Convergent molecular evolution of toxins in the venom of advanced snakes (Colubroidea)

Xie, B.

Chapter 1. Introduction and outline of this thesis

Introduction

Early snakes likely possessed mixed serous-mucous oral glands inherited from the last common ancestor of Toxicoferan reptiles (Fry, et al. 2006). Extant snakes have evolved a number of different gland morphologies from this ancestral state (Fry, et al. 2008; Fry, et al. 2013). Some basal snake genera such as *Cylindrophis* and *Eryx* retain some serous gland tissue which likely produces appreciable quantities of protein (Phisalix and Caius 1918; Fry, et al. 2013). In contrast, in the derived basal snake lineages which have secondarily evolved powerful constriction as a novel form of prey capture, the glands have evolved towards primarily secreting mucous in order to lubricate the large furred or feathered prey, in order to facilitate their ingestion (Fry, et al. 2013). However, in the constricting snakes trace levels of proteins are still secreted as an evolutionary relic. These proteins are evolved from ancestral toxins and can be detected by SDS-PAGE gel electrophoresis or via PCR amplification of the encoding genes, and remain sufficiently similar to other snake toxins to produce false positives in antibody-based snake venom detection kits (Jelinek, et al. 2004; Fry, et al. 2013).

The explosive radiation of the advanced snakes (superfamily Colubroidea (Hsiang, et al. 2015) was associated with the partitioning of the mixed glands into two discrete glands, one devoted to venom production, the other for mucous (Fry, et al. 2008; Jackson, et al. 2017). The venom gland has subsequently evolved into an extraordinary diversity of morphological forms (Fry, et al. 2008; Fry, Sunagar, et al. 2015; Jackson, et al. 2017). The venom systems of the lamprophiid lineage (including the genera *Atractaspis* and *Homoroselaps*), elapids, and viperids are homoplastic in that they have convergently evolved hollow fangs linked via tube-like ducts to the venom glands which are enclosed by powerful compressor muscles to increase the speed and efficiency of venom delivery. Similar compressor muscles have also evolved in at least three other lineages (*Brachyophis revoili, Dispholidus typus,* and *Gonionotophis capensis*) without the additional refinement of syringe-like venom delivering dentition (Fry, et al. 2008; Fry, Sunagar, et al. 2015).

Numerous morphological and developmental studies have ascertained that the venom producing glands of all advanced snakes are homologous structures that develop from the primordium at the posterior end of the dental lamina (Kochva 1963a; Kochva 1963b; Kochva and Gans 1965; Fry, et al. 2008; Vonk, et al. 2008; Jackson, et al. 2017). Despite this demonstrated homology, the gland of rear-fanged snakes is often distinguished in the literature from that of front-fanged snakes through the use of the term 'Duvernoy's gland'. The use of this term perpetuates a historical mistake that was made in the original designation of the gland by Taub in 1967 (Taub 1967). At that time, Taub agreed with earlier work that the post-orbital gland of rear-fanged snakes produced venom, citing studies from the early 1900s (Alcock and Rogers 1902; Phisalix and Caius 1918), but considered the glands of viperids and elapids to be non-homologous to each other, and thus assumed the gland of the rear-fanged snakes were also not homologous to elapids or viperids. Crucially, the phrase 'Duvernoy's gland' was not even suggested to highlight this erroneous non-homology, but was simply suggested as a replacement for the name 'parotid gland' which was also occasionally assigned to this structure in rear-fanged snakes. It is now considered

well-established that the venom glands of all colubroid snakes are homologous (Fry, et al. 2008). In fact, elapid and viperid glands are more closely related to the glands of rear-fanged snakes than they are to each other. Thus, the use of the term 'venom gland' for the homologous glands of all advanced snakes is the more appropriate than to refer to a paraphyletic array of morphologies as 'Duvernoy's gland'.

Just as the venom glands of advanced snakes are homologous, so are the venom delivering teeth (fangs). Developmental uncoupling of the posterior sub-region of the tooth forming epithelium facilitated evolution of a wide range of highly modified posterior teeth, with tremendous diversity both in size and morphology (Fry, Sunagar, et al. 2015; Cleuren, et al. 2021). Enlarged rear-fangs—which have evolved on a myriad of occasions—and the three independent evolutions of hollow front fangs, all evolved from the same posterior teeth (Vonk, et al. 2008). These teeth evolved to be farther forward in the mouth of the three front-fanged lineages due to shortening of the maxillary bone and the loss of more anterior teeth (Vonk, et al. 2008).

In a similar fashion to the homology of the morphological aspects of the venom system, modern evidence has accumulated for the homology venom gland toxins expressed across the advanced snakes. The discovery of three-finger toxins (3FTx) for the first time outside of elapid snakes (Fry, Lumsden, et al. 2003a) stimulated a phylogenetic analysis of all known toxin types, revealing the non-monophyly of a myriad of toxin types relative to the organismal relationships (Fry and Wüster 2004). Several toxin families were found to be shared across all advanced snakes including 3FTx, acetylcholinesterase, C-type natriuretic peptides (CNP), kallikrein enzymes, hyaluronidase, kunitz, lectins, and snake venom metalloproteases (SVMP). As the species and their venom secretion and delivery system have diversified, so too have the venom proteins themselves. Accelerated evolution and rapid neofunctionalization are common traits of snake venom gene families (Fry, Wüster, et al. 2003; Sunagar, et al. 2013; Junqueira-de-Azevedo, et al. 2016; Dashevsky, et al. 2018; Dashevsky and Fry 2018; Barua and Mikheyev 2020; Dashevsky, et al. 2021). These genes are frequently duplicated, which can lead to functional and structural diversification (Moura-da-Silva, et al. 1996; Slowinski, et al. 1997; Afifiyan, et al. 1999; Kordiš and Gubenšek 2000), as well as faster rates of sequence evolution (Nakashima, et al. 1993; Kini and Chan 1999). This variety could be due to selection for the ability to kill and digest prey (Daltry, et al. 1996), or it could be the outcome of a predator–prey arms race (e.g., (Poran, et al. 1987; Heatwole and Poran 1995).

Despite the evolutionary novelty of snake venom proteins, a comprehensive reconstruction of the molecular evolutionary history of these major shared toxin types has not been undertaken. This has been in part due to the relatively low number of sequences available from rear-fanged species. Elapid and viperid snake species have been the focus of much more research because of their medical importance (Saviola, et al. 2014; Jackson, et al. 2019). To carry out these broad analyses, we obtained venom gland transcriptomes from eight rear-fanged species spanning the families Colubridae subfamilies Dipsadinae (*Helicops leopardinus, Heterodon nasicus*) and Natricinae (*Rhabdophis subminiatus*), Homalopsidae (*Homalopsis buccata*), and the Lamprophiidae subfamily Psammophiinae (*Malpolon monspessulanus, Psammophis schokari, Psammophis subtaeniatus,* and *Rhamphiophis oxyrhynchus*) as well as two viperid species spanning that family's phylogenetic range (*Pseudocerastes urarachnoides* and *Vipera*

transcaucasiana). With these sequences, we were able to reconstruct the molecular evolutionary history of the shared toxins and map their diversification patterns relative to the organismal relationships and functional changes in the venom. This allows us to evaluate the relative order of evolutionary events such as the diversification of Colubroidea, the partitioning of the venom glands into discrete mucous and protein-secreting units, diversifications in toxin families, and structural and functional novelties in toxin sequences.

Outline of this thesis

In **Chapter 1**, we introduce the background of this thesis. We describe how this thesis involves the *de novo* sequencing and of venom gland tissue from multiple snake species, including rarely-studied rear-fanged snakes. Because of the large amount of new data generated and analysed, we devote each of chapters 2-5 to a different family of toxins for which we found particularly important results.

In **Chapter 2**, we prepare the ground for future chapters by addressing some problematic issues in RNA-seq based transcriptomics, when used to study venom toxin evolution. This approach has been widely used in the study of the evolution, ecology, function, and pharmacology of animal venoms. However, the accuracy and completeness of venom profiles determined by transcriptomics are limited by the cross-contamination between samples and the performance of the transcriptome assembly. To solve these problems, we sequenced eight species of RFS and two FFS, and applied several commonly used *de novo* assembly methods to recover the authentic venom profiles followed by a strict criterion to discard chimeric transcripts. Evaluation of the pipelines and the software performance was carried out on the basis of the recovery of non-toxin and confidently-identified toxin transcripts. Serious misrepresentation of the diversity of the toxin families and their relative transcripts abundance are demonstrated here. The authentic toxin transcripts are then used in Chapter 3-5 to reveal the full-scale molecular evolutionary history and the patterns of selection.

In **Chapter 3**, we analyse the evolution of Kunitz-type toxins. These toxins, found in reptile venoms have exhibited extensive diversity of structures and functions, from enzyme inhibitors to channel-blocking neurotoxins. However, their detailed evolutionary trends and patterns remain a mystery. We therefore conducted a large-scale phylogenetic and selection analysis. This revealed that the kunitz-type toxins evolved by gene duplication and rapid diversification and showed that: (1) the main ancestral function plasmin inhibitors in Viperidae are under neutral selection while in non-vipers they are under purifying selection; (2) neurotoxic toxins are only found in non-viper clades and different neurotoxic types clustered in separate distinct clades under positive selection.

In **Chapter 4**, we examine C-type lectins, one of the largest toxin families in mammals and reptiles. These toxins are known to have various functions including defence against predator. Since the Viperidae split off from the remaining caenophidian snakes, a novel heterodimeric lectin type evolved by the loss of carbohydrate-binding ability. However, the evolutionary trends and patterns of C-type lectins remain a mystery. We therefore conducted a large-scale phylogenetic and selection analysis. We recovered multiple variants from non-viperid snakes that possessed the diagnostic cysteine of the dimeric lectin

form, but forms with and without the glutamine motif form were present. We also found that the α-subunit and β-subunit were subject to different selection pressures.

In **Chapter 5**, we study SVMP (snake venom metalloproteinase) toxins. These toxins serve as a model system for looking at the evolutionary mechanisms that lead to changes in protein function and structure. Extensive gene duplication and domain loss has occurred. And the ancestral P-III SVMPs have indicated a much more complex structure and functional diversity than P-I and P-II SVMPs. P-IIId SVMP and truncated SVMPs are recently emerged novel traits, but the evolution of P-IIId SVMP and the truncated SVMPs remain a mystery. The aim of this study was to investigate the evolutionary process that resulted in the structural and functional diversification within the P-IIId subfamily and the truncated SVMP propeptide type. We found structural convergences in including the evolution of cysteines for form heteromeric complexes within SVMP, and *de novo* evolution of new toxin families within the propeptide region occurring in the SVMP gene.

In **Chapter 6**, we summarize and discuss the results of Chapters 2-5.

In **Chapter 7**, we provide supplementary materials for this thesis.

References

Afifiyan F, Armugam A, Tan CH, Gopalakrishnakone P, Jeyaseelan K. 1999. Postsynaptic α-neurotoxin gene of the spitting cobra, *Naja naja sputatrix*: structure, organization, and phylogenetic analysis. Genome research, 9(3), 259-266.

Alcock AW & Rogers L. 1902. On the toxic properties of the saliva of certain 'non-poisonous' Colubrines. Proceedings of the Royal Society of London, 70(459-466), 446-454.

Barua A & Mikheyev AS. 2020. Toxin expression in snake venom evolves rapidly with constant shifts in evolutionary rates. Proceedings of the Royal Society B, 287(1926), 20200613.

Cleuren SG, Hocking DP & Evans AR. 2021. Fang evolution in venomous snakes: adaptation of 3D tooth shape to the biomechanical properties of their prey. Evolution. In Press

Daltry JC, Wüster W, Thorpe RS. 1996. Diet and snake venom evolution. Nature, 379(6565), 537-540.

Dashevsky D, Debono J, Rokyta D, Nouwens A, Josh P, Fry BG. 2018. Three-finger toxin diversification in the venoms of cat-eye snakes (Colubridae: *Boiga*). Journal of Molecular Evolution, 86(8), 531-545.

Dashevsky D, Fry BG. 2018. Ancient diversification of three-finger toxins in *Micrurus* coral snakes. Journal of Molecular Evolution, 86(1), 58-67.

Dashevsky D, Rokyta D, Frank N, Nouwens A, Fry BG. 2021. Electric Blue: Molecular Evolution of Three-Finger Toxins in the Long-Glanded Coral Snake Species *Calliophis bivirgatus*. Toxins, 13(2), 124.

Fry BG, Lumsden NG, Wüster W, Wickramaratna JC, Hodgson WC, Kini MR. 2003. Isolation of a neurotoxin (α-colubritoxin) from a nonvenomous colubrid: evidence for early origin of venom in snakes. Journal of Molecular Evolution, 57(4), 446-452.

Fry BG, Scheib H, van der Weerd L, Young B, McNaughtan J, Ramjan SR. 2008. Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). Molecular & Cellular Proteomics, 7(2), 215-246.

Fry BG, Sunagar K, Casewell NR, Kochva EL, Roelants K, Scheib H. 2015. The origin and evolution of the Toxicofera reptile venom system. In: *Venomous Reptiles and Their Toxins: Evolution, Pathophysiology and Biodiscovery*. New York: Oxford University Press. p.1-31

Fry BG, Undheim EA, Ali SA, Jackson TN, Debono J, Scheib H. 2013. Squeezers and leaf-cutters: differential diversification and degeneration of the venom system in toxicoferan reptiles. Molecular & Cellular Proteomics, 12(7), 1881-1899.

Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SR. 2006. Early evolution of the venom system in lizards and snakes. Nature, 439(7076), 584-588.

Fry BG, Wüster W. 2004. Assembling an arsenal: origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences. Molecular Biology and Evolution, 21(5), 870-883.

Fry BG, Wüster W, Ryan R, Sheik F, Jackson T, Martelli P. 2003. Analysis of Colubroidea snake venoms by liquid chromatography with mass spectrometry: evolutionary and toxinological implications. Rapid Communications in Mass Spectrometry, 17(18), 2047-2062.

Heatwole H, Poran NS. 1995. Resistances of sympatric and allopatric eels to sea snake venoms. Copeia, 136-147.

Hsiang AY, Field DJ, Webster TH, Behlke AD, Davis MB, Racicot RA. 2015. The origin of snakes: revealing the ecology, behavior, and evolutionary history of early snakes using genomics, phenomics, and the fossil record. BMC evolutionary biology, 15(1), 1-22.

Jackson TN, Jouanne H, Vidal N. 2019. Snake venom in context: neglected clades and concepts. Frontiers in Ecology and Evolution, 7, 332.

Jackson TN, Young B, Underwood G, McCarthy CJ, Kochva E, Vidal N. 2017. Endless forms most beautiful: the evolution of ophidian oral glands, including the venom system, and the use of appropriate terminology for homologous structures. Zoomorphology, 136(1), 107-130.

Jelinek GA, Tweed C, Lynch D, Celenza T, Bush B, Michalopoulos N. 2004. Cross reactivity between venomous, mildly venomous, and non-venomous snake venoms with the Commonwealth Serum Laboratories Venom Detection Kit. Emergency Medicine, 16(5-6), 459-464.

Junqueira-de-Azevedo IL, Campos PF, Ching AT, Mackessy SP. 2016. Colubrid venom composition: an-omics perspective. Toxins, 8(8), 230.

Kini RM, Chan YM. 1999. Accelerated evolution and molecular surface of venom phospholipase A 2 enzymes. Journal of Molecular Evolution, 48(2), 125-132.

Kochva, E. 1963. Development of the venom gland and trigeminal muscles in *Vipera palaestinae*. Cells Tissues Organs, 52(1-2), 49-89.

Kochva, E. 1963. The phylogenetic significance of the venom apparatus in snakes. American Zoologist, 3(4). 487.

Kochva E, Gans C. 1965. The venom gland of *Vipera palaestinae* with comments on the glands of some other viperines. Cells Tissues Organs, 62(3), 365-401.

Kordiš D, Gubenšek F. 2000. Adaptive evolution of animal toxin multigene families. Gene, 261(1), 43-52.

Moura-da-Silva AM, Theakston RD, Cramptonl JM. 1996. Evolution of disintegrin cysteine-rich and mammalian matrix-degrading metalloproteinases: gene duplication and divergence of a common ancestor rather than convergent evolution. Journal of Molecular Evolution, 43(3), 263-269.

Nakashima K, Ogawa T, Oda N, Hattori M, Sakaki Y, Kihara H. 1993. Accelerated evolution of *Trimeresurus flavoviridis* venom gland phospholipase A2 isozymes. Proceedings of the National Academy of Sciences 90(13), 5964-5968.

Phisalix M, Caius R. 1918. L'extension de la fonction venimeuse dans l'ordre entière des ophidiens et son existence chez des familles ou elle n'avait pas été soupçonnée jusqu'içi. Journal de Physiologie et de Pathologie générale, 17, 923-964.

Poran NS, Coss RG, Benjamini E. 1987. Resistance of California ground squirrels (*Spermophilus beecheyi*) to the venom of the northern Pacific rattlesnake (*Crotalus viridis oreganus*): a study of adaptive variation. Toxicon, 25(7), 767-777.

Saviola AJ, Peichoto ME, Mackessy SP. 2014. Rear-fanged snake venoms: an untapped source of novel compounds and potential drug leads. Toxin Reviews, 33(4), 185-201.

Slowinski JB, Knight A, Rooney AP. 1997. Inferring species trees from gene trees: a phylogenetic analysis of the Elapidae (Serpentes) based on the amino acid sequences of venom proteins. Molecular phylogenetics and evolution, 8(3), 349-362.

Sunagar K, Jackson TN, Undheim EA, Ali S, Antunes A, Fry BG. 2013. Three-fingered RAVERs: rapid accumulation of variations in exposed residues of snake venom toxins. Toxins, *5*(11), 2172-2208.

Taub AM. 1967. Comparative histological studies on Duvernoy's gland of colubrid snakes. Bulletin of the AMNH; v. 138, article 1.

Vonk FJ, Admiraal JF, Jackson K, Reshef R, de Bakker MA, Vanderschoot K. 2008. Evolutionary origin and development of snake fangs. Nature, 454(7204), 630-633.

Abbreviations

The following abbreviations for toxins are used in this thesis:

3ftx	Three-finger toxin
AChE	Acetylcholinesterase
BPP	Bradykinin-potentiating peptide
C3/CVF	Complement C3 of cobra venom factor
CNP	C-type natriuretic peptide
CTL	C-type lectin
CRISP	Cysteine-rich secretory protein
ESP	Epididymal secretory protein
HYAL	Hyaluronidase
KTT	Kunitz type toxin
LAAO	L-amino acid oxidase
NGF	Nerve growth factor
PDE	Phosphodiesterase
PLA2_II_E	Phospholipase_A2_type_II_E
PLB	Phospholipase_B
RAP	Renin aspartate protease
RNase	Ribonuclease
SVMP	Snake venom metalloproteinase
SVSP	Snake venom serine protease
VEGF	Vascular endothelial growth factor