. 2012). This sets the similarity threshold of the clusters to 45% and sorts the clusters by the number of sequences they contain.

### Tests for Selection

Coding DNA sequences for denmotoxin-like sequences were compiled from GenBank (Benson et al. 2013). The sequences were trimmed to only include those codons which translate to the mature protein, translated, aligned, and reverse translated using AliView and the MUSCLE algorithm (Edgar 2004; Larsson 2014). The resulting codon alignments can be found in Supplementary Files 03–09.

Phylogenetic trees for each clade were generated from the resulting codon alignments using the same methods as described above. These tree topologies were used for all subsequent analyses.

We used several of the tests for selection implemented in HyPhy version 2.220150316beta due to their different emphases (Pond et al. 2005). The Analyze Codon Data analysis generates overall  $\omega$  values for an alignment, while the Fast Unconstrained Bayesian AppRoximation (FUBAR) method gauges the strength of consistent positive or negative selection on individual amino acids (Murrell et al. 2013). In contrast, the Mixed Effects Model of Evolution (MEME) method identifies individual sites that were subject to episodes of positive selection in the past (Murrell et al. 2012).

### Protein Modelling

Custom models for each clade of 3FTx were generated by inputting representative sequences (Clade

A—*Boiga\_irregularis\_Brisbane\_3FTx\_00*; Clade B—*Boiga\_dendrophila\_dendrophila\_3FTx\_14*; Clade C—*Boiga\_irregularis\_A0A0B8RS39*; Clade D—*Trimorphodon\_biscutatus\_A7X3S0*; Clade E—*Boiga\_dendrophila\_dendrophila\_3FTx\_28*; Clade F—*Boiga\_dendrophila\_dendrophila\_3FTx\_38*; Clade G—*Boiga\_dendrophila\_dendrophila\_3FTx\_29*) to the Phyre2 webserver using the Intensive option (Kelley et al. 2015).

Alignments of each clade were trimmed to match these structures and attribute files were created from FUBAR and MEME results. Conservation scores were calculated using the default settings of AL2CO (Pei and Grishin 2001). The structures were rendered and colored according to these attributes in UCSF Chimera version 1.10.2 (Pettersen et al. 2004).

## Results

### Transcriptomics

The summary statistics describing the concatenated transcriptomes for the various species are very similar, yet they contained highly variable numbers of unique 3FTx isoforms (Table 2). The most unusual transcriptome in terms of these statistics is that of the *B. irregularis* from Sulawesi which had a much greater portion of the reads left unpaired, more contigs, and a greater N50 length; despite these differences, the number of 3FTx isoforms in this sample roughly similar to the number from our *B. irregularis* from Brisbane. Similarly, the *B. cynodon* assembly contained only one full-length 3FTx (and fragmentary contigs indicating the likely presence of a second isoform at low expression levels) despite containing the second highest number of contigs. Our transcriptomics also confirm the presence of sequences with one or both the additional cysteines that characterize the iridotoxin dimer complex in a range of taxa beyond *B. irregularis*, where it was originally discovered (Table 3). Interestingly, we did not recover toxins with either iridotoxin-like

**Table 2** Assembly statistics for concatenated transcriptomes

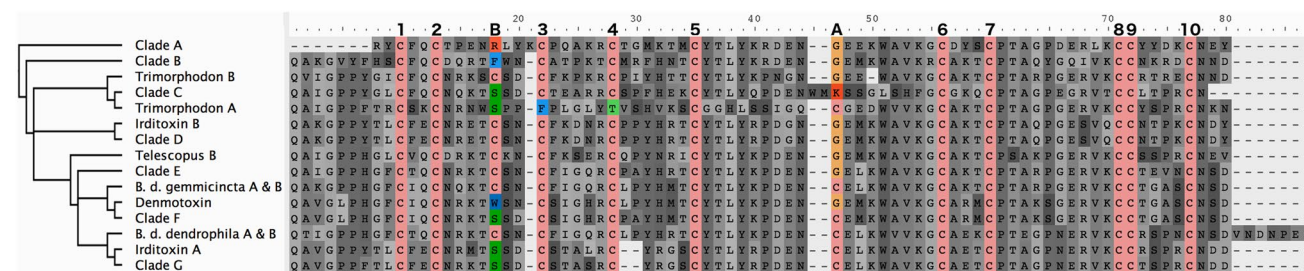
Sample	Reads (merged)	Reads (unmerged pairs)	Reads (total)	Contigs	N50 length	3FTx
<i>Boiga cynodon</i>	1,301,770	70,802	1,372,572	258,965	47,418	1
<i>B. dendrophila dendrophila</i>	821,221	41,517	862,738	203,518	42,443	43
<i>B. d. gemmicincta</i>	935,846	50,831	986,677	250,557	49,671	12
<i>B. irregularis Brisbane</i>	1,003,728	48,410	1,052,138	210,826	42,996	19
<i>B. irregularis Sulawesi</i>	8544	905,435	913,979	297,697	55,737	14
<i>B. nigriceps</i>	859,952	39,583	899,535	221,907	44,572	7
<i>B. trigonata</i>	1,023,112	38,033	1,061,145	175,745	35,975	7

**Table 3** Presence and absence of iridotoxin-like cysteines

Taxon	Iridotoxin-like B	Iridotoxin-like A
<i>Boiga cynodon</i>	✓	
<i>Boiga dendrophila dendrophila</i>	✓	✓
<i>Boiga dendrophila gemmicincta</i>	✓	✓
<i>Boiga irregularis Brisbane</i>	✓	✓
<i>Boiga irregularis Sulawesi</i>	✓	✓
<i>Boiga nigriceps</i>	✓	✓
<i>Boiga trigonata</i>		
<i>Telescopus dhara</i>	✓	
<i>Trimorphodon biscutatus</i>		✓
<i>Trimorphodon lambda</i>	✓	

cysteine from the *B. trigonata* transcriptome, even though several taxa more distantly related to *B. irregularis* contain one or both. It remains unclear if this is a genuine result or an artefact of the assembly process, especially given that *B. trigonata*'s assembly had the lowest number of contigs and N50 lengths (Table 2).

The presence of single toxins which contained the additional cysteines that characterize both the A and B subunits of iridotoxin (Fig. 1) was an unexpected finding. While it is possible that these sequences are chimeric or otherwise represent assembly error when we examined the coverage of the contigs and the sequences of the individual reads from which they were assembled, there was no evidence of assembly error. The potential structural and functional impacts of a novel cysteine pair within the same sequence on the mature toxins' structures and activity are entirely unknown.



**Fig. 1** Mature peptide sequences of notable denmotoxin-like sequences and additional representatives of each subclade with the sites of the ten canonical cysteines and iridotoxin A and B cysteines highlighted in color. These sequences include denmotoxin (a previously characterized monomeric toxin), iridotoxin A and B subunits (a previously characterized dimeric toxin), sequences with cysteines in the characteristic position of Iridotoxin A or B from *Trimorphodon* and *Telescopus*, and the two sequences which with iridotoxin-like cysteines at both positions. Phylogeny and subclades are derived from Fig. 2. Corresponding IDs for the selected sequences: Clade A—*Boiga irregularis Brisbane*\_3FTx\_00; Clade B—*Boiga den-*

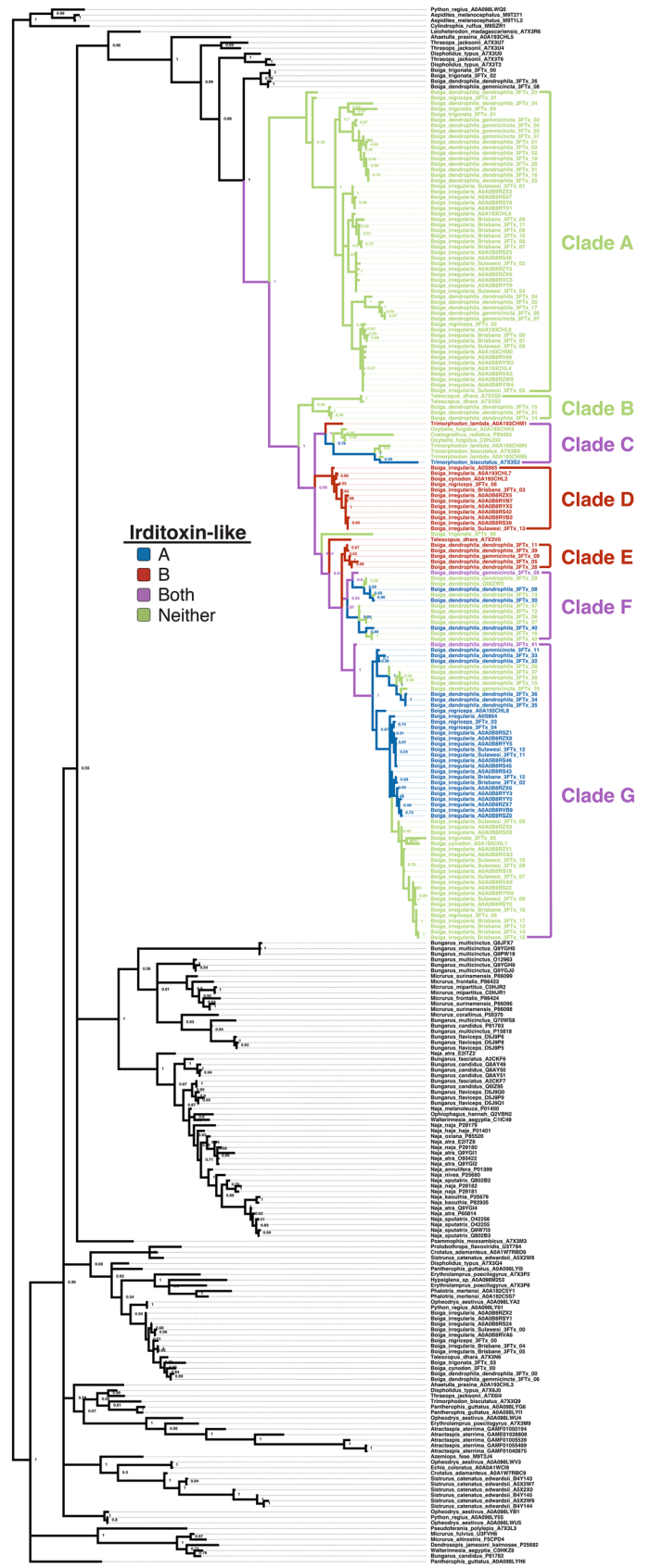
## Phylogenetics and Protein Similarity Network

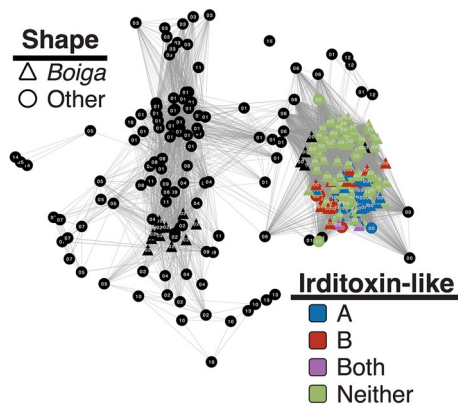
We found that 56.4% of published 10-cysteine 3FTx come from the genus *Boiga*. This is likely due in large part to publication bias: *Boiga* have long been recognized as venomous, and *B. irregularis* is one of the most notorious invasive species in the world (Minton 1990; Rodda and Savidge 2007). The other well-studied NFFS venom is that of *Dispholidus typus*, the dangerously toxic boomslang, which is primarily composed of P-III snake venom metalloproteinases and so contributes relatively few sequences to our data (Debono et al. 2017; Pla et al. 2017). Among the front-fanged snakes, the stereotypical toxins for each lineage are something besides 10-cysteine 3FTx: viperid venoms are primarily enzymatic (Mackessy 2010), elapid venoms often dominated by derived 8-cysteine 3FTx (Fry 1999), and atractaspid venoms by blood pressure acting serotoxin peptides and procoagulant enzymes (Kochva et al. 1982; Terrat et al. 2013; Oulion et al. 2018). Together, these factors make it so that there are far more 10-cysteine 3FTx known from *Boiga* than any other snake lineage. However, the sheer number of sequences from *Boiga* does indicate that the 3FTx toxin family has undergone extensive duplication during their evolution.

Of these *Boiga* toxins, 92.5% belong to a clade that we refer to as denmotoxin-like (Fig. 2). Independent protein clustering and protein similarity network analyses confirmed this subdivision within our data set (Fig. 3). Within the denmotoxin-like sequences, we designated monophyletic clades for further analysis (Figs. 1, 2): Clade A is composed exclusively of *Boiga* sequences, they retain the ancestral state of only 2 amino acids before cysteine #1 (as opposed to 9 in all other denmotoxin-like

*dendrophila\_dendrophila*\_3FTx\_14; *Trimorphodon B*—*Trimorphodon lambda\_A0A193CHM1*; Clade C—*Coelognathus radiatus\_P83490*; *Trimorphodon A*—*Trimorphodon biscutatus\_A7X3S2*; Iridotoxin B—*Boiga irregularis\_A0S865*; Clade D—*Boiga cynodon\_A0A193CHL2*; *Telescopus B*—*Telescopus dhara\_A7X3V0*; Clade E—*Boiga dendrophila gemmicincta\_3FTx\_09*; *B. d. gemmicincta A & B*—*Boiga dendrophila gemmicincta\_3FTx\_05*; Denmotoxin—*Boiga dendrophila\_Q06ZW0*; Clade F—*Boiga dendrophila\_dendrophila\_3FTx\_08*; *B. d. dendrophila A & B*—*Boiga dendrophila\_dendrophila\_3FTx\_41*; Iridotoxin A—*Boiga irregularis\_A0S864*; Clade G—*Boiga nigriceps\_3FTx\_03*. (Color figure online)

**Fig. 2** Phylogenetic tree of all publicly available 10-cysteine 3FTx and those from our transcriptome assemblies. Denmotoxin-like sequences are divided into monophyletic clades for further analyses and colored according to their iridotoxin-like cysteines. (Color figure online)





**Fig. 3** Protein similarity network based on BLAST  $e$  values. Labels correspond to CD-HIT clusters. Both the network and the clusters confirm the distinction between the denmotoxin-like sequences and others. Denmotoxin-like sequences are colored according to their iriditoxin-like cysteines. (Color figure online)

clades), and none possess either of the derived iriditoxin-like cysteines; Clade B contains *Boiga* and *Telescopus* sequences none of which possess either of the iriditoxin-like cysteines; Clade C contains sequences from *Coelognathus*, *Oxybelis*, and *Trimorphodon*, the latter of which includes toxins with the iriditoxin-like A and B cysteines separately (though the *T. biscutatus* sequence also lacks the canonical cysteines #2 and #3); Clades D and E both contain only *Boiga* sequences with the iriditoxin-like B cysteine; and Clades F and G are also composed entirely of *Boiga* sequences, but contain a mix of those with the iriditoxin-like A cysteine, neither, and one sequence each with both the iriditoxin-like A and B cysteines. Clades E, F, and G are part of a polytomy that is separated from Clade D by two monotypic branches including a *Telescopus* sequence with the iriditoxin-like B cysteine.

## Signals of Selection

We examined the signals of selections on the various clades of denmotoxin-like sequences using whole gene as well as site-specific algorithms (Table 4). We found Clade D to be subject to negative selection, Clade C neutral, Clades A, E, F, and G positive, and Clade B extreme positive selection. However, Clades B, C, and E included few sequences which increases the potential error and decreases the statistical power of these estimates. This is illustrated by the fact that, despite Clade B's extreme  $\omega$  value across the whole gene, our site-specific analyses found relatively few sites that met their statistical significance thresholds. A possible reason for the high  $\omega$  value in Clade B is because the taxa within it span a broad range of the colubrines. Though many sites failed to reach statistical significance, the pattern of estimated selection across the proteins for each clade can be seen in Fig. 4 (full size images for each clade can be found in Supplementary Figs. 1–7).

## Proteomics

The combined 1D and 2D gel approach revealed that both *T. blandingii* and *B. d. dendrophila* were dominated by 3FTx (Fig. 5). Comparison of reduced and non-reduced *B. d. dendrophila* 1D gels revealed the presence of a band in the non-reduced gel between the 15 and 20 kDa markers (Fig. 5a), while in the reduced gel, it was apparent that the monomeric 3FTx band sitting just below the 10 kDa marker was correspondingly thicker and LC-MSMS identified the additional band as 3FTx (Fig. 5a, Supplementary Table 2). This confirmed the presence of a dimeric complex made up of two 3FTx. Homologous bands were either absent entirely from the *T. blandingii* venom or was too faint to detect in both the reduced and non-reduced gels (Fig. 5a). The 2D gels (Fig. 5b, c) further demonstrate the diversity of 3FTx isoforms found in both venoms.

**Table 4** Tests of selection

Clade	Sequences	$\omega$	FUBAR (–) <sup>a</sup>	FUBAR (+) <sup>b</sup>	MEME <sup>c</sup>	FUBAR (+) & MEME <sup>d</sup>
A	54	1.25	5	12	18	11
B	5	3.02	0	4	1	0
C	6	0.97	4	1	5	1
D	12	0.67	2	0	1	0
E	5	1.42	0	2	0	0
F	13	1.4	4	6	6	2
G	54	1.45	4	16	14	11

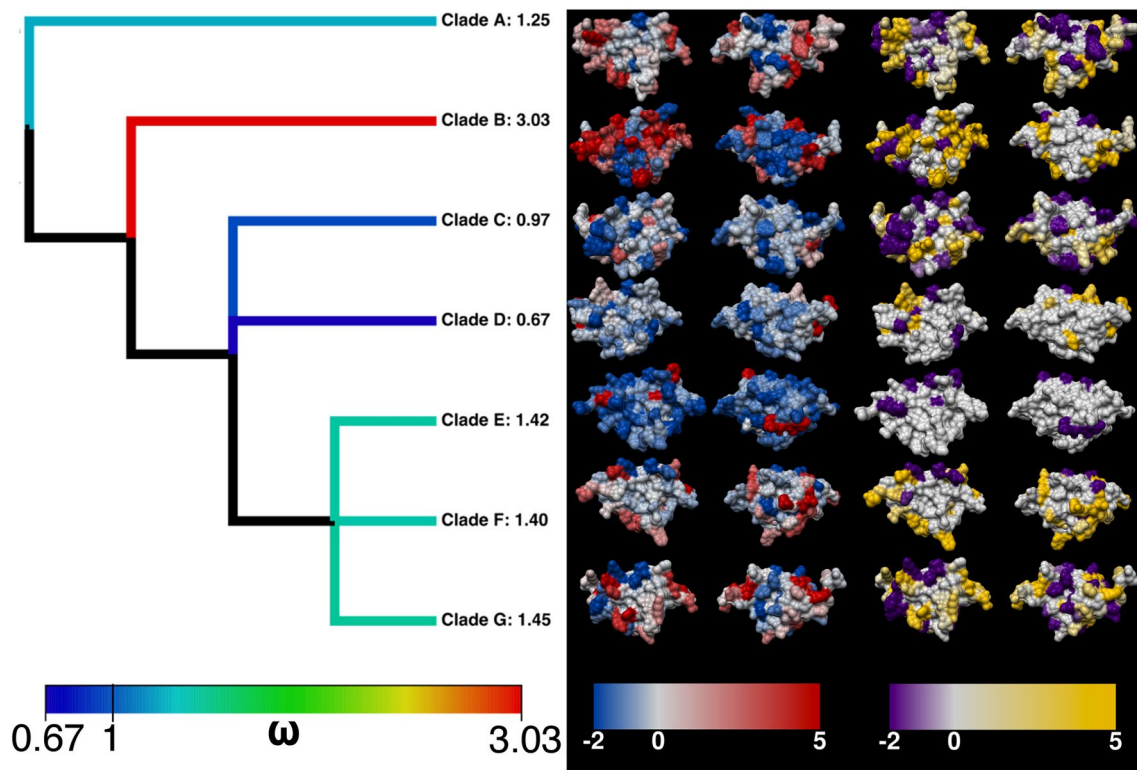
<sup>a</sup>Number of codons under negative selection according to FUBAR

<sup>b</sup>Number of codons under positive selection according to FUBAR

<sup>c</sup>Number of codons under episodic positive selection according to MEME

<sup>d</sup>Number of codons that fit criteria <sup>b</sup> and <sup>c</sup>





**Fig. 4** Schematic phylogeny of the clades within the denmotoxin-like 3FTx. Branches for each clade are colored and tips are labeled according to  $\omega$  values. Protein models show front and back views

colored according to the estimated strength of selection ( $\beta - \alpha$ ) from FUBAR (left) and MEME (right). (Color figure online)

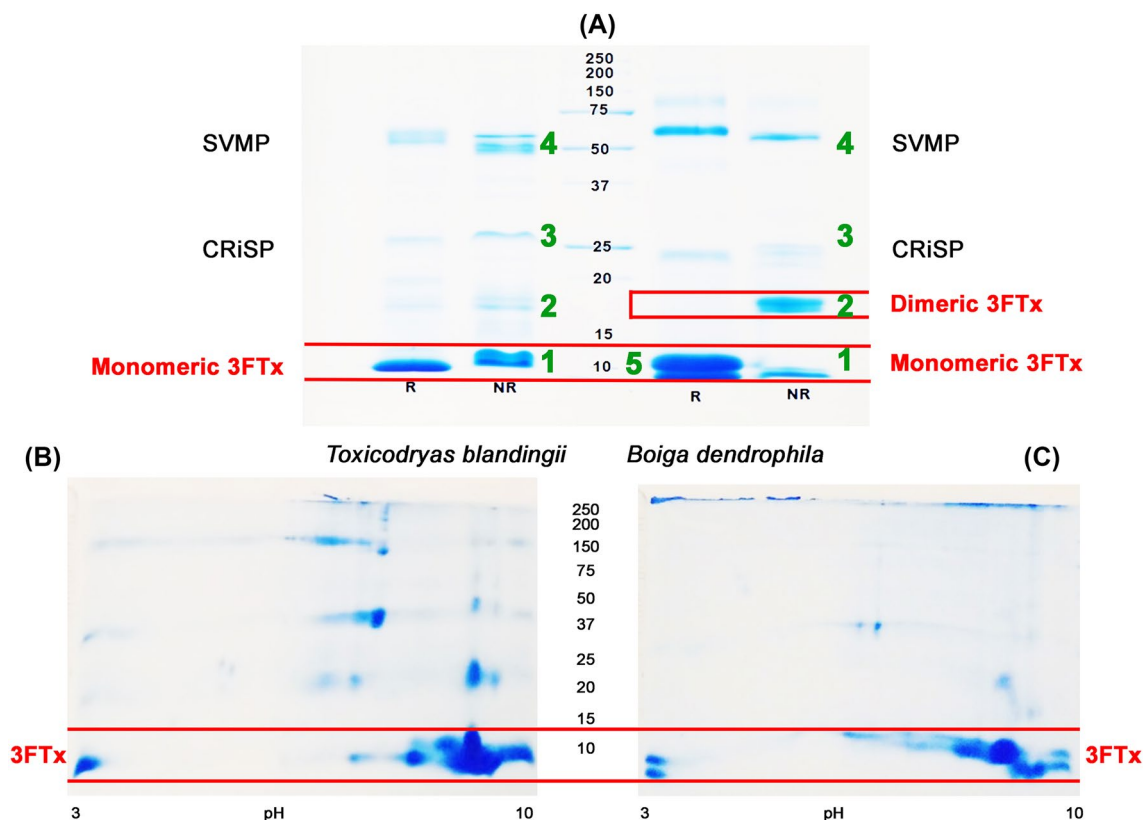
## Discussion

NFFS venoms remain a neglected area of research. This is in part due to the comparative difficulty of obtaining venom samples (Hill and Mackessy 1997), the fact that bites from most of these snakes have negligible impacts on human health (Minton 1990), and because it has only been relatively recently that the research community has begun to regard many of these snakes as venomous at all (Fry et al. 2003a, 2015; Fry and Wuster 2004; Nelsen et al. 2014). This study focuses on the evolutionary dynamics of one toxin family and one genus of colubrids; there remains much work to be done on other toxin types and other NFFS.

Proteomic and transcriptomic studies of *Boiga* venoms revealed that diapsid-specific neurotoxic 3FTx are the predominant toxins secreted by this amphibian, bird, and reptile feeding genus. The previous investigation into this genus revealed a unique disulphide-bond-linked 3FTx dimer called irditoxin isolated from *B. irregularis*, an invasive species responsible for devastating decreases in bird populations on the island of Guam (Savidge 1986, 1988; Pawlak et al. 2008). This was especially significant, because it was the first known covalently linked 3FTx–3FTx dimer toxin. Irditoxin was shown to have a particularly strong

diapsid-specific (bird and reptile) neurotoxicity, tenfold that of monomeric toxins (Pawlak et al. 2008). Despite this toxin being isolated 8 years ago, the presence or absence of homologous toxins in other *Boiga* species has not been thoroughly investigated until now.

Our transcriptomic analyses revealed for the first time that species besides *B. irregularis* produce transcripts with the necessary additional cysteines to form both the A and B chains of irditoxin homologues. Such pairs of transcripts were recovered from *B. dendrophila dendrophila*, *B. d. gemmicincta*, and *B. nigriceps* (Table 3). We also found, from the alignment of our data and previously published toxin sequences, that two papers had published toxins with cysteines at the characteristic irditoxin sites prior to this paper. Fry et al. (2008) published a toxin sequence from *Trimorphodon biscutatus* (*Trimorphodon\_biscutatus\_A7X3S2*) with the A cysteine and one with the B cysteine from *Telescopus dhara* (*Telescopus\_dhara\_A7X3V0*) the year prior to the discovery of irditoxin (Pawlak et al. 2009). Thus, the earlier study did not recognize the significance of these cysteines. Modahl and Mackessy (2016) discussed the similarity of their two *B. cynodon* sequences to irditoxin A (*Boiga\_cynodon\_A0A193CHL1*, 83% identical) and B (*Boiga\_cynodon\_A0A193CHL2*, 93% identical); however,



**Fig. 5** **a** 1D SDS-PAGE glycine produced under reducing and non-reducing conditions stained with colloidal coomassie brilliant blue G-250. Toxins of interest were identified by LC/MS–MS of excised bands (green numbers correspond to Supplementary Tables 1 and 2 of LC–MS/MS results). From left; *Toxicodryas blandingii* (Tb) reduced (R), non-reduced (NR), ladder (250–10 kDa), *Boiga dendrophila* (Bdd) reduced (R), non-reduced (NR). **b** 2D SDS-PAGE reduced electrophoresis mini gel. *Toxicodryas blandingii* crude venom stained with colloidal coomassie brilliant blue G-250. First dimension: isoelectric focusing (pH 3–10 non-linear gradient); sec-

ond dimension molecular weight 250–10 kDa: 12% SDS-PAGE. The pH gradient and the molecular weight marker positions are shown. 3FTx were identified by weight and comparison to 1D results. **c** 2D SDS-PAGE reduced electrophoresis mini gel. *Boiga dendrophila* crude venom stained with colloidal coomassie brilliant blue G-250. First dimension: isoelectric focusing (pH 3–10 non-linear gradient); second dimension molecular weight 250–10 kDa: 12% SDS-PAGE. The pH gradient and the molecular weight marker positions are shown. 3FTx were identified by weight and comparison to 1D results. (Color figure online)

only the sequence similar to the B subunit actually contained the characteristic cysteine and the presence of this cysteine was not commented upon. Modahl and Mackessy (2016) also published—but did not discuss—a *B. nigriceps* sequence that contained the A cysteine and a *Trimorphodon lambda* sequence that contained the B cysteine; nor did they discuss the presence or absence of the cysteines necessary for forming covalent bonds between the subunits when they speculated on the possible occurrence of dimeric complexes homologous to irditoxin in species besides *B. irregularis*.

Conspicuously, our phylogenetic, clustering, and network analyses (Figs. 2, 3) demonstrate that sequence similarity does not reliably distinguish between toxins with or without the A and B cysteines. Our transcriptomic analyses confirm the finding of toxins with the A cysteine from *B. nigriceps*, but they did not recover any sequence with the B cysteine from *B. cynodon*. Because their technique of sequencing

mRNA from venom rather than gland tissue resulted in very low yields, Modahl and Mackessy (2016) did not recover toxins with both the A and B cysteines from any species including *B. irregularis*. Thus, our analyses are the first evidence that species besides *B. irregularis* transcribe both the A and B subunits.

Despite the fact that our data include *Trimorphodon* toxins with cysteines at the characteristic site of both the A and B chains of irditoxin, they are found in different species with the A chain in *T. biscutatus* and the B chain in *T. lambda*. The *T. biscutatus* toxin with the A cysteine is particularly unusual, because it has additionally mutated the third and fourth canonical cysteines to phenylalanine and threonine, respectively. Whether either species actually secretes both chains and whether they form a dimeric toxin will require a thorough proteomic investigation of their venoms. The function of the putative toxin is also of interest; while it seems

somewhat unlikely these terrestrial snakes would possess a bird-specific toxin, these toxins are also potent against other diapsids such as lizards. However, *B. trigonata*, which is morphologically and ecologically convergent with *Telescopus* and *Trimorphodon*, did not yield contigs with either of the characteristic iridotoxin cysteines. Specific biochemical investigation of *B. trigonata* venom would be necessary to conclusively confirm or refute this finding.

The distribution of these toxins—with the *Trimorphodon* A subunit more closely related to the *Trimorphodon* B subunit than the *Boiga* A subunit and the *Telescopus* B subunit nested within the *Boiga* B subunits (Fig. 2)—suggests that the B subunit may have arisen in a common ancestor before much of the colubrine radiation and that toxins similar to the A subunit has evolved independently in *Boiga* and *Trimorphodon*. However, given how closely related the A and B subunits are in our phylogeny, it must also be considered that this denmotoxin-like clade of 3FTx may have been exapted for forming dimeric complexes and that both subunits could have evolved independently on multiple occasions. While the latter scenario is less parsimonious, to settle the issue, a greater diversity of colubrine sequences must be isolated and characterized to elucidate the likely ancestral forms of these proteins, as has been done with dimeric phospholipase A<sub>2</sub> toxins from crotaline venoms (Whittington et al. 2018).

Somewhat surprisingly, we did not find evidence for widespread positive selection across the denmotoxin-like 3FTx. In snake venom proteins, extensive duplication is typically accompanied with high  $\omega$  values (e.g., Margres et al. 2013; Sunagar et al. 2013; Dashevsky and Fry 2018). The only extreme positive selection we measured was in Clade B ( $\omega = 3.02$ ), and as discussed earlier (See “[Signals of Selection](#)”), this value may be an artefact from low sample size and wide taxonomic range that would decrease with further sampling. On the other hand, we found Clade D, which includes the canonical iridotoxin B chain, to be under fairly strong negative selection ( $\omega = 0.67$ ). Perhaps, this indicates that the structural constraints of dimerization may strengthen the negative selection acting on a toxin, as has been shown for non-covalently linked dimeric 3FTx (Sunagar et al. 2013). The overall low  $\omega$  values may help explain how these denmotoxin-like 3FTx have remained relatively similar across much of the colubrine radiation and the extensive diversification within *Boiga*.

We also expected to see a pattern among site-specific signals of selection similar to that found in the 8-cysteine 3FTx, where the loops of the toxins—the domains that interact with the toxins’ targets—tended to be enriched in sites experiencing positive selection and the cores which are essential to the overall structure of the protein were enriched for those experiencing negative selection (Sunagar et al. 2013; Dashevsky and Fry 2018). However, no such pattern is apparent in the clades of denmotoxin-like 3FTx (Fig. 4).

This may be partly due to the overall strength of positive selection being lower. We also see little variation in the predicted protein structures of the various clades. This is certainly due in part to the fact that only three structures have ever been published of 10-cysteine 3FTx: iridotoxin B chain (Clades A, B, C, and D), denmotoxin (Clades E and F), and iridotoxin A chain (Clade G). However, the relatively high similarity between the clades and low  $\omega$  values also plays a role in this. For these reasons, we suspect that the main source of functional diversity within the denmotoxin-like 3FTx is the actual dimerization, i.e., all the monomeric toxins are likely to exhibit similar activity to other monomeric toxins such as denmotoxin, while all the *Boiga* dimeric toxins are likely to be very similar to iridotoxin. As discussed earlier, the function of the putative dimeric *Trimorphodon* toxin remains speculative, but diapsid-specific toxicity seems likely. Given the apparent conservation of function within the denmotoxin-like 3FTx, it seems likely that the extensive duplication serves the proximal evolutionary purpose of increasing venom yield through gene dosage effects, a pattern of venom evolution that has been demonstrated in rattlesnakes (Margres et al. 2017).

*Toxicodryas blandingii* and *B. d. dendrophila* were used for our venom proteomics, because they bracket much of the genetic, morphological, geographical, and ecological diversity found within this clade of colubrine snakes: *T. blandingii* is restricted to Africa, where this radiation is thought to have started, while *B. d. dendrophila* is representative of the recently derived Southeast Asian lineage within *Boiga* (Pyron et al. 2013a; Figueroa et al. 2016). Proteomic investigations via 1D and 2D SDS-PAGE confirmed that both venoms were dominated by 3FTx (Fig. 5). It was hypothesised that *B. d. dendrophila* would produce 3FTx homologous to those of its close relative, *B. irregularis*, since they occupy similar ecological niches and we had assembled transcripts of such homologues. A series of analyses including 1D (reduced and non-reduced) SDS-PAGE, 2D (reduced) SDS-PAGE, LC-MS/MS peptide sequencing, and transcriptomics supported this hypothesis (Fig. 5, Supplementary Tables 1 and 2). The protein band between 15 and 20 kDa that can be seen in the *B. d. dendrophila*—but not *T. blandingii*—venom under non-reducing conditions matches what we would expect from iridotoxin: the combined mass of its A and B chains is approximately 17 kDa (Fig. 5a, b; Pawlak et al. 2008). It is apparent that this band contains dimeric complexes, because under reducing conditions (which break disulphide bonds), this heavier band disappears and a new monomeric 3FTx band appears at a lower molecular weight (Fig. 5a). The double band pattern is also seen in the 2D reduced gel for *B. d. dendrophila*, but not *T. blandingii* (Fig. 5b, c). This suggests that *B. d. dendrophila* produces toxins of similar molecular mass and biochemical structure

to iridotoxin, while *T. blandingii* does not. These toxins were confirmed to be 3FTx by LC–MS/MS sequencing of proteins isolated from gel bands (Supplementary Tables 1 and 2). This indicates that the evolution of dimeric toxins with both subunits homologous to those of iridotoxin occurred after the genus *Boiga* separated from its sister genera and may help explain the patterns of migration and speciation we see in the genus.

Our results suggest that the idea that the evolution of dimeric toxins was the key evolutionary innovation exclusive to *B. irregularis* which facilitated it being a unique and highly successful invasive species is likely too simplistic, because similar toxins can be found in several other species. Perhaps, all *Boiga* species, or at least those expressing iridotoxin-like dimers, pose similar risks and it is merely chance that it was *B. irregularis* invaded Guam rather than another species. If this is the case, then iridotoxin may have underpinned the rapid evolution and migration of the modern *Boiga* species. The havoc *B. irregularis* wreaked on Guamanian bird populations demonstrates how effective these toxins and the snakes employing them might have been as they encountered naïve prey populations during their invasion of Southeast Asia and Oceania. A positive feedback between their arboreal lifestyle and this potent, diapsid-specific, toxin could explain why there are so many species of *Boiga* occupying highly arboreal niches throughout the Asian region. However, this would raise the question of why none of the other species have invaded other locations: perhaps, Guam's ecology was unique in the opportunities and *Boiga* was given to invade in the first place, the lack of native snake species, or some other factor. While it is possible that there is something else distinctive about *B. irregularis* that makes it more noxious than its close relatives, no aspect of their ecology is an obvious culprit. For instance, *B. irregularis* is known to be highly food-motivated and somewhat generalist in feeding habits (Chiszar et al. 1993) which are traits often found in invasive species, but many other *Boiga* species have similarly broad diets (Stuebing et al. 1999). Thus, the invasion of Guam and surrounding islands by *B. irregularis* may simply have been a chance event based on which species was first transported to these islands and that other large species such as *B. cynodon* or *B. dendrophila* may have been just as invasive had either been given the opportunity. Iridotoxin itself could still be the answer if the *B. irregularis* toxins happen to be more potent than their homologs in closely related species. However, neurotoxicity testing has shown that *B. cynodon*, for example, is of comparable toxicity (Lumsden et al. 2004). Whatever the reason, the fact remains that *B. irregularis* is a proven threat as an invasive species, while other *Boiga* species are not. The currently available evidence neither disproves nor supports the idea that this is due to differences in their biology rather than a contingent historical fact.

**Acknowledgements** DD was funded by a UQ Centennial Scholarship from The University of Queensland DD and JD were funded by Research Training Program scholarships from the Australian Government Department of Education and Training. DRR was funded by the National Science Foundation (NSF DEB 1638902). BGF was funded by an Australian Research Council Future Fellowship and by the University of Queensland.

## References

- Ali SA, Yang DC, Jackson TN, Undheim EA, Koludarov I, Wood K, Jones A, Hodgson WC, McCarthy S, Ruder T (2013) Venom proteomic characterization and relative antivenom neutralization of two medically important Pakistani elapid snakes (*Bungarus sindanus* and *Naja naja*). *J Proteomics* 89:15–23
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Baig KJ, Masroor R, Arshad M (2011) Biodiversity and ecology of the herpetofauna of Cholistan Desert, Pakistan. *Russ J Herpetol* 15:193–205
- Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2013) GenBank. *Nucleic Acids Res* 41:D36–D42
- Boquet P (1979) History of snake venom research. Springer, New York, pp. 3–14
- Brahma RK, McCleary RJR, Kini RM, Doley R (2015) Venom gland transcriptomics for identifying, cataloging, and characterizing venom proteins in snakes. *Toxicon* 93:1–10
- Broaders M, Ryan MF (1997) Enzymatic properties of the Duvernoy's secretion of Blanding's tree snake (*Boiga blandingi*) and of the mangrove snake (*Boiga dendrophila*). *Toxicon* 35:1143–1148
- Burbrink FT, Lawson R (2007) How and when did Old World ratsnakes disperse into the New World? *Mol Phylogenet Evol* 43:173–189
- Cascardi J, Young BA, Husic HD, Sherma J (1999) Protein variation in the venom spat by the red spitting cobra, *Naja pallida* (Reptilia: Serpentes). *Toxicon* 37:1271–1279
- Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG (2013) Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol Evol* 28:219–229
- Chiszar D, Dunn TM, Smith HM (1993) Response of brown tree snakes (*Boiga irregularis*) to human blood. *J Chem Ecol* 19:91–96
- Cipriani V, Debono J, Goldenberg J, Jackson TNW, Arbuckle K, Dobson J, Koludarov I, Li B, Hay C, Dunstan N et al (2017) Correlation between ontogenetic dietary shifts and venom variation in Australian brown snakes (*Pseudonaja*). *Comp Biochem Physiol Part C* 197:53–60
- Daltry JC, Wüster W, Thorpe RS (1996) Diet and snake venom evolution. *Nature* 379:537
- Das I (2012) A naturalist's guide to the snakes of Southeast Asia. John Beaufoy Publishing, Oxford
- Dashevsky D, Fry BG (2018) Ancient diversification of three-finger toxins in *Micrurus* coral snakes. *J Mol Evol* 86:58–67
- Debono J, Cochran C, Kuruppu S, Nouwens A, Rajapakse NW, Kawasaki M, Wood K, Dobson J, Baumann K, Jouiaei M (2016) Canopy venom: proteomic comparison among new world arboreal pit-viper venoms. *Toxins* 8:210
- Debono J, Dobson J, Casewell NR, Romilio A, Li B, Kurniawan N, Mardon K, Weisbecker V, Nouwens A, Kwok HF, Fry BG (2017) Coagulating colubrids: evolutionary, pathophysiological and biodiversity implications of venom variations between boomslang (*Dispholidus typus*) and twig snake (*Thelotornis mosambicanus*). *Toxins* 9:171

- Deufel A, Cundall D (2006) Functional plasticity of the venom delivery system in snakes with a focus on the poststrike prey release behavior. *Zool Anzeiger A* 245:249–267
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Figueroa A, McKelvy AD, Grismer LL, Bell CD, Lailvaux SP (2016) A species-level phylogeny of extant snakes with description of a new colubrid subfamily and genus. *PLoS ONE* 11:e0161070
- Fritts TH, McCoid MJ, Haddock RL (1994) Symptoms and circumstances associated with bites by the brown tree snake (Colubridae: *Boiga irregularis*) on Guam. *J Herpetol* 28:27–33
- Fry BG (1999) Structure–function properties of venom components from Australian elapids. *Toxicon* 37:11–32
- Fry BG (2005) From genome to “venome”: molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. *Genome Res* 15:403–420
- Fry BG, Wüster W (2004) Assembling an Arsenal: origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences. *Mol Biol Evol* 21:870–883
- Fry BG, Lumsden NG, Wüster W, Wickramaratna JC, Hodgson WC, Kini RM (2003a) Isolation of a Neurotoxin ( $\alpha$ -colubritoxin) from a nonvenomous colubrid: evidence for early origin of venom in snakes. *J Mol Evol* 57:446–452
- Fry BG, Wüster W, Ramjan R, Jackson SF, Martelli T, P., and Kini RM (2003b) Analysis of Colubroidea snake venoms by liquid chromatography with mass spectrometry: evolutionary and toxicological implications: LC/MS analysis of Colubroidea snake venoms. *Rapid Commun Mass Spectrom* 17:2047–2062
- Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SFR, Kuruppu S, Fung K, Hedges SB, Richardson MK et al (2006) Early evolution of the venom system in lizards and snakes. *Nature* 439:584–588
- Fry BG, Scheib H, Weerd L van der, Young B, McNaughtan J, Ramjan SFR, Vidal N, Poelmann RE, Norman JA (2008) Evolution of an arsenal structural and functional diversification of the venom system in the advanced Snakes (Caenophidia). *Mol Cell Proteomics* 7:215–246
- Fry BG, Roelants K, Champagne DE, Scheib H, Tyndall JDA, King GF, Nevalainen TJ, Norman JA, Lewis RJ, Norton RS et al (2009a) The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu Rev Genomics Hum Genet* 10:483–511
- Fry BG, Vidal N, van der Weerd L, Kochva E, Renjifo C (2009b) Evolution and diversification of the Toxicofera reptile venom system. *J Proteomics* 72:127–136
- Fry BG, Undheim EAB, Ali SA, Jackson TNW, Debono J, Scheib H, Ruder T, Morgenstern D, Cadwallader L, Whitehead D et al (2013) Squeezers and leaf-cutters: differential diversification and degeneration of the venom system in toxicoferan reptiles. *Mol Cell Proteomics* 12:1881–1899
- Fry BG, Sunagar K, Casewell NR, Kochva E, Roelants K, Scheib H, Wüster W, Vidal N, Young B, Burbrink F et al (2015a). The origin and evolution of the Toxicofera reptile venom system. In: Fry BG (ed) *Venomous reptiles and their toxins: evolution, pathophysiology and biodiscovery*. Oxford University Press, Oxford, pp 1–31
- Fry BG, Sunagar K, Casewell NR, Kochva E, Roelants K, Scheib H, Wüster W, Vidal N, Young B, Burbrink F, Pyron RA, Vonk FJ, Jackson TNW (2015b) The origin and evolution of the Toxicofera reptile venom system. In: Fry BG (ed) *Venomous reptiles and their toxins: evolution, pathophysiology and biodiscovery*. Oxford University Press, New York, pp 1–31
- Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28:3150–3152
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29:644
- Hayes WK, Herbert SS, Harrison JR, Wiley KL (2008) Spitting versus biting: differential venom gland contraction regulates venom expenditure in the black-necked spitting cobra, *Naja nigricollis nigricollis*. *J Herpetol* 42:453–460
- Heyborne WH, Mackessy SP (2013) Identification and characterization of a taxon-specific three-finger toxin from the venom of the Green Vinesnake (*Oxybelis fulgidus*; family Colubridae). *Biochimie* 95:1923–1932
- Hill RE, Mackessy SP (1997) Venom yields from several species of colubrid snakes and differential effects of ketamine. *Toxicon* 35:671–678
- Jackson K (2003) The evolution of venom-delivery systems in snakes. *Zool J Linn Soc* 137:337–354
- Jackson TN, Fry BG (2016) A tricky trait: applying the fruits of the “function debate” in the philosophy of biology to the “venom debate” in the science of toxinology. *Toxins* 8(9):263
- Jackson TNW, Sunagar K, Undheim EAB, Koludarov I, Chan AHC, Sanders K, Ali SA, Hendrikx I, Dunstan N, Fry BG (2013) Venom down under: dynamic evolution of Australian elapid snake toxins. *Toxins* 5:2621–2655
- Jackson TNW, Young B, Underwood G, McCarthy CJ, Kochva E, Vidal N, van der Weerd L, Nabuurs R, Dobson J, Whitehead D et al (2016) Endless forms most beautiful: the evolution of ophidian oral glands, including the venom system, and the use of appropriate terminology for homologous structures. *Zoology* 136:107–130
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Pyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 10:845–858
- Kessler P, Marchot P, Silva M, Servent D (2017) The three-finger toxin fold: a multifunctional structural scaffold able to modulate cholinergic functions. *J Neurochem* 142:7–18
- Kochva E, Viljoen CC, Botes DP (1982) A new type of toxin in the venom of snakes of the genus *Atractaspis* (Atractaspidinae). *Toxicon* 20:581–592
- Krueger F (2015) Trim Galore!: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files
- Larsson A (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30:3276–3278
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760
- Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658–1659
- Li M, Fry BG, Kini RM (2005) Eggs-only diet: its implications for the toxin profile changes and ecology of the marbled sea snake (*Aipysurus eydouxii*). *J Mol Evol* 60:81–89
- Lumsden NG, Fry BG, Kini RM, Hodgson WC (2004) In vitro neuromuscular activity of ‘colubrid’ venoms: clinical and evolutionary implications. *Toxicon* 43:819–827
- Lumsden NG, Fry BG, Ventura S, Kini RM, Hodgson WC (2005) Pharmacological characterisation of a neurotoxin from the venom of *Boiga dendrophila* (Mangrove catsnake). *Toxicon* 45:329–334
- Mackessy SP (2010) Evolutionary trends in venom composition in the western rattlesnakes (*Crotalus viridis sensu lato*): toxicity vs. tenderizers. *Toxicon* 55:1463–1474
- Mackessy SP, Sixberry NM, Heyborne WH, Fritts T (2006) Venom of the Brown Treesnake, *Boiga irregularis*: Ontogenetic shifts and taxa-specific toxicity. *Toxicon* 47:537–548

- Mapleson D, Garcia Accinelli G, Kettleborough G, Wright J, Clavijo BJ (2017) KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics* 33:574–576
- Marçais G, Kingsford C (2011) A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* 27:764–770
- 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 Genom* 14:1
- Margres MJ, Bigelow AT, Lemmon EM, Lemmon AR, Rokyta DR (2017) Selection to increase expression, not sequence diversity, precedes gene family origin and expansion in rattlesnake venom. *Genetics* 206:1569–1580
- McGivern JJ, Wray KP, Margres MJ, Couch ME, Mackessy SP, Rokyta DR (2014) RNA-seq and high-definition mass spectrometry reveal the complex and divergent venoms of two rear-fanged colubrid snakes. *BMC Genomics* 15:1061
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M et al (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303
- Minton SA (1990) Venomous bites by nonvenomous snakes: an annotated bibliography of colubrid envenomation. *J Wilderness Med* 1:119–127
- Mitchell SW (1861) Researches upon the venom of the rattlesnake: with an investigation of the anatomy and physiology of the organs concerned. Smithsonian Institution, Washington, DC
- Modahl CM, Mackessy SP (2016) Full-length venom protein cDNA sequences from venom-derived mRNA: exploring compositional variation and adaptive multigene evolution. *PLOS Negl Trop Dis* 10:e0004587
- Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Pond SLK (2012) Detecting individual sites subject to episodic diversifying selection. *PLOS Genet* 8:e1002764
- Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Pond K, L. S., and Scheffler K (2013) FUBAR: a fast, unconstrained Bayesian AppRoximation for inferring selection. *Mol Biol Evol* 30:1196–1205
- Nelsen DR, Nisani Z, Cooper AM, Fox GA, Gren ECK, Corbit AG, Hayes WK (2014) Poisons, toxungens, and venoms: redefining and classifying toxic biological secretions and the organisms that employ them. *Biol Rev* 89:450–465
- Oulion B, Dobson JS, Zdenek CN, Arbuckle K, Lister C, Coimbra FCP, op den Brouw B, Debono J, Rogalski A, Violette A et al (2018) Factor X activating *Atractaspis* snake venoms and the relative coagulotoxicity neutralising efficacy of African antivenoms. *Toxicol Lett* 288:119–128
- Panagides N, Jackson TNW, Ikonomopoulou MP, Arbuckle K, Pretzler R, Yang DC, Ali SA, Koludarov I, Dobson J, Sanker B et al (2017) How the cobra got its flesh-eating venom: cytotoxicity as a defensive innovation and its co-evolution with hooding, aposematic marking, and spitting. *Toxins* 9:103
- Pawlak J, Mackessy SP, Fry BG, Bhatia M, Mourier G, Fruchart-Gaillard C, Servent D, Ménez R, Stura E, Ménez A et al (2006) Denmotoxin, a three-finger toxin from the colubrid snake *Boiga dendrophila* (Mangrove Catsnake) with bird-specific activity. *J Biol Chem* 281:29030–29041
- Pawlak J, Mackessy SP, Sixberry NM, Stura EA, Le Du MH, Ménez R, Foo CS, Ménez A, Nirthanan S, Kini RM (2008) Irditoxin, a novel covalently linked heterodimeric three-finger toxin with high taxon-specific neurotoxicity. *FASEB J* 23:534–545
- Pei J, Grishin NV (2001) AL2CO: calculation of positional conservation in a protein sequence alignment. *Bioinformatics* 17:700–712
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612
- Pla D, Sanz L, Whiteley G, Wagstaff SC, Harrison RA, Casewell NR, Calvete JJ (2017) What killed Karl Patterson Schmidt? Combined venom gland transcriptomic, venomomic and antivenomic analysis of the South African green tree snake (the boomslang), *Dispholidus typus*. *Biochim Biophys Acta* 1861:814–823
- Pla D, Petras D, Saviola AJ, Modahl CM, Sanz L, Pérez A, Juárez E, Frieze S, Dorresteijn PC, Mackessy SP et al (2018a) Transcriptomics-guided bottom-up and top-down venomomics of neonate and adult specimens of the arboreal rear-fanged Brown Treesnake, *Boiga irregularis*, from Guam. *J Proteomics* 174:71–84
- Pla D, Petras D, Saviola AJ, Modahl CM, Sanz L, Pérez A, Juárez E, Frieze S, Dorresteijn PC, Mackessy SP, Calvete (2018b) Transcriptomics-guided bottom-up and top-down venomomics of neonate and adult specimens of the arboreal rear-fanged Brown Treesnake, *Boiga irregularis*, from Guam. *J. Proteomics* 174:71–84
- Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679
- Pyron RA, Burbrink FT, Wiens JJ (2013a) A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evol Biol* 13:93
- Pyron RA, Kandambi HKD, Hendry CR, Pushpamal V, Burbrink FT, Somaweera R (2013b) Genus-level phylogeny of snakes reveals the origins of species richness in Sri Lanka. *Mol Phylogenet Evol* 66:969–978
- Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842
- Ramadhan G, Iskandar DT, Subasri DR (2010) A new species of Cat Snake (Serpentes: Colubridae) morphologically similar to *Boiga cynodon* from the Nusa Tenggara Islands, Indonesia. *Asian Herpetol Res* 1:22–30
- Rochelle MJ, Kardong KV (1993) Constriction versus envenomation in prey capture by the brown tree snake, *Boiga irregularis* (Squamata: Colubridae). *Herpetologica* 1993:301–304
- Rodda GH, Fritts TH (1992) The impact of the introduction of the colubrid snake *Boiga irregularis* on Guam's lizards. *J Herpetol* 26:166–174
- Rodda GH, Savidge JA (2007) Biology and impacts of Pacific island invasive species. 2. *Boiga irregularis*, the brown tree snake (Reptilia: Colubridae). *Pac Sci* 61:307–324
- Rokyta DR, Lemmon AR, Margres MJ, Aronow K (2012) The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genomics* 13:312
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542
- Savidge JA (1986) The Role of Disease and Predation in The Decline of Guam's Avifauna (Extinction, *Boiga irregularis*). PhD Thesis. University of Illinois at Urbana-Champaign, Champaign
- Savidge JA (1988) Food habits of *Boiga irregularis*, an introduced predator on Guam. *J Herpetol* 22:275–282
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504
- Shivik JA (2006) Are vultures birds, and do snakes have venom, because of macro- and microscavenger conflict? *BioScience* 56:819–823
- Stuebing RB, Inger RF, Tan FL (1999) Field guide to the snakes of Borneo. Natural History Publications (Borneo), Kota Kinabalu

- Sunagar K, Jackson T, Undheim E, Ali S, Antunes A, Fry B (2013) Three-fingered RAVERS: rapid accumulation of variations in exposed residues of snake venom toxins. *Toxins* 5:2172–2208
- Sunagar K, Undheim EAB, Scheib H, Gren ECK, Cochran C, Person CE, Koludarov I, Kelln W, Hayes WK, King GF et al (2014) Intraspecific venom variation in the medically significant Southern Pacific Rattlesnake (*Crotalus oreganus helleri*): Biodiscovery, clinical and evolutionary implications. *J Proteomics* 99:68–83
- Terrat Y, Sunagar K, Fry BG, Jackson TNW, Scheib H, Fourmy R, Verdenaud M, Blanchet G, Antunes A, Ducancel F (2013) *Atractaspis aterrima* toxins: the first insight into the molecular evolution of venom in side-stabbers. *Toxins* 5:1948–1964
- The Uniprot Consortium (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45:D158–D169
- Utkin YN, Kukhtina VV, Kryukova EV, Chiodini F, Bertrand D, Methfessel C, Tsetlin VI (2001) “Weak toxin” from *Naja kaouthia* is a nontoxic antagonist of  $\alpha 7$  and muscle-type nicotinic acetylcholine receptors. *J Biol Chem* 276:15810–15815
- Utkin Y, Sunagar K, Jackson T, Reeks T, Fry B (2015). Three-finger toxins (3FTxs). In: Fry BG (ed) *Venomous reptiles and their toxins: evolution, pathophysiology and biodiscovery*. Oxford University Press, Oxford, pp. 215–227
- Vonk FJ, Casewell NR, Henkel CV, Heimberg AM, Jansen HJ, McCleary RJ, Kerckamp HM, Vos RA, Guerreiro I, Calvete JJ (2013) The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc Natl Acad Sci USA* 110:20651–20656
- Whittington AC, Mason AJ, Rokyta DR (2018) A single mutation unlocks cascading exaptations in the origin of a potent pitviper neurotoxin. *Mol Biol Evol* 35(4):887–898
- Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, Huang W, He G, Gu S, Li S (2014) SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* 30:1660–1666
- Zhang J, Kobert K, Flouri T, Stamatakis A (2013) PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30:614–620