Cover Page

# Universiteit Leiden

**Author:** Vonk, Freek Jacobus
**Title:** Snake evolution and prospecting of snake venom
**Date:** 2012-09-06

# Snake Evolution and Prospecting of Snake Venom

**Freek Jacobus Vonk**

**To my family and (snake) friends, with special thanks to my parents…..**

**Freek Jacobus Vonk**

Snake Evolution and Prospecting of Snake Venom

Dissertation Leiden University

Cover photo: a wild king cobra (*Ophiophagus hannah*), Indonesia (Java). By Freek Vonk

Printed by: Wöhrmann Print Services, Zutphen.

# Snake Evolution and Prospecting of Snake Venom

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van de Rector Magnificus Prof. Mr. P.F. van der Heijden,

volgens besluit van het College voor Promoties

te verdedigen op donderdag 6 september 2012

klokke 10:00 uur

door

**Freek Jacobus Vonk**

Geboren te Dordrecht,

Nederland 24 februari, 1983.

# Promotiecommissie

| | |
|---|---|
| **Promotor** | **Prof. Dr. Michael K. Richardson** |
| | |
| Overige leden | Prof. Dr. Carel J. ten Cate |
| | Prof. Dr. Herman P. Spaink |
| | Dr. Bart Vervust (University of Antwerp) |
| | Prof. Dr. Rob E. Poelman (Leiden University Medical Centre) |
| | Prof. Dr. Peter G.L. Klinkhamer |

# Contents

# Chapter 1: Introduction

Snakes are limbless predators that have evolved two different ways of incapacitating prey: suffocation and envenomation. The evolution of these life-history strategies lie in close relation with the evolution of body-elongation and limblessness. Taken together, these adaptations underlie the massive radiation of snakes which occurred after the K-T boundary and extinction of the dinosaurs, presumably because multiple niches became available and many of their predators got extinct. In this thesis, I examine the evolution of these adaptations. A general question to address would be: what underlies these adaptations in snakes and how do they correspond with their ecology? In addition to these evolutionary questions, I also provide a summary in chapter one on the current status of using snake venom toxins in drug discovery and design, because the evolution of venom is closely related to the evolution of a serpentine body form and fangs, but also to the question of "why" snake venom is so interesting for the pharmacological prospector. The high potency and specifity of snake venom toxins is generated by millions of years of evolution and strong selective pressures, so it is my believe that the different chapters in my thesis are tightly bound together. It is believed that it was Aristotle who first suggested – on the basis of symptoms observed in snakebite victims – that snakes may also be used to cure certain diseases (it was not yet known that venom was responsible for the observed effects). And in 1835 venom was already (although unsuccessfully) used to treat rabies. This field of "bioprospecting" has accelerated especially in the last decade, mainly because the development of many new sequencing and screening techniques. In chapter two, we go into the development and evolution of the different snake venom-conducting fangs. This has been a matter of great controversy among scientists for the last century and many different hypotheses have been proposed. This was particularly because the two groups of snakes with fangs in the front of their upper jaw (Elapidae and Viperidae), have only relatively recently in the 90s been shown not to form a monophyletic group, hinting towards an independent origin and evolution of front fangs. By using modern techniques such as *in situ* hybridization of snake embryos, and careful histological analysis we were able to reconstruct the evolutionary history and make a solid hypothesis for

1

the origin and evolution of fangs. In the third chapter of this thesis we look at the evolution of body-elongation and deregionalization in snakes, since that presumably has also been a major driver of venom evolution by providing the ability for snakes to coil up and act as coiled spring that can strike and envenomate prey at a distance. Since *Hox* determine the basic structure of an embryo we have performed many *in situ* hybridisations in snake embryos and carefully compared the expression profiles of different *Hox* genes. For the fourth chapter of this thesis we sequenced a draft genome of the largest venomous snake in the world, the king cobra (*Ophiophagus hannah*) and annotated many of its venom genes. We wanted to find out whether we could find evidence of gene duplication and / or gene modification of physiological genes for use in the venom gland. The birth-and-death model of gene evolution is the canonical framework for venom evolution, but was never actually thoroughly studies due to lack of genomic resources. Our genomic sequences let us study the king cobra venom genes in relation to its 'normal' physiological genes. In the fifth chapter, we have set out to develop a new efficient and rapid method of analysing snake venom mixtures and performing a biological assay at the same time. We used the acetylcholine binding protein (AChBP) as biological target, because acetylcholine receptors are also often involved in neurological disorders, such as Alzheimer's and Myasthenia gravis. In the last chapter of this thesis, I discuss this work and try to put it into broader perspective. The field of evolutionary genomics, molecular evolution as well as bioprospecting is huge and broad. However, I strongly believe that working on the edge of different scientific disciplines will provide you with a better understanding of nature and allows you to answer more questions.

# Chapter 2: Snake venom: from fieldwork to the clinic.

Vonk FJ, Jackson K, Doley R, Madaras F, Mirtschin PJ and N Vidal.
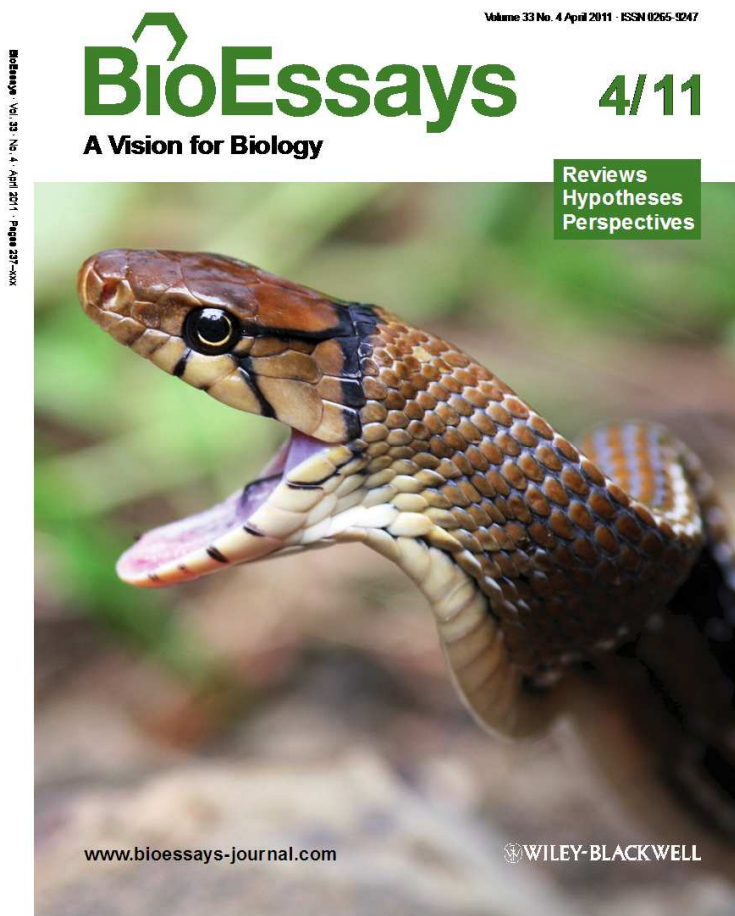
## Abstract

Snake venoms are recognized here as a grossly under-explored resource in pharmacological prospecting. Discoveries in snake systematics demonstrate that former taxonomic bias in research has led to the neglect of thousands of species of potential medical use. Recent discoveries reveal an unexpectedly vast degree of variation in venom composition among snakes, from different species down to litter mates. The molecular mechanisms underlying this diversity are only beginning to be understood. However, the enormous potential that this resource represents for pharmacological prospecting is clear. New high-throughput screening systems offer greatly increased speed and efficiency in identifying and extracting therapeutically useful molecules. At the same time a global biodiversity crisis is threatening the very snake populations on which hopes for new venom-derived medications depend. Biomedical researchers, pharmacologists, clinicians, herpetologists and conservation biologists must combine their efforts if the full potential of snake venom derived medications is to be realized.

## Introduction

Snakes are represented on earth today by some 3150 species [1] . Of these the vast majority (ca. 2700 species, see Fig. 1) represent a single massive diversification event which occurred after the K-T boundary and extinction of the dinosaurs. This large and relatively recent group is known as Caenophidia or "advanced snakes" and characterized by the possession of a venom-delivery system or components of such a system [2]. Snakes traditionally considered venomous are the 600 or so species with tubular front fangs, muscularized venom glands, and a bite significantly dangerous to humans (Viperidae, Elapidae and Atractaspidinae) - including well-known examples as the cobras, seasnakes, vipers and rattlesnakes. The remaining caenophidians were traditionally classified as "Colubridae", meaning snakes with a venom gland  whose venom poses no danger to humans, and who lack the fangs at the front of the mouth for injecting it. The "Colubridae" has been shown to be paraphyletic, and most of its subfamilies have recently been elevated to a familial rank in order to reflect their evolutionary distinctiveness (Fig. 1) [1,3].

"Colubrid" snakes have been largely neglected in venom research, because of the sole fact that bites have not been perceived as of medical importance, except few species as the African boomslang  (*Dispholidus typus*) and twigsnakes (*Thelotornis* spp.) and the Asian yamakagashi (*Rhabdophis tigrinus*). However, during the last few years, there has been a trend towards studying the venoms of these harmless venomous snakes to, primarily, increase our understanding of venom evolution. It became evident that these harmless snakes do secrete a strong acting venom with powerful toxins, although in a significantly lower amount and without an efficient injection mechanism [2,4]. This influenced the ongoing quest for venom molecules that may be useful in fighting human disease - a snake need not be dangerous to humans for its venom to still have a profound effect upon the human body. This field of biodiscovery is now rapidly emerging, especially due to the development of modern high-throughput screening assays that allow rapid identification of potential therapeutic agents.
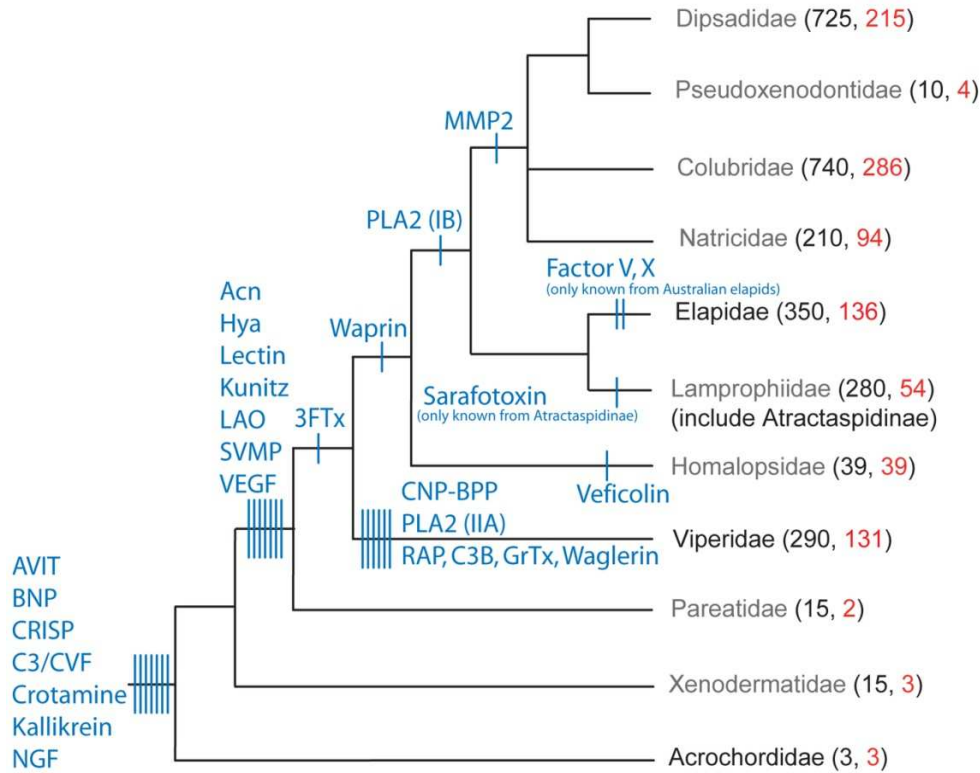
**Figure 1** | Phylogenetic tree showing the distribution of 28 snake venom protein families among advanced snakes, with the number of currently known species of each family and the number of red listed species. Families in grey used to make up the old traditional "Colubridae". Phylogeny based on [1,5].

Snake venom contains a mixture of powerful proteins and peptides that have evolved to be targeted to receptors, ion channels or enzymes [6], in addition to some carbohydrates, nucleosides, lipids, and metal ions, whose functions are not all known [7,8]. They interact with a wide variety of mammalian proteins and can disrupt the central and peripheral nervous systems, the blood coagulation cascade, the cardiovascular and neuromuscular systems, and homeostasis in general. These venom proteins act with great precision – different toxins recognize different subtypes of certain receptors with only subtle differences – and are very biologically active. The precision and power with which they work lies right at the centre of why they form such a valuable resource to biochemists, biomedical researchers, evolutionary biologists and others.

6

Several major human drugs or diagnostics have been developed based on snake venom components. In addition, some fundamental biological processes have been revealed by using toxins as probes to study cells and their receptors. Here we review the state of this field and emphasize that the emerging field of high-throughput screening assays applied to a wide range of unstudied venoms, has the potential to provide a solid basis for the discovery of new lead compounds for new drugs.

## Forces driving evolutionary divergence in venom composition: diet, phylogeny, biogeography and ontogeny

Only in relatively recent years has the remarkable variability of venom composition at the genus, species, subspecies, population and even individual levels been fully appreciated, and the underlying ecological forces and mechanisms of mutation at the molecular level started to be identified and understood. Variation in venom composition has far-reaching implications. For the treatment of snakebite, for example, the importance of using pooled venom (mixtures of venom from several individuals of the same species representing different ages and geographical origins) has been emphasized in the production of antivenom to – in some cases - produce a serum that will be maximally effective against a bite from any snake of that species [9]. However, it should be noted that many antivenoms provide excellent – sometimes even better – cross reactions against antigens not even included in the original mixture, a phenomenon which is not yet well understood.

Since different species of prey differ in their physiological reaction to venom molecules, it follows that venom composition is linked to diet and varies from the species level all the way to individuals. In saw-scaled vipers (*Echis*), venom from those species that feed on arthropods was highly toxic to scorpions [10]. By contrast, scorpions were unaffected by venom from those species that feed on mammals [10]. Phylogenetic analysis revealed repeated instances of co-evolution of venom composition with prey preference. Other examples of variation in venom composition, at the individual to the species level, correlated with diet have been found in coral snakes (*Micrurus*) [11], rattlesnakes (*Crotalus* and

*Sistrurus*) [9,12-14], Malayan pitvipers (*Calloselasma*) [15], lanceheads (*Bothrops*) [16,17], puff adders (*Bitis arietans*) [18], and others.

Specificity of venom to prey type has also been found by studies approaching the question by analyzing the composition of the venom. A single species of coral snake (*Micrurus s. surinamensis*), distinctive within its genus in feeding on fish, was found to have venom containing neurotoxins lethal to fish and not known from the venom of other *Micrurus* species [11]. Significant differences in the protein composition were also found in the venoms of three subspecies of the rattlesnake *Sistrurus catenatus*, attributable to dietary differences between the subspecies [19]. Of the 11 protein families represented in the venom of this species, variation between the subspecies was found in all protein families, but variation was greater in some than in others. The metalloproteinases proved to be the most conserved while the PLA2s were the most divergent. PLA2s are often associated with neurotoxins – i.e the neurotoxin molecules also have PLA2 activity (e.g. notexin, taipoxin, crotoxin, β-Bungarotoxin etc.), so it could be that the neurotoxins vary more to accommodate the  different prey types.   Some toxins may assist in prey breakdown (digestion) but others – like neurotoxins causing paralysis or haemotoxins causing rapid low blood pressure or circulation – are perhaps more critical, they may determine whether the snake gets the meal in the first place.

In addition to diet, other variables such as phylogeny, biogeography and ontogeny may also be forces driving the evolutionary divergence of venom proteins (these factors may well be correlated with diet in many cases) – and perhaps just even separation time between populations due to genetic drift [20]. Although most studies documenting variation attribute it to a single factor, considering these categories in isolation may result in an incomplete understanding of the true scenario. For example, individuals of the common lancehead (*Bothrops asper*)  from two different localities and of different ages displayed significant differences in venom proteome [21]. Variation in symptoms of bite victims from individuals of this species varying in age or geographic origin has been observed though never formally documented. Ontogenetic differences in venom compositions are most probably explained by ontogenetic shifts in diet

[22]. The venom of neonates was found to differ significantly from that of adults, with a trend toward increasing complexity with age [22]. In the proteomes of neonate and adult individuals of the rattlesnake *Crotalus simus* the venom of adults and neonates had only 50% of their proteome in common, with a trend from neurotoxic changing to hemorrhagic properties from neonate to adult [23].

While ontogenetic variation certainly, and biogeographical variation in some cases, may well be explained by variation in prey type, observations of variation in the venom proteome of male versus female neonates from a single litter of the South American viper *Bothrops jararaca* [24] are puzzling and underscore just how much remains to be learned about the evolutionary forces driving variation in venom composition.

For the pharmacological prospector, variation is right at the heart of the potential goldmine of molecules that snake venom represents since a high degree of variation in venom composition increases the number of potentially useful novel molecules. Recognizing the crucial biomedical significance of variation in venom underscores the importance of efforts by those field biologists working with snakes to understand and conserve biological diversity, from the species down to the individual level. This emphasizes the importance of fieldwork in biodiscovery and the collection of venom samples from a wide range of species and specimens of different geographical localities and ages (Fig. 2).

**Figure 2** | Due to the enormous diversity in venom composition and the accelerated evolution of different isoforms of each of the toxin types in Fig. 1, fieldwork is essential to obtain samples from a wide range of species and specimens of different geographical localities and ages to exploit the full pharmacological potential of venom. A: a rare species of "colubrid", the green tree snake (*Dipsadoboa viridis*), in the Republic of Congo (Africa), photo by K.J. B: one of the authors (F.J.V.) with a wild king cobra (*Ophiophagus hannah*) – the longest venomous snake in the world - on the island of Java in Indonesia, photo by Smarley. C. The same author obtaining a venom sample from an Australian King brown (*Pseudechis rossignoli*), photo by H-W Herrmann. D. one of the authors (K.J.) removing a watersnake (*Grayia ornata*), mimic of the venomous Water cobra (*Naja annulata*) from a net in the Republic of Congo.

## Sources of toxin diversity: multiple splicing, exon insertion, exon switching, post-translational modification and domain switching

Snake toxin genes are the result of gene-duplications of normal body proteins that are subsequently selectively expressed in the venom gland [25], often followed by accelerated point mutations in the protein coding regions [26]. Gene duplication creates redundancy and allows the duplicated gene to escape the pressures of negative selection and acquire new functions through accelerated adaptive molecular evolution [26]. Acetylcholinesterase is the only currently known exception, because both the toxin and the normal enzyme are encoded by the same gene but differentially expressed using alternative splicing [27].

The molecular mechanisms that cause accelerated evolution – a bias towards nucleotide mutations that lead to amino acid changes (nonsynonymous substitutions) compared to those that don't (synonymous substitutions) – are currently not understood. When the first snake genome becomes available, this may shed light upon the molecular mechanisms that operate on the venom genes. Interestingly, nucleotide sequences appear to determine the accelerated rate of point mutations [27]. Specific triplets were found to be more 'stable' with regards to mutations than others, and the stable triplets were found to be higher in abundance in venom introns while the non-stable triplets were found higher in abundance in venom exons. There also appears to be a bias for transversions over transitions in nucleotide substitutions in some toxins [28].

There are several mechanisms by which molecular diversity of venom toxins is generated. First, alternative splicing allows multiple different functional proteins to be created using the same exons (Fig. 3a) [29,30] as it may cause binding to different receptors [31] or change of target altogether. Second, exons may be inserted into existing genes - as in denmotoxin (Fig. 3b), a three-finger toxin in the venom of the "colubrid" mangrove catsnake (*Boiga dendrophila*) with bird-specific activity [32]. Third, part of the intron may be retained in the mRNA due to error in splicing (Fig. 3c). Also, a recent discovered mechanism termed accelerated segment switch in exons to alter targeting (ASSET) may play an important role in

generating the molecular diversity in snake venom molecules [33] (Fig. 3d). During ASSET certain parts of exons are changed through accelerated segment switch and generate a functionally new toxin with a conserved structural fold. Sometimes, synergistic action between two different toxins may enhance their potency [34,35].

Post-translational modifications such as disulphide bridge formation and proteolysis play important roles in structural modification and acquisition of new functional sites. Both covalent and non-covalent interactions between similar and dissimilar proteins can form complexes that may exhibit a much higher pharmacological activity compared to the individual components. Sometimes formation of hetero/homo dimeric or trimeric complexes may lead to recognition of new targets as protein-protein interaction in complexes may expose critical amino acid residues that were otherwise buried in monomers [28].

New interaction sites are also formed in certain proteins to exhibit higher pharmacological potencies through domain swapping. In domain swapping, exchange of identical structural elements takes place between two or more molecules to form dimers or oligomers. For example, with the exchange of domains by α and β subunits of IX/X binding protein, isolated from the venom of the Okinawa habu (*Protobothrops flavoviridis*) [36], the hinge region forms a concave structure between the subunits and provides a new functional site in the heterodimer [28]. Without this swapping the loop would fold back and the ligand binding site would be absent.
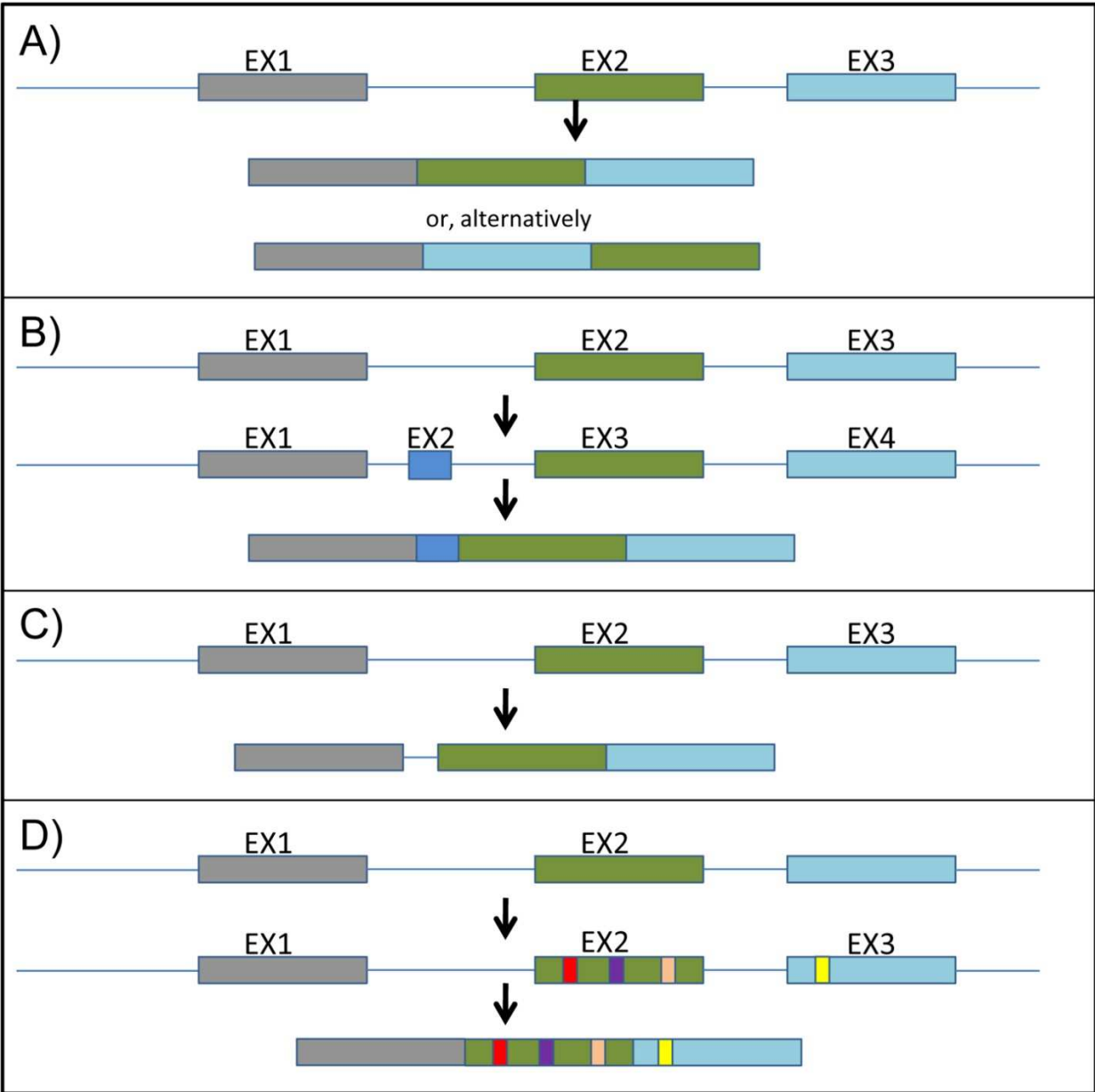
**Figure 3** | Some molecular mechanisms by which the diversity of toxin proteins is achieved. A: normal and alternative splicing. B: insertion of exon, as in denmotoxin [32]. C: intron retention due to error in splicing. D: accelerated segment switch in exons to alter targeting (ASSET) sas in a 3FTX from the venom of the rattlesnake *Sistrurus catenatus* [33]. Colored segments (red, purple, pink, yellow) represent exchanges of segments.

## Snake venom enzymes and toxins

Of the following enzymes, some have currently been described in all snake venoms, and others from only a limited number of species: PLA2s, metalloproteinases, serine proteases, acetylcholinesterases, LAOs, and hyaluronidases (Fig. 1). Two important and diverse families are the PLA2s and the metalloproteinases. PLA2 enzymes exhibit a wide variety of pharmacological effects in prey/victims and are therefore interesting for the pharmacological prospector [37]. Catalytically inactive Lys49 PLA2 homologs – found in many viperid venoms and being strongly myotoxic – work through a mechanism independent of hydrolysis [38]. Furthermore, although different in their pharmacological and enzymatic activity, the PLA2-like and PLA2 toxins are highly similar in sequence and structure and differ only for the substation of Asp-49 with Lys or Ser [39] - an example of how puzzling the mechanisms of toxicity can be. Some of the neurotoxic PLA2 are either homo/heterodimeric complexes. For example, β–bungarotoxin contains a covalently linked kunitz-type serine protease inhibitor which is involved in the blocking activity [40].

SVMPs are responsible for major local symptoms in snakebite - causing hemorrhage, edema, hypotension, hypovolemia, inflammation and necrosis. They are divided into three groups (P-I to P-III) based on the presence of other domains in the mature protein [41]. PI SVMP's have a prodomain and a single metalloproteinase domain that overall causes less hemorrhagic action then the other types, but still displays a variety of biological activities. The PII SVMP's are composed of a metalloproteinase and disintegrin domain, along with a prodomain.

Non-enzymatic venom proteins include 3FTXs, Kunitz-type serine protease inhibitors, sarafotoxins, CRISP, disintegrins, C-type lectins, waprins, veficolins and vespryns. 3FTXs have three beta-stranded loops resembling three-fingers and are mainly found in the venoms of elapids, some "colubrids" [32,42] and, in low quantity, in some viperids [43,44]. The structure is stabilized by four conserved disulphide bridges [45,46]. However, these structurally related molecules differ greatly in their biological functions [47,48]. Functional characterization of this family of proteins has contributed significantly to our

understanding of the mechanisms of venom toxicity and of normal physiological processes [48]. For example, characterization of α-bungarotoxin, a three-finger neurotoxin found in the venom of the banded krait (*Bungarus multicinctus*), enabled the isolation of the human nAChR [49] and contributed to our understanding of myasthenia gravis [50]. Significant contributions have been made in determining the distribution of specific receptors or ion channels in particular tissues or cells, identification of subtypes of receptors, imaging receptor trafficking [51-53] as well as in the development of therapeutic agents for treatment of adrenomyeloneuropathy and multiple sclerosis [54].

Snake venom kunitz-type serine protease inhibitors like dendrotoxin, calcicludine and the B chain of β-bungarotoxin, act as $Ca^{2+}$ and $K^+$ channel blockers respectively [55-57]. Textilinin, a kunitz-type serine protease inhibitor from the venom of the highly dangerous Australian brown snake (*Pseudonaja textilis*) is a reversible plasmin inhibitor and has promising potential for development of anti-bleeding agent [58].

Waprins have only recently been identified in snake venom and show homology to whey acidic proteins [59]. So far their functions are not really understood except for omwaprin, isolated from the Inland taipan (*Oxyuranus microlepidotus*), which possesses selective antimicrobial activities [60], useful for developing potential antibiotics.

CRISP toxins have molecular weights of 20-30 kDa and 16 conserved cysteine residues. Functional characterization of some CRISP has revealed that they are involved in blocking cyclic nucleotide-gated ion channels and block potassium-stimulated smooth muscle contraction [61]. Therefore in addition to their involvement in disrupting the normal physiological functions of victims they can be used for studying ion channel chemistry. A CRISP toxin from the venom of the "colubrid" snake the Patagonia green racer (*Philodryas patagoniensis*) was recently shown to cause damage to the murine gastrocnemius muscle, an action never before shown for any CRISP toxin [62] – showing the potential "colubrid" venoms have to find new toxins.

Snaclecs comprise two subclasses of protein, C-type lectins (CTLs) and C-type lectin-related proteins (CLRPs), found in venoms of most families of advanced snakes. Snake venom CTLs are involved in hemagglutinating and platelet aggregation activities during envenomation [63-66]. The CLRPs are involved in anticoagulant, procoagulant and agonist/antagonist of platelet activation [67,68]. Pharmacological characterization reveals that they either enhance or inhibit the function of coagulation factors which underscores their potential in drug discovery for blood related diseases. Their immaculate specificity also helps in understanding platelet physiology.

Disintegrins are a class of non-enzymatic that bind to integrins expressed on platelets and other vascular endothelial cells as well as some tumor cells [69] – they are an important class of cell surface receptors that are critically involved in cell–cell and cell–matrix interactions and are therefore good candidates to understand the interaction between extracellular matrix and cells [70]. In addition to their role in antiplatelet activity, these molecules are used in the diagnosis of cardiovascular diseases as well as serving as prototypes for therapeutic molecules in treatment of cancer.

## The medicinal use of snake venom

A large number of venom proteins affect the haemostatic system [71] and can have procoagulant, anticoagulant, fibrinolytic or platelet active activities. Ancrod (Arvin) from the venom of the Malayan pitviper (*Calloselasma rhodostoma)*, Batroxobin (Reptilase) from the common lancehead (*Bothrops atrox*), Crotalase from the Eastern diamondback rattlesnake (*Crotalus adamanteus*) have all been used as defibrinogenating agents for several clinical conditions including deep vein thrombosis, myocardial infarction, pulmonary embolus, and many others [72]. Venoms with anticoagulant properties are extensively studied for possible medical applications. The drug Aggrastat (tirobifan) was developed from a compound in the venom of the saw-scaled viper (*Echis carinatus*), and is used as an anti-platelet drug (glycoprotein IIb/IIIa inhibitors) [73] and given to those with unstable angina (Fig. 4). Many venoms with procoagulant properties find application in the diagnosis of clotting abnormalities, and most of the clotting pathways can be assayed by some venom component. For example "Reptilase time" (*B. atrox)* assays for thrombin

16

inhibitors [74], "Ecarin" (*E. carinatus*) and "taipan time" (*O. scutellatus*) assays for phrothrombin, and

Russell's viper (*Daboia russelii*) venom assays for factor X and for monitoring anticoagulant therapy [75].
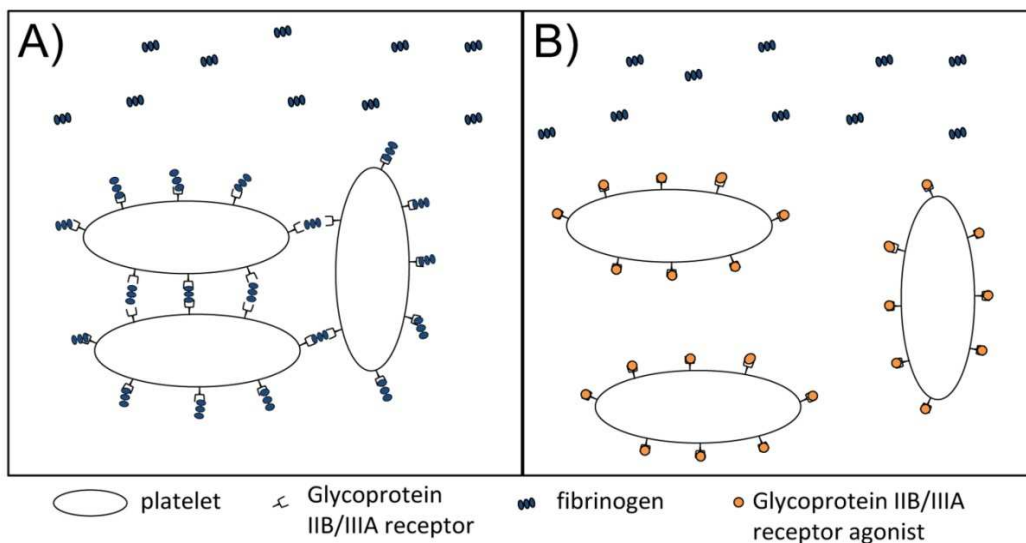


**Figure 4** | Mechanisms of action of the anticoagulant Aggrastat developed from a compound in the venom of the Indian saw-scaled viper (*Echis carinatus*). A: aggregated platelets by formation of fibrinogen bridges between the glycoprotein IIb/IIIa receptors. B: glycoprotein IIB / IIIA receptor antagonists like Aggrastat prevent platelet aggregation, and are mainly used in patients with acute coronary syndromes.

A number of snake venoms create a transient condition of depressed blood pressure in envenomed patients. ACE-inhibitors were developed from a bradykinin potentiating enzyme isolated from the venom of the Brazilian pitviper (*Bothrops jararaca*) and approved in 1979 by the FDA [76] to treat high-blood pressure and heart disease (Fig. 5). They work by blocking the switch between angiotensin-I and angiotensin-II, the latter being a vasoconstrictor. These inhibitors are now prescribed worldwide and have saved the lives of millions.

Many venoms have analgesic properties. Hannalgesin, derived from the venom of the king cobra (*Ophiophagus hannah*, see Fig. 2b) is already in clinical trials [77]. Promising toxins have been isolated from the tropical rattlesnake (*Crotalus durissuss terrificus*) and some other related species. Compounds derived from the Asiatic cobra (*Naja kaouthia*) are already in use in alternative medicine. Being more

powerful than morphine, cobra venom was used in the 1930's for treating intractable pain in cancer sufferers.

There is great deal of research currently been done into the anti-cancer properties of venoms. For example malignant brain and spinal-cord tumors (gliomas) are not curable by surgery because they invade the surrounding brain tissue without clear boundaries, making removal impossible. Disintegrins, like contortrostatin from American copperhead (*Agkistrodon contortrix*) venom, prevent cells from sticking together, and inhibit their interaction with surrounding tissue and result in a blockage of cell motility and invasiveness [78].

It has been demonstrated that fibrin(ogen) plays separate and distinctive roles at different stages of tumor growth and dissemination. At the primary site, fibrin deposition around the tumor could form a protective barrier, but also limit tumor progression. On the other hand, fibrin deposits formed by metastatic tumor cells may help disseminating these tumor cells [79]. Anticoagulants and the removal of fibrin could be an effective therapy. One of the earliest reports on the successful use of a venom defibrinogenating enzyme was that of Wood and Hilgard [80].
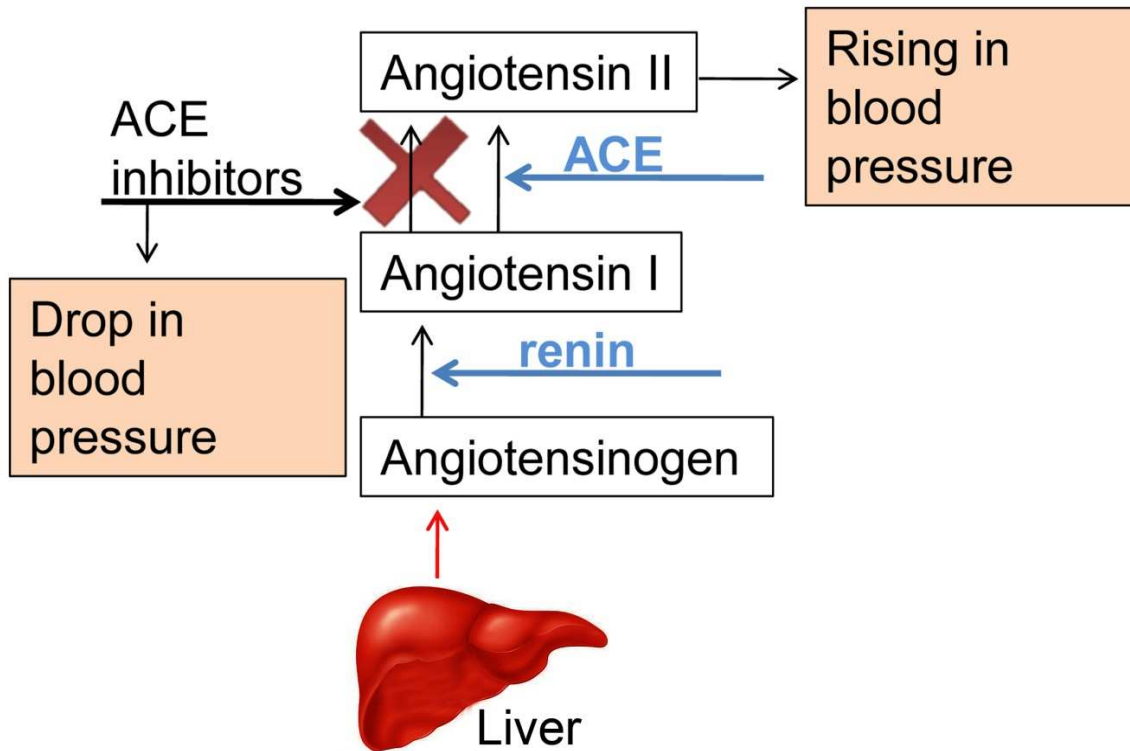
**Figure 5** | Mechanism of action of ACE inhibitors developed from a bradykinin potentiating enzyme isolated from the venom of the Brazilian pitviper (Bothrops jararaca) and approved in 1979 by the FDA [75] to treat high-blood pressure and heart disease. Angiotensin I causes vasoconstriction which raises the blood pressure.

Many cobra venoms contain Cobra Venom Factor (CVF), that activates and depletes the mammalian immune-complement system [81]. They are structural and functional analogues of the mammalian serum complement factor C3 [82]. CVF is used as a tool to study various aspects of the complement system and in this respect CVF is a uniquely useful venom component that has also been found useful as an immuno suppressant agent in tissue transplantation and in cancer therapy [83].

Many venoms have antibacterial properties. For example, Stiles *et al* [84] found 2 antibacterial bioactive L-amino acid oxidase components in King brown (*Pseudechis australis*) venom that were 70 and 17.5 times more effective *in-vitro* than tetracycline, a drug of choice for *Aeromonas* infections.

Antibacterial and antiparasitic effects of the venom of the Marajó lancehead (*Bothrops marajoensis*) were shown to be caused by PLA2 and and L-amino acid oxidase toxins [85].

Antiviral activity has been demonstrated in several venoms, although no commercialization of any of these compounds has yet taken place. The venom of the tropical rattlesnake (*Crotalus durissus terrificus*) from Brazil has been shown to be active against the measles virus [86]. The purified PLA2 venom neurotoxin "taipoxin" from the coastal taipan (*Oxyuranus scutellatus*), Nigexine from the African black necked cobra (*Naja nigricolis*) and a basic PLA2 from the Mozambique spitting cobra (*Naja mossambica*) have been shown to have potent anti-viral activities against HIV-1 virus [87].

Myasthenia gravis, a chronic autoimmune disorder affecting 2 out of every 100,000 people, results in progressive skeletal muscle weakness with rapid fatigue and loss of strength. It primarily affects mastication, facial and swallowing muscles and in advanced cases, respiratory muscles. Autoimmune antibodies destroy acetylcholine receptor sites at the neuromuscular junctions preventing nerve impulses from reaching the muscles. A rapid, quantitative and sensitive radioimmunoassay using human acetylcholine receptor, affinity labeled with purified neurotoxin, alpha-bungaratoxin, from the venom of the Taiwan krait (*Bungarus multicinctus*) is used to diagnose the condition [88,89]. Most venoms possessing alpha neurotoxins are potential candidates for this purpose.

## High-throughput screening to identify potential new medicines in venoms

A major obstacle in exploring snake venom for new leads into drug discovery is the low amount of venom usually obtained during the process of milking (extracting venom from the venom glands of a live snake, see Fig. 2c). In addition, it is difficult to extract venom from many of the "colubrid" snakes. Because of this, venom research has focused mainly on those snakes that are dangerous to humans, but this doesn't necessarily reflect the potential of their venom for drug discovery. Some of the harmless "colubrid" venom toxins have been shown to be as potent as those of some deadly elapids, only produced

in small amounts [4,90]. For example, the venoms of the Patagonia green racer (*Phildryas patagoniensis*) and Lichtenstein's green racer (*P. olfersii*) have been shown to posses high proteolytic activity, degrading fibrinogen and the vascular wall [91], and strong edematogenic and myotoxic activity [92], respectively. In addition, a completely new family of toxins has recently been found in the venom of the Asian dog-faced water snake (*Cerberus rynchops*). These toxins, called veficolins, may induce platelet aggregation and/or initiate complement activation [93]. Hence, the bias towards deadly snakes in exploring venoms for drug discovery is completely unwarranted.

Alternatively with the use of molecular tools the venom gland transcriptome can be complied. This approach not only allows us to understand the expression profile and the evolution of the venom proteins [94], but also reveals the lowly expressed proteins which will expand our resource for pharmaceutically active molecules.

Many of the medically significant venomous snakes produce hundred to a couple hundred milligrams of dry venom in a single milking [95]. Large species of elapids and viperids may produce from several hundred mg's up to more than a gram of venom in a single milking [95]. The recently described giant spitting cobra (*Naja ashei*) has been shown to produce the immense amount of 3 gm's of weight of dry venom [96], the largest amount ever collected during a single milking. The amount of venom that a snake produces during milking is determined by the species, its geographic origin, its body size and relative head size, and by the time of the year that it is milked, as well as by interactions among these factors, body size being the primary factor [97].

When biomedical researchers are looking for lead therapeutic agents against diseases, they can first look at those venoms of which the snakebite symptoms involve the same pathways as the disease, for example, using venom that causes vasodilatation to look for potential blood pressure regulators. However, snake venom is very complex and contains many different isoforms, some of which may only be present in very low quantities in the venom. For example, nerve growth factor is only present in a concentration

of about 0.1-0.5% [98]. A fraction that stimulates neurite outgrowth was first identified in tumour cell extracts, but venom from cottonmouth vipers (*Agkistrodon piscivorus*), was then found to be up to 6,000 times more potent [99] – a discovery that allowed the authors to study mechanisms which regulate cell and organ growth and for which they were awarded the Nobel prize in 1986. The problem is that venom compounds that are present only in low concentrations don't necessarily contribute to the bite symptomology and may thus be easily overlooked, although they can be of high potential biomedical interest. To exploit the full potential of snake venoms, a wide variety of them need to be fractionated and all fractions tested separately in a high-throughput screen of interest. Mice have typically been used as model organisms, in for example identifying genes involved in pain and anxiety, and in screening for analgesic peptides [100]. However, using mice is ethically controversial and low-throughput, screens also require relatively large amounts of the precious venoms. Zebrafish (*Danio rerio*) are now more and more used as model organisms to screen for new drugs, and this provides a significant window of opportunity for using snake venom (or venom from other venomous animals) to screen. Zebrafish are cheaper to maintain than rodents, their eggs and embryos are transparent so phenotypic changes can be visualized, and they can be easily scaled up into high-throughput assays [101] that require minimal amounts of venom, and automated imaging and analysis systems are available [102,103]. Morpholino knockdowns can be performed by injecting the yolk sac of embryos, allowing live imaging studies, and many disease-related genes identified in humans have orthologs in zebrafish [101]. Although zebrafish will never replace mammalian models completely in the drug development pipeline, they do form a cost-effective bridge between cell-based models on the one hand and rodent whole-organism models on the other [101].

## Conservation of snakes

A recent report by conservation biologists [104] documented a disturbing trend - a global decline in non-related snake species over the world during the same time period. Although they only looked at a small number of snakes (17 populations of eight species), we need to keep in mind that any species of snake that goes extinct may have held a new drug in its venom. Species having small home ranges, sedentary

habits and ambush feeding strategies are likely to be the most vulnerable as they rely on sites with specific types of ground cover that are disrupted by anthropogenic activities, and as ambush foraging is associated with a suite of life-history traits that involve low rates of feeding, growth and reproduction [105].

Current research on snake venoms is phylogenetically biased. Of the 1622 snake venom toxin sequences available on Universal Protein Resource database, only 49 are sequenced from snakes that do not possess any danger to humans. But with the high-throughput systems currently being developed this may hopefully soon change.

Very little data actually exist on the true conservation status of snakes worldwide. Out of 2711 extant caenophidian species, 26 are CITES listed, i.e. less than 1%, while 967 are listed on the IUCN Red List of Threatened Species, i.e. 36%. The main causes identified for their threatened status are caused by humans: annual & perennial non-timber crops (272 sp.), logging & wood harvesting (179 sp.), housing & urban areas (143 sp.), livestock farming & ranching (98 sp.), hunting & trapping (64 sp.) and natural system modifications (57 sp.). Although the majority of the species listed have been categorized as Least Concern (60%), this does not necessarily reflect their true status [104], and for another 22% of the species the data are insufficient to categorize them accordingly. In addition, 87% of the species (837) have been added to the Red List since 2007, highlighting the lack of proper attention in the years before.

If biogeography is considered, relatively well protected regions are North America (123 red listed sp.), Meso America (325 sp.), the Philippines (84 sp.) and Europe (29 sp.). On the other hand, several biodiversity hotspots such as South America (113 sp.), Sub-Saharan Africa (82 sp.), and the Caribbean Islands (16 sp.), among others are neglected while these areas may represent the locations with the highest potential for discovering snake species new to science, each representing new potential for advancing the frontiers of clinical medicine.

## Conclusions and future perspectives

While the use of snake venom for medicinal purposes dates back to ancient times, the past few decades have seen several new drugs derived from components of snake venom becoming available to patients worldwide. Here, we have reviewed recent advances in areas ranging from evolutionary to molecular biology that taken altogether indicate new possibilities for contributions to clinical medicine and medical research from snake venom. First, advances in snake systematics demonstrate a high degree of phylogenetic bias in previous molecular prospecting in snake venom, increasing the number of species from which medically useful molecules may be obtained. Secondly, discoveries in snake venom composition have revealed an unexpectedly high degree of variation of snake venoms of the same species linked to variables such as diet, geographical distribution, ontogeny and others. These two factors increase the potential pool that snake venom represents for pharmacological prospecting. Finally, new and innovative techniques in high throughput systems offer increased speed and efficiency in identifying and extracting desirable molecules from snake venom. In conclusion, we emphasize the challenges faced by the conservation of snake biodiversity, since it is ultimately on this that hopes for the development of new therapeutic agents from snake venom depends.

# Chapter 3: Evolutionary Origin and Development of Snake Fangs

Vonk FJ, Admiraal JF, Jackson K, Reshef R, de Bakker MAG, Vanderschoot K, van den Berge I, van Atten M, Burgerhout E, Beck A, Mirtschin PJ, Kochva E, Witte F, Fry BG, Woods A and MK Richardson.

All supplementary material can be found here:

http://www.nature.com/nature/journal/v454/n7204/suppinfo/nature07178.html

Many advanced snakes use fangs — specialised teeth associated with a venom gland [106-108] — to introduce venom during prey capture or defence. Various front- and rear-fanged groups are recognised, according to whether fangs are positioned anterior (e.g. cobras and vipers) or posterior (e.g. grass snakes) in the upper jaw [109-111] . A fundamental controversy in snake evolution is whether or not front and rear fangs share the same evolutionary and developmental origin [107,109-114]. Solving this controversy would point to a major evolutionary transition that underlies the massive radiation of advanced snakes, and to the underlying developmental events. We examine this issue by visualising the tooth-forming epithelium in the upper jaw of 96 snake embryos, covering eight species. We use the *sonic hedgehog (shh)* gene as a marker [115-118], and reconstruct the development in 3D in 41 of these. Here we show that front fangs develop from the posterior end of the upper jaw, and we reveal a striking similarity in morphogenesis between front and rear fangs that argue for their homology. In front-fanged snakes, the anterior part of the upper jaw lacks *shh* expression, and ontogenetic allometry displaces the fang from its posterior developmental origin to its adult front position — suggesting a posterior evolutionary origin. In rear-fanged snakes, the fangs develop from an independent posterior dental lamina and retain their posterior position. In light of our findings, we suggest a possible new model for the evolution of snake fangs; a posterior sub-region of the tooth-forming epithelium became developmentally uncoupled from the remaining dentition. This would have allowed the posterior teeth to evolve independently and in close association with the venom gland, becoming highly modified in different lineages. This developmental event could have facilitated the massive radiation of advanced snakes in the Cenozoic era, resulting in the spectacular diversity of snakes seen today [5,107,119].

Many advanced snakes (Caenophidia, *sensu* ref. 22) use venom, with or without constriction, to subdue their prey [5,107]. Their venom-delivery system includes a post-orbital venom gland associated with specialised venom-conducting fangs[108]. Fangs can occupy various positions on the upper jaw, but are always located on the maxilla and never on any other tooth-bearing bone [120] (Fig. 6C). Viperidae (vipers & pitvipers), *Atractaspis* (Lamprophiidae, *sensu* ref. 22), and Elapidae (cobras and relatives) have tubular front fangs  (Fig. 6B,C). The remaining lineages are all non-front-fanged and called either 'non-fanged' (no distinguishable enlarged posterior tooth) or rear-fanged [107,120]  (Fig. 6B,C). Rear fangs can be solid, slightly or deeply grooved, but never tubular [120].

There has been active debate concerning the evolutionary origin of these different fang types, and their relationships to the simple, unmodified teeth of non-fanged [121] basal snakes [107,109-114,121] such as pythons and boas (Boidae). Proposed hypotheses include: (i) front-fanged snakes form a monophyletic group and their fangs are derived from rear-fangs [113,122,123] ; (ii) elapid fangs are derived from front teeth and viperid fangs from rear fangs [124,125] ; (iii) elapid and viperid fangs are both derived from rear-fangs [112] ; and (iv) front and rear fangs have independent origins [107]. Establishing the origin and evolutionary transformation series between these dentition types requires a robust phylogeny to map the characters onto. Since recent molecular phylogenies of caenophidians place the front-fanged Viperidae as relatively basal and the front-fanged Elapidae as more recently derived [1,107] (Fig. 6A), the current evidence seems to support the 'independent-origin' hypothesis [107].
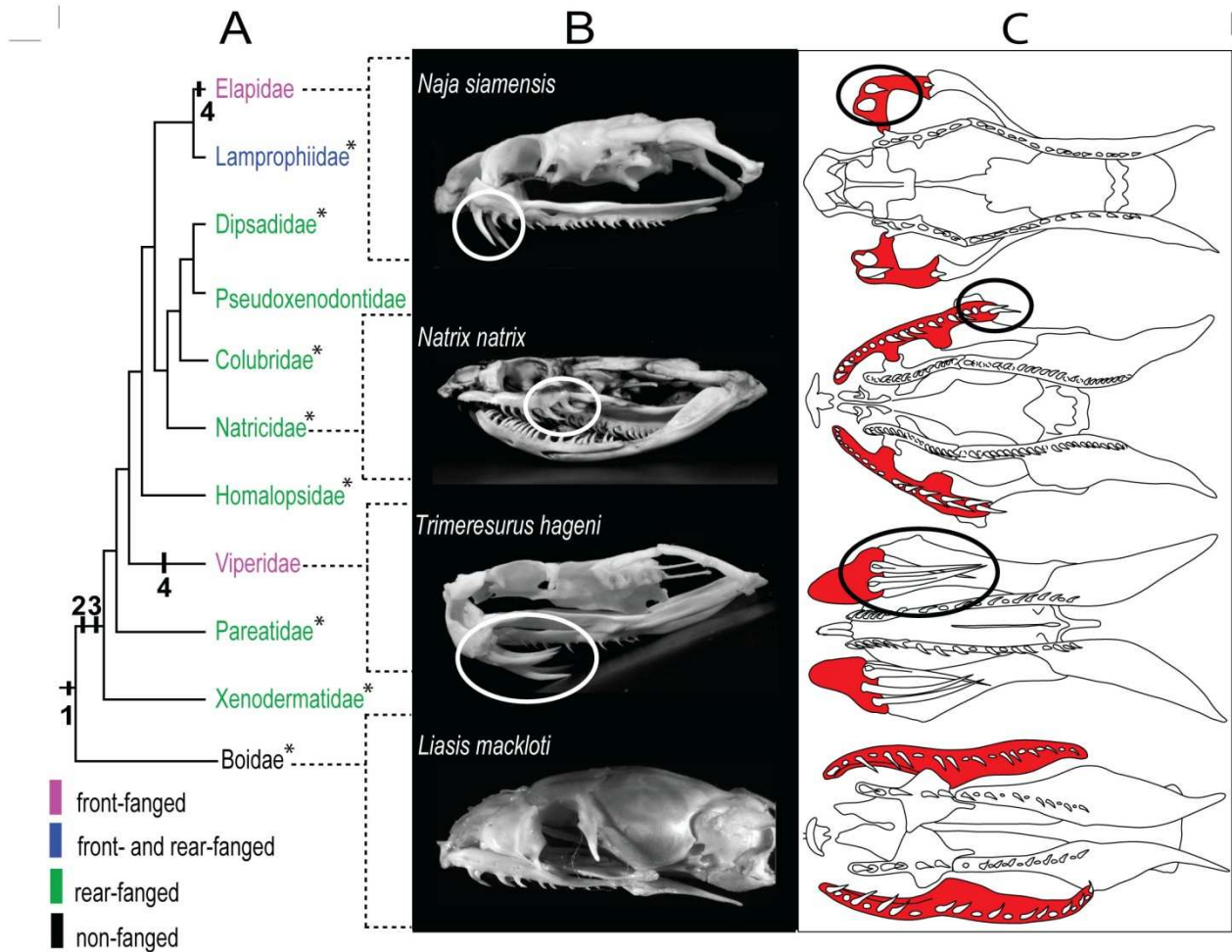
**Figure 6** | Adult maxillary dentition mapped onto a molecular snake phylogeny to show relative positions of the various fang types. a, Phylogeny from ref. 16. b, c, Adult skulls (Supplementary Table 4): lateral views (b); palate, schematic ventral views (c; maxilla coloured, fangs circled). Asterisks indicate species studied by electron microscopy (Supplementary Fig. 5, Supplementary Table 3). The evolutionary changes leading from an unmodified maxillary dentition to the different fang types in advanced snakes are indicated at the nodes: (1) continuous maxillary dental lamina, no specialized subregions—ancestral condition for advanced snakes; (2) evolution of posterior maxillary dental lamina—developmental uncoupling of posterior from anterior teeth; (3) starting differentiation of the posterior teeth with the venom gland; (4) loss of anterior dental lamina and development of front fangs.

To date, this issue has not been examined using a molecular developmental approach, not least because of the difficulty of obtaining snake embryos of all the different species. Development of fangs and venom glands has been studied before in viperids [126,127], *Natrix* [128,129] (Natricidae), *Spalerosophis* [130], *Thamnophis* and *Telescopus* [106,108] (Colubridae). Those morphological studies identified a common primordium of the venom gland and fangs [106,126,128], but did not visualise the odontogenic band (tooth-forming epithelium: a band of epithelial tissue that invaginates and forms a dental lamina). Therefore, no conclusions could be drawn about the origin and evolutionary transformation series of the fangs.

Here, we have carried out *in situ* hybridisation of the *sonic hedgehog (shh)* gene in 96 snake embryos of multiple stages, and reconstructed the development of the maxillary dentition in 3D in 41 of these through serial sections. We use nine advanced snake species — comprising two front-, one rear- and one 'non-fanged' lineage. As outgroup we included the non-fanged waterpython (*Liasis mackloti*: Boidea), which is basal to advanced snakes (Fig. 6). *Shh* is expressed in the odontogenic band in different vertebrate species [115-117]. By visualising this band, we aim to find evidence for the ancestral condition of the maxillary dentition. A list of all material studied can be found in Supplementary Tables 1-4. We map our characters onto the recently published, robust molecular phylogeny of advanced snakes obtained by Vidal et. al. 2007 (ref. 22, and Fig. 6A).

In the waterpython (*L. mackloti*), *shh* expression reveals one continuous maxillary odontogenic band (Fig. 7a). As confirmed by serial sections of embryos ranging from young to old, this band invaginates to form one dental lamina — a single continuous, invaginating epithelium that will develop a row of teeth (Fig. 8a & Fig. 9j-l). This is consistent with a recent morphological study of *Python sebae* [131]. The odontogenic band, with its associated lamina, appears along the entire rostrocaudal extent of the upper jaw — from the premaxilla to the mandibular articulation (Fig. 7a). This suggests that the ancestral condition for the maxillary dentition of advanced snakes is one dental lamina that appears along the entire rostrocaudal extent of the upper jaw, lacking specialised sub-regions.

The early odontogenic band in the non-front-fanged grass snake *Natrix natrix* (Natricidae) and rat snake *Elaphe obsoleta* (Colubridae) is similar to that of the waterpython (Fig. 7c & Suppl. Fig. 2f,i) However, we show for the first time that there are two dental laminae which invaginate separately (Fig. 8b,e & Fig. 9a-c & Suppl. Fig. 8e-g) and fuse during development (Fig. 9c & Suppl. Fig. 3h-i). The anterior lamina bears only teeth (Fig. 9c & Suppl. Fig. 3f) and is similar in development to the one in the waterpython (Fig. 9j,k&l). The posterior lamina, however, bears teeth and forms the common primordium with a post-orbital gland (Fig. 8b & Fig. 9b-c & Suppl. Fig. 3i). These develop into the rear fangs and venom gland in the grass snake, and probably represent the first differentiation of the posterior teeth with a venom gland in the rat snake. The latter observation is consistent with a recent MRI and histology study [132], showing the presence of a small gland in rat snakes. To verify that the anterior and posterior dental laminae are truly developmentally independent, we ablated the primordium of the anterior lamina in isolated developing upper jaws of the dice snake *Natrix tessellate* [133] (Suppl. Fig. 3m-o) [134]. After cultivation under the yolk sac membrane, we found that the posterior lamina, with its venom gland and fangs, developed normally in the absence of the anterior lamina (Suppl. Fig. 4), showing that they are developmentally independent.
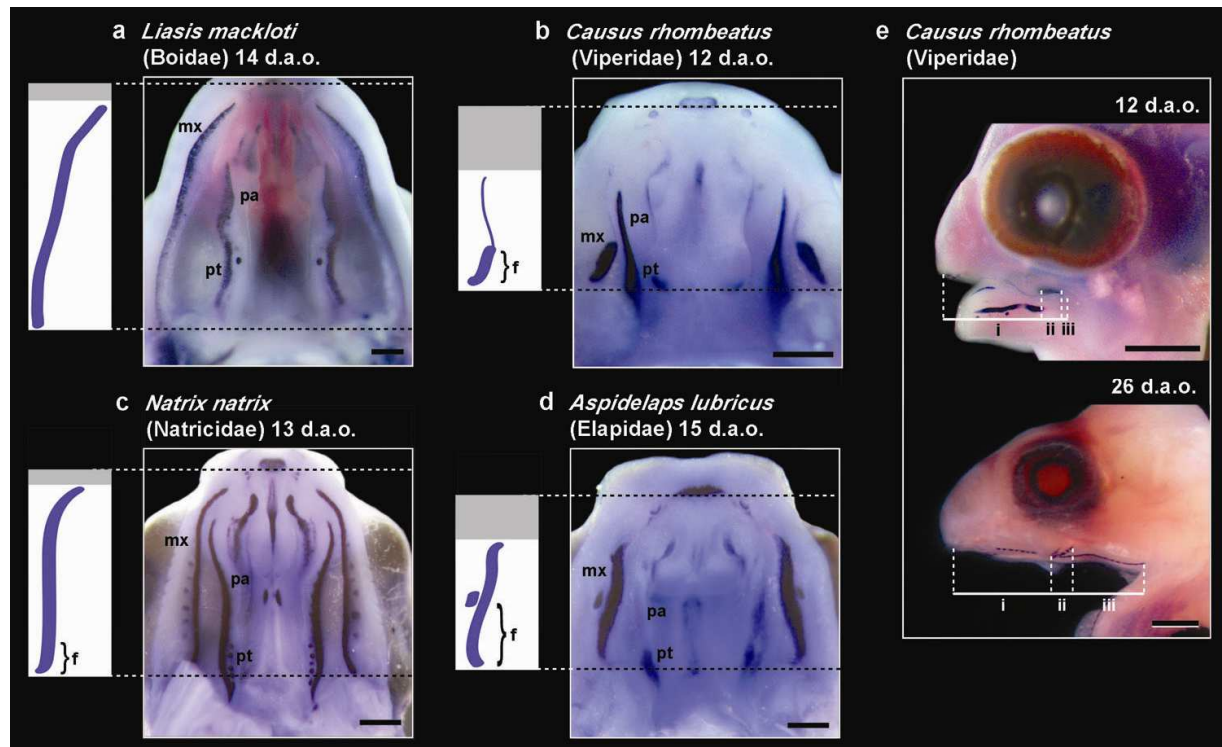
**Figure 7** | *Shh* expression in the embryonic snake palate, showing the posterior developmental origins of front fangs. a–d, Palate, ventral view: top, anterior; scale bar, 0.5 mm; dotted lines, upper jaw (posterior margin of premaxilla to attachment of the mandible); boxes, schemes of maxillary odontogenic band (purple, *shh* expression; grey, no *shh* expression). Positions of fangs in b–d were identified histologically (Fig. 3, Supplementary Fig. 3). The odontogenic band in the front-fanged species is located posterior in the upper jaw (b, d). In the non-fanged outgroup (a) and the rear-fanged *Natrix* (c), the odontogenic band extends along the entire upper jaw. f, fang; mx, maxillary odontogenic band; pa, palatine odontogenic band; pt, pterygoid odontogenic band. e, Ontogenetic allometry in the fang in the front-fanged *Causus* displaces the fang along the upper jaw (Supplementary Figs 5–9, Supplementary Tables 5–9). Scale bars, 1 mm. We note the change in relative size of the upper jaw subregions: i, anterior; ii, fang; iii, posterior. d.a.o., days after oviposition.
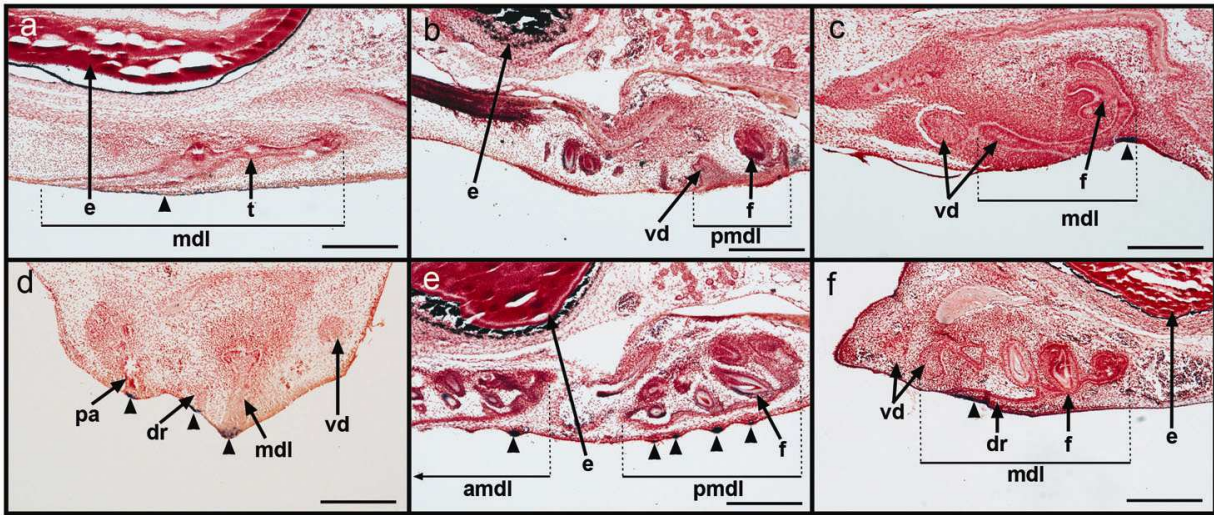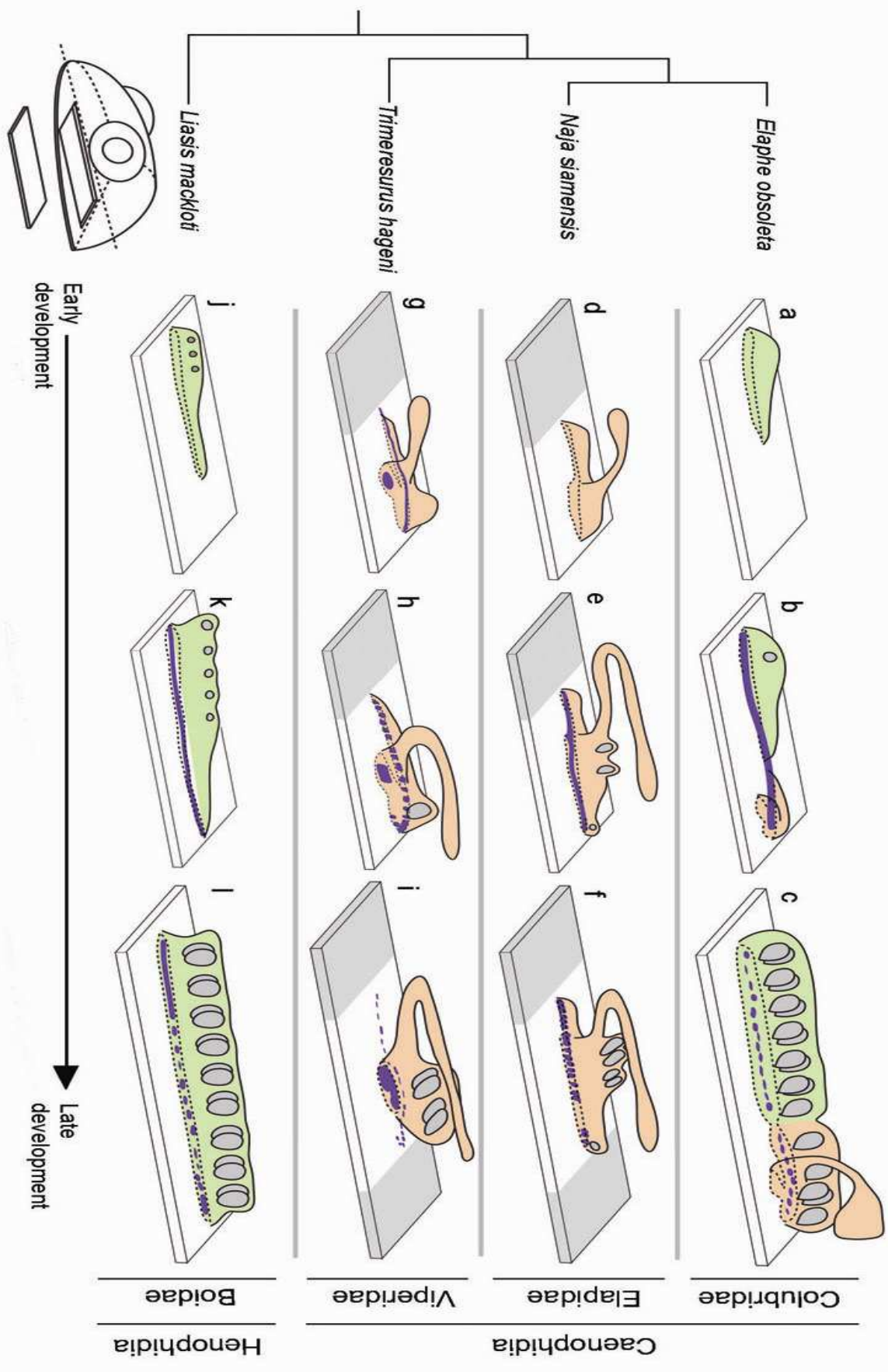
**Figure 8** | Sections of the shh in situ hybridizations of the embryonic upper jaw in five snake species, showing the posterior and anterior dental laminae. a–c, e–f, Sagittal sections, anterior to the left, of *L. mackloti* (Boidae) 22 d.a.o. (a), *N. natrix* (Natricidae) 22 d.a.o. (b), *Calloselasma rhodostoma* (Viperidae) 8 d.a.o. (c), *N. natrix* 22 d.a.o. (e), *Naja siamensis* (Elapidae) 23 d.a.o. (f). d, transverse section, medial to the left, of *Trimeresurus hageni* (Viperidae) 8 d.a.o. The posterior maxillary dental laminae in b and e are similar in morphogenesis to the maxillary dental laminae in all front-fanged species examined (c, d, f; see also Fig. 4). Arrowheads, shh expression; amdl, anterior maxillary dental lamina; dr, dental ridge; e, eye; f, fang; mdl, maxillary dental lamina; pa, palatine dental lamina; pmdl, posterior maxillary dental lamina; t, tooth bud; vd, primordium of venom gland; scale bars, 300 mm.

In the five front-fanged species examined (Viperidae and Elapidae), the maxillary odontogenic band is found in the posterior part of the upper jaw (Fig. 7b,d & Suppl. Fig. 2a,d,g). There is no *shh* expression nor dental lamina in the anterior region (verified by histology, data not shown). By contrast, in the waterpython, grass and rat snake, the odontogenic band and associated dental laminae appear along the entire rostrocaudal extent of the upper jaw. We find that, during development, the 'rear' fang is displaced to its adult 'front' position by ontogenetic allometry (Suppl. Figs. 6 & 7 and Suppl. Table 6 for statistical analyses); suggesting a posterior evolutionary origin of the front fangs. Histology shows that although the odontogenic band invaginates normally and forms one dental lamina (in contrast to the non-front-fanged snakes examined above), in all front-fanged species the fangs develop from the posteriormost part of this lamina and there are no developing teeth in the anterior part (Fig 8f & Fig. 9d-

32

h). This apparently toothless part of the dental lamina has been described before only in *Vipera aspis* (Viperidae) and termed the 'dental ridge' [128]. We show it here for the first time in Elapidae (Fig. 8f & Suppl. Fig. 3j-l). The fact that viperids and elapids share; (i) the 'dental ridge' and (ii) a posterior developmental origin for their front fangs is interesting, since they are phylogenetically not closely related (Fig. 6A).

**Figure 9** | (page 34) Schematic three-dimensional reconstructions showing the similarity in morphogenesis between the rear and front fangs. Derived from serial sections (Fig. 8, Supplementary Fig. 3); materials analysed are listed in Supplementary Tables 1, 2, 4. Left-hand side of the upper jaw is depicted, and only epithelial components are shown. Purple, *shh* expression; grey, tooth buds; green, unspecialized maxillary dental lamina; orange, specialized maxillary dental lamina that bears fangs. The specialized dental lamina is dilated into a bifurcated epithelial sac, the lateral part giving rise to the venom duct and venom gland by growing rostrad (see also Fig. 3b–d, f), then turning caudad to reach the post-orbital region (as previously described for vipers [126,127], *Natrix* [128,129] and *Spalerosophis* [130]. In *Elaphe obsoleta* (a–c) and *Natrix natrix* (data not shown), fangs develop rostrally and caudally alongside the base of the venom duct; in *Naja siamensis* (d–f) and *Trimeresurus hageni* (g–i) the rostral part regresses, remaining visible only as the dental ridge, whereas in b and c this part bears fangs and fuses with the anterior dental lamina. The unspecialized dental lamina in *E. obsoleta* (a–c) and the outgroup *Liasis mackloti* (j–l) starts developing anterior and grows caudad.

Early development → Late development

Liasis mackloti

Trimeresurus hageni

Naja siamensis

Elaphe obsoleta

a

b

c

d

e

f

g

h

i

j

k

l

Boidae

Viperidae

Elapidae

Colubridae

Henophidia

Caenophidia

34

Because Elapidae and Viperidae do not form a monophyletic group (Fig. 1A), the 'dental ridge', the posterior developmental origin of the fangs, and the allometric growth in both lineages may still reflect convergent evolution. However, our 3D reconstructions reveal that there is a striking similarity in morphogenesis of all front and rear fangs examined (Fig. 9b-i), despite the large variation in adult morphology. The toothless 'dental ridge' seen in elapids and viperids is similar to the part of the posterior dental lamina that fuses with the anterior dental lamina in the grass snake and the rat snake (Fig. 9). Although developmental similarity is not conclusive proof of structural homology, this is especially interesting in light of the posterior developmental origin of the front fangs in both elapids and viperids mentioned above. These results are difficult to reconcile with the 'independent-origin' hypothesis, but are consistent with the hypothesis that elapid and viperid front fangs, and the posterior dental lamina in non-front-fanged snakes, represent homologous structures.

Our results suggest a possible model for the evolution of snake fangs. A posterior sub-region of the ancestral tooth-forming epithelium became developmentally uncoupled from the remaining dentition, resulting in a posterior and anterior dental lamina that are developmentally independent (Suppl. Fig. 1). This condition is retained in the non-front-fanged snakes, such as the grass and rat snake. This model would imply that the front-fanged elapids and viperids have independently and secondarily lost the anterior dental lamina (Fig. 6), which is supported by the lack of *shh* expression anterior in their upper jaws.

Since obtaining developmental data for each non-front-fanged advanced snake lineage is impracticable, convergence cannot be ruled out completely. We have, therefore, examined the adult maxillary tooth morphologies through scanning electron microscopy in the waterpython, the grass and rat snake, and a further wide range of non-front-fanged advanced snake species (Fig. 6A). We aimed to find differences in the maxillary dentition, which might suggest the presence of two maxillary dental laminae in additional lineages. Our results show, indeed, that there is a consistent difference in anterior versus

posterior tooth morphologies in additional advanced snake lineages (Suppl. Fig. 5d-x, and Suppl. Table 3). In contrast, the maxillary teeth from the Boidae do not show a morphological difference between anterior and posterior teeth (Suppl. Fig. 5f). These results suggest the possible presence of two dental laminae in other non-front-fanged advanced snake lineages, and provides additional support to our proposed model.

The developmental uncoupling of the posterior from the anterior tooth region represents a possible mechanism that could have allowed the posterior teeth to evolve independently and in close association with the venom gland. Subsequently, it could became modified and form the fang-gland complex — an event which underlies the massive radiation of advanced snakes during the Cenozoic era [107,119,132].

## Methods

**Snake embryos.** Snake eggs and embryos were acquired in accordance with local and international regulations from European and Australian breeders and zoos. Eggs were incubated at 30°C and embryos fixed in 4% paraformaldehyde in PBS at 4°C overnight. They were dehydrated through graded methanols and stored at -18°C.

***In situ* hybridization.** The RNA probe is based on the partial PCR product of SHH using the cDNA of a 1-day-old Rhombic Night Adder (*Causus rhombeatus*) embryo as template (GeneBank accession number EU236145). Hybridization was performed according to standard protocols. In all species examined, the odontogenic band (tooth-forming epithelium) always expressed *shh* (Fig. 2 & Suppl. Fig 2). This shows that, as in other vertebrate groups [115-117], *shh* is also a marker for odontogenic epithelium in snakes. A list of embryos studied can be found in Suppl. Table. 1.

**Histology.** Embryos where dehydrated through ethanols or methanols, cleared in HistoClear or tetrahydronaphthalene, and embedded in paraffin. Sections were cut at 5-7 microns and counterstained using Neutral Red or H+E. A list of embryos studied can be found in Suppl. Table. 2.

**3D modeling.** Schematic 3D models were drawn from analyses of the serial sections of the embryos using a Nikon Eclipse E800 microscope. All models were drawn using Adobe Illustrator and Adobe Photoshop.

**Scanning electron microscopy.** The maxillary bone on one side was dissected out of the specimen, allowed to dry and mounted on a stub, using double-sided tape, with teeth pointing upwards. Specimens were sputter-coated with gold, and examined using a JEOL JSM-T300 scanning electron microscope, at an acceleration voltage of 15kV. A list of specimens examined with their museum numbers can be found in Suppl. Table. 3.

**Ablation experiment.** Was performed as previously described [133].

**Author Contributions** F.J.V.: study concept and design; embryo and skull collection, acquisition and processing; probe synthesis; *in situ* hybridizations; histology; figures; paper. J.F.A.: three-dimensional reconstructions, figures. K.J.: scanning electron microscopy. R.R.: study concept, ablation experiment, histology. E.K.: study concept, embryological data. K.V., I.v.d.B., M.v.A.: in situ hybridizations, histology. M.A.G.d.B.: probe design, in situ hybridizations. A.B.: Naja embryo collection. P.J.M.: supply of Naja material. B.G.F.: study concept, supply of Causus material. A.W.: provision of laboratory space. E.B., F.W.: morphometric analyses. M.K.R. (project leader): study concept and design, provision of funding and laboratory space.

# Chapter 4: Axial Patterning in Snakes and Caecilians: Evidence for an Alternative Interpretation of the *Hox* Code

Woltering JM, Vonk FJ, Müller H, Bardine N, Tuduce IL, de Bakker MA, Knöchel W, Sirbu IO, Durston AJ and MK Richardson.

It is generally assumed that the characteristic deregionalized body plan of species with a snake-like morphology evolved through a corresponding homogenization of *Hox* gene expression domains along the primary axis. Here, we examine the expression of *Hox* genes in snake embryos and show that a collinear pattern of *Hox* expression is retained within the paraxial mesoderm of the trunk. Genes expressed at the anterior and most posterior, regionalized, parts of the skeleton correspond to the expected anatomical boundaries. Unexpectedly however, also the dorsal (thoracic), homogenous rib-bearing region of trunk, is regionalized by unconventional gradual anterior limits of *Hox* expression that are not obviously reflected in the skeletal anatomy. In the lateral plate mesoderm we also detect regionalized *Hox* expression yet the forelimb marker *Tbx5* is not restricted to a rudimentary forelimb domain but is expressed throughout the entire flank region. Analysis of several *Hox* genes in a caecilian amphibian, which convergently evolved a deregionalized body plan, reveals a similar global collinear pattern of *Hox* expression. The differential expression of posterior, vertebra-modifying or even rib-suppressing *Hox* genes within the dorsal region is inconsistent with the homogeneity in vertebral identity. Our results suggest that the evolution of a deregionalized, snake-like body involved not only alterations in *Hox* gene *cis*-regulation but also a different downstream interpretation of the *Hox* code.

The patterning and morphogenesis of the vertebrate skeleton during embryonic development is regulated by *Hox* genes. In tetrapods, the *Hox* gene family consists of 39 closely related homeobox genes consisting of 13 paralogous gene groups organized in 4 clusters (named A, B, C and D) (reviewed in [135]. These genes are expressed in collinear spatial domains along the antero-posterior body axis and act as effectors of vertebral identity. The sequential expression pattern of different *Hox* genes is often referred to as "the *Hox* code" [136] which, combined with different inductive properties of the genes, provides positional information for the antero-posterior regionalization of the axial skeleton. Typically, *Hox* genes form nested expression series and the domains of anterior genes overlap with those of more posterior genes. Where several genes are co-expressed, posterior genes act dominantly over anterior genes, a form of phenotypic suppression called "posterior prevalence" [137]. The evolutionary conservation of the *Hox* patterning system is reflected in the fact that expression boundaries of *Hox* genes correlate with the same anatomical transitions in different vertebrate species [138,139] and changes in *Hox* expression are often suggested to underlie macro-evolutionary modifications of the bodyplan[140].

Evolution has produced an astonishing diversity of body shapes adapted to different lifestyles. An important contributor to the shape of animals with backbones is the number of bones (vertebrae) that make up this structure. Some animals have gone to extremes (Fig. 10), none more so than snakes, which have more vertebrae than any other living animal — often more than 300, with some species having more than 500. One of the most striking modifications of the vertebrate axial skeleton is the evolution of a 'snake-like', elongated body plan as adopted by at least 7 extant taxa of reptiles, fish and amphibians independently[141] (Fig. 10). The evolutionary transition to a snake-like body plan is accompanied by a profound deregionalization of the axial skeleton which includes loss or reduction of the limbs and limb girdles, as well as sternal elements [141,142]. These species have a greatly increased number of vertebrae[141,143,144] and, with the exception of the atlas, all of the pre-cloacal vertebrae are usually rib bearing. In contrast to non-elongated species, the dorsal (in squamates the rib-bearing region of the trunk skeleton is termed 'dorsal' instead of 'thoracic') region of the skeleton is extremely homogeneous with

few or no differences between ribs and vertebrae along the antero-posterior axis. It has been hypothesized that this peculiar "deregionalized" limbless anatomy results from a corresponding homogenization of *Hox* expression domains along the snake's primary axis[145]. In this study we set out to analyse *Hox* expression in detail, including previously unexamined posterior paralog members in the corn snake (*Pantherophis guttatus*) and to a lesser extent, because of material scarcity, in a caecilian amphibian (*Ichthyophis* cf. *kohtaoensis*), which evolved a comparable body plan independently.
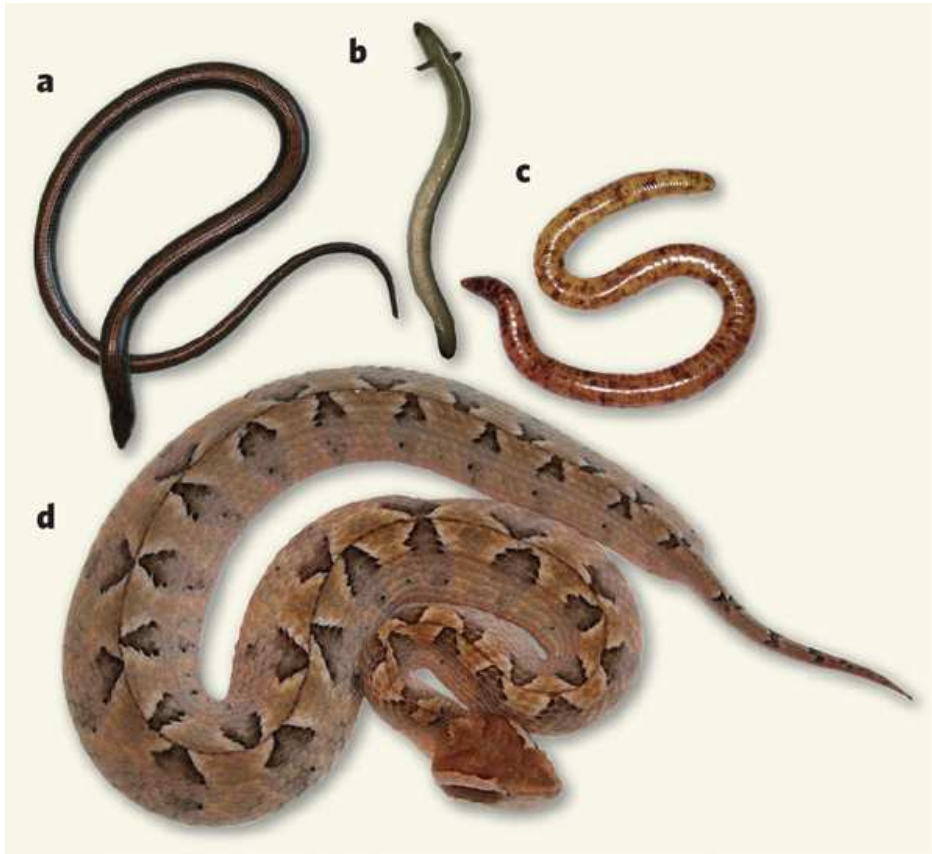


**Figure 10** | Body elongation, including increased numbers of vertebrae and a longer backbone, has evolved independently many times among vertebrates. a, Eastern glass lizard (*Ophisaurus ventralis*). b, European eel (*Anguilla anguilla*). c, Sao Tome caecilian ((*Schistometopum thomense*). d, Malayan pit viper (*Calloselasma rhodostoma*). Photos not to scale. (a,b,d, F.J.V. & M.K.R.; c, S. Blair Hedges.)

# Results

## Spatially collinear *Hox* expression in the snake somitic mesoderm

The axial skeleton of the corn snake consists of a total of 308 vertebrae (Fig. 11) and the anatomical subdivisions are listed in Table 1. Characteristic features include the complete absence of limb girdles, sternal elements, forelimbs and hindlimbs and the presence of homogeneous ribs on all but the 3 most anterior pre-cloacal vertebrae.
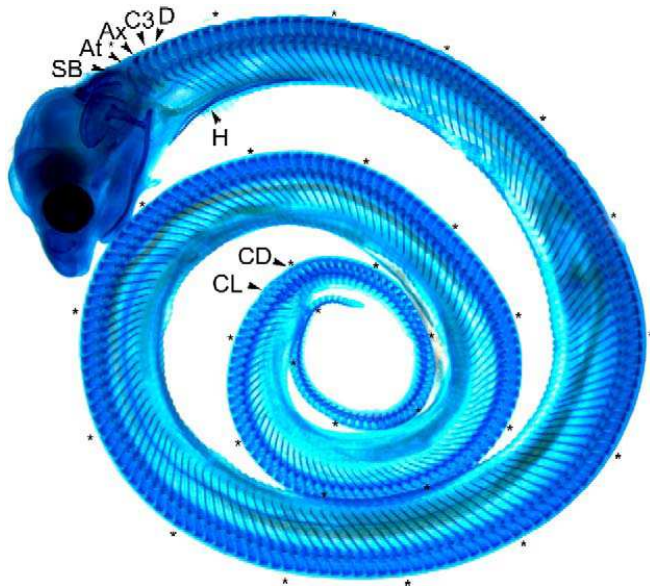


**Figure 11** | Alcian blue staining of an advanced corn snake embryo to show the skeletal anatomy. The axiak skeleton consists of 308 vertebrae; the atlas (At), the axis (Ax), 1 additional cervical vertebrae (C3) and 231 dorsal vertebrae (D), 4 cloacal vertebrae (CL) and 70 caudal vertebrae (CD). The first doral, cloacal and caudal vertebrae are indicated. SB; skull bone, H; hyoid.

We examined the expression of a panel of *Hox* genes in tailbud stage embryos (see Supplementary Material, Fig. S1). For all *Hox* genes examined, we observe, as reported previously[145], a strong homogenous expression in the posterior part of the trunk, but also a clear spatially collinear expression in the anterior part of their expression domains (Fig. 12 A-T and Table 1).

**Table 1 |** Corn snake axial formula and anterior somitic *Hox* expression limits

**Corn snake axial formula**

| Vertebra type | Position | Number | Rib |
|---|---|---|---|
| Atlas | 1 | 1 | No |
| Axis | 2 | 1 | No |
| Cervical | 3 | 1 | No |
| Dorsal (thoracic) | 4–234 | 231 | Yes |
| Cloacal | 235–238 | 4 | 'Forked' |
| Caudal | 239–308 | 70 | No |

**Corn snake *Hox* expression boundaries**

| Gene | Pre-vertebra level | Anatomical boundary |
|---|---|---|
| *HoxA3* | 1 | Atlas |
| *HoxB4* | 2 | Axis |
| *HoxC5* | 6 | – |
| *HoxA6* | 4 | D1 |
| *HoxB6* | 3 | C3 |
| *HoxC6* | 11–31 | – |
| *HoxA7* | 32–44 | – |
| *HoxB7* | 6–8 | – |
| *HoxB8* | 5–7 | – |
| *HoxC8* | 33–53 | – |
| *HoxB9* | 17–37 | – |
| *HoxC10* | ~200 | – |
| *HoxC13* | ~238 | ~Cloaca/caudal vertebrae |

Indicated here are the vertebral types, their position and number, and whether they are rib-bearing, the lymphapophyses of the snake cloacal vertebrae are indicated as forked ribs. Anterior somatic Hox expression boundaries are given below. Where *Hox* expression boundaries coincide with an anatomical transition this is indicated. A pre-vertebrae range is given in case of a diffuse expression limit.

The anterior expression boundaries of the most anterior and most posterior *Hox* genes examined, *HoxA3*, *HoxB4* and *HoxC13*, coincide with homologous morphological boundaries in the mouse and chicken [4, 12] (Fig. 12 A, B, S, T and Table 1). They are located at the atlas, axis and cloaca respectively and are thus expressed at expected positions within regionalized parts of the snake trunk.
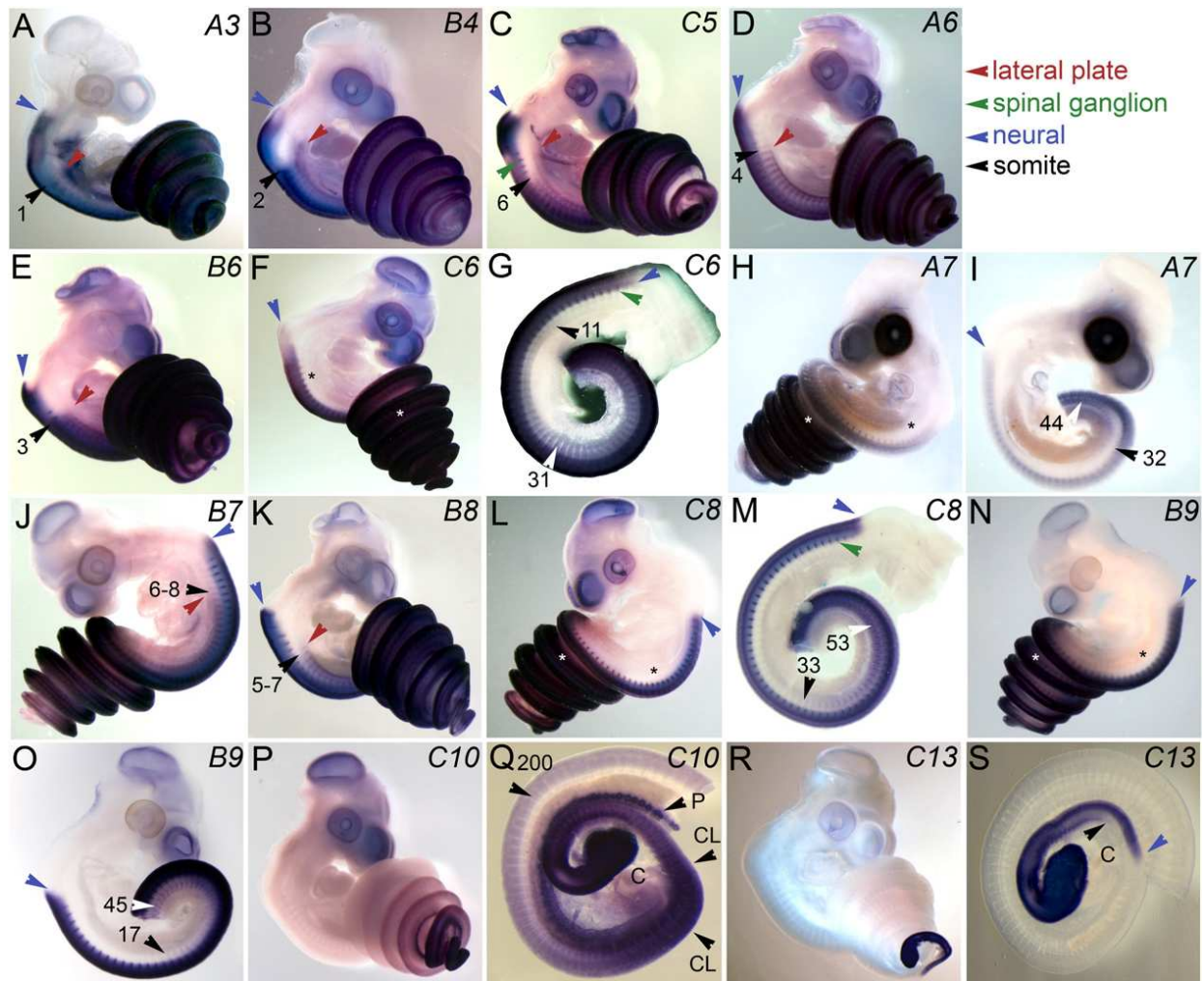


**Figure 12** | *Hox* gene expression in corn snake embryos. All embryos are at the second day after oviposition. *Hox* gene names are indicated in the upper right corner of each picture. (A-E,J,K). Expression of *HoxA3*, *HoxB4*, *HoxC5*, *HoxA6*, *HoxB6*, *HoxB7* and *HoxB8*. These anterior genes have sharp anterior expression boundaries in the somitic mesoderm indicated with a black arrowhead and a number referring to the pre-vertebra level. Red arrowheads indicate the anterior expression boundaries in the lateral plate mesoderm and blue arrowheads indicate the anterior limit of expression in the neural tissue. In some panels an example of spinal ganglion expression is indicated with a green arrowhead. For these *Hox* genes additional expression data is provided in the supplementary data (Fig.S2). (F-I, L-O) Expression of *HoxC6*, *HoxA7*, *HoxC8* and *HoxB9*. These anterior genes

45

have gradual anterior expression limits in the somitic mesoderm which are difficult to visualize exactly in the whole mount views (F,H,L,N,O). In these whole mounts a black asterisk is placed in the anterior somitic domain where no expression is detected and a white asterisk is placed in the posterior domain where the genes are expressed robustly. The expression limits for these genes are indicated more precisely in flat mounts of the anterior trunk region (G,H,M,O). In these images anterior expression limits in the somitic mesoderm are indicated with a black arrowhead and a number indicating the pre-vertebra level. This somitic domain is indicated again more posteriorly in a region of more robust expression with a white arrowhead and the indication of pre-vertebra level. In some panels an example of spinal ganglion expression is indicated with a green arrowhead and the anterior start of the neural expression domain is indicated with a blue arrowhead. For these *Hox* genes additional expression data is provided in the supplementary material (Fig.S3-6). (P-Q) *HoxC10* expression in whole mount (p) and flatmount of trunk/tail transition (Q). HoxC10 is expressed in the somotic mesoderm anterior to the cloaca (indicated "C") and cloacal vertebrae (indicated with black arrowheads and "CL") anterior to approximately pre-vertebra 200 (indicated with a black arrowhead and number) in a region of the trunk that will give rise to rib bearing vertebrae. Expression in pronephric vesicles is indicated with a black arrowhead and "P". Additional *HoxC10* expression data is provided in the supplementary material (Fig.S7). (R,S) Expression of *HoxC13* in whole mount (R) and flat mount of trunk/tail transition (S). *HoxC13* is expressed in the somitic mesoderm up till approximately the cloaca (indicated "C"), expression indicated with a black arrowhead. The anterior limit of neural expression is indicated with a blue arrowhead.

The dorsal region in snakes extends for most of the body length and has remarkable similar vertebrae. Despite this very poor regionalization, we do detect collinear expression in the dorsal region of the *Hox* genes belonging to paralogous groups 5-9. These genes are differentially expressed in the trunk and have been experimentally linked to the patterning of the thoracic and lumbar region in mouse and chicken [138,146-148] (Fig 12 C-O, Supplementary Material Fig S2, 3 and Table 1). The anterior boundaries of the most anteriorly expressed of these paralogous genes are sharply demarcated (*HoxC5*, *HoxA6* and *HoxB6*) or decrease in intensity over several somites (*HoxB7* and *HoxB8*). Genes expressed more posteriorly (*HoxC6*, *HoxA7*, *HoxC8* and *HoxB9*) show unconventional gradual anterior limits of expression, slowly fading out over the course of approximately 10 somites, but being completely absent from the anterior part of the trunk paraxial mesoderm. We analyzed expression of several genes at earlier stages (~150 somites: *HoxA6*, *HoxB6*, *HoxC6* and *HoxB9*; data not shown) or later stages (~ 5 days:

*HoxC8* and *HoxB9*; data not shown) and detected expression in the same pattern as observed in the stages shown in Fig. 12.

Snakes lack a lumbar region and rib-bearing vertebrae are present up to the cloaca. Genes from the *Hox-10* group suppress thoracic and induce lumbar fate in the mouse and expression of these genes is normally absent in the thoracic mesoderm[138,149-151]. In the corn snake by contrast, we detect expression of *HoxC10* within the rib-bearing part of the trunk (Fig. 12 Q, R, Fig. S4) and this expression thus conflicts with its presumed function as a rib suppressor (at least in snakes). As it has been suggested that *Hox-10* genes exert their rib suppressing activity in the pre-somitic mesoderm [150], we analyzed embryos at earlier stages and we also detected *HoxC10* expression in the pre-somitic mesoderm of rib-bearing somites (Supplementary Material Fig. S4).

Contrary to expectations, therefore, we find that the somitic mesoderm, which will give rise to the dorsal (thoracic) homogeneous rib-bearing part of the skeleton, is regionalized by *Hox* genes that are known to possess vertebra-modifying or even rib-suppressing activity in the mouse. Despite this regionalization, the dorsal skeleton shows no or very little regionalisation of vertebral anatomy.

## Comparison of *Hox* expression at the cervico-dorsal transition in snake and lizard embryos

Snakes have no well-defined transition from trunk to neck (cervico-dorsal) which in other squamates is marked by the position of the forelimb girdle (squamates often have rib-bearing vertebrae in the neck region hence presence of ribs is not a useful anatomical marker)[140,142]. The expression boundaries of *HoxC5*, *HoxA6*, *HoxB6*, *HoxC6* and *HoxB8* are located around the cervico-dorsal (thoracic) transition in mouse and chicken[138,148,152,153] and expression of the three *Hox-6* paralogs coincides at the level of 1st thoracic vertebra[138,153]. In the corn snake *HoxA6* is expressed as far anterior as pre-vertebra 4, *HoxB6* as far as pre-vertebra 3 while *HoxC6* is expressed gradually up to approximately pre-vertebra 11 (Fig. 13 and Supplementary Material Fig. S2). The expression of these genes thus seems shifted relative to each other, possibly in relation to the deregionalization of the snake body plan. However, due to lack of *Hox* expression data from non-snake squamates these differences in expression could also represent a synapomorphic squamate character. Therefore, we investigated the expression boundaries of *hoxC5*, *HoxC6*, *HoxB6* and *HoxB*8 in the bearded dragon lizard (*Pogona vitticeps*) and of *HoxA6* in the green anole lizard (*Anolis carolinensis*) (Fig. 14). These species have 4 rib-less neck vertebrae and the forelimbs are positioned between vertebra 6 and 7 (data not shown). *HoxC5* and *HoxB8* have anterior expression boundaries at the level of the 6th pre-vertebra, *HoxA6* is expressed at the level of 5th pre-vertebra (the first rib-bearing vertebra), *HoxC6* at the 7th pre-vertebra and *HoxB6* at the 3rd pre-vertebra (Fig. 14 A-E). This shows that a relatively normal cervico-dorsal transition, although slightly elongated, is patterned in snakes (Fig. 14 F). Interestingly, the expression of *HoxA6* corresponds to the transition from rib-bearing to rib-less vertebrae at pre-vertebra level 4 in the snake and pre-vertebra 5 in the anole lizard. However, all the other skeletal anatomical features associated with this transition are absent in snakes.
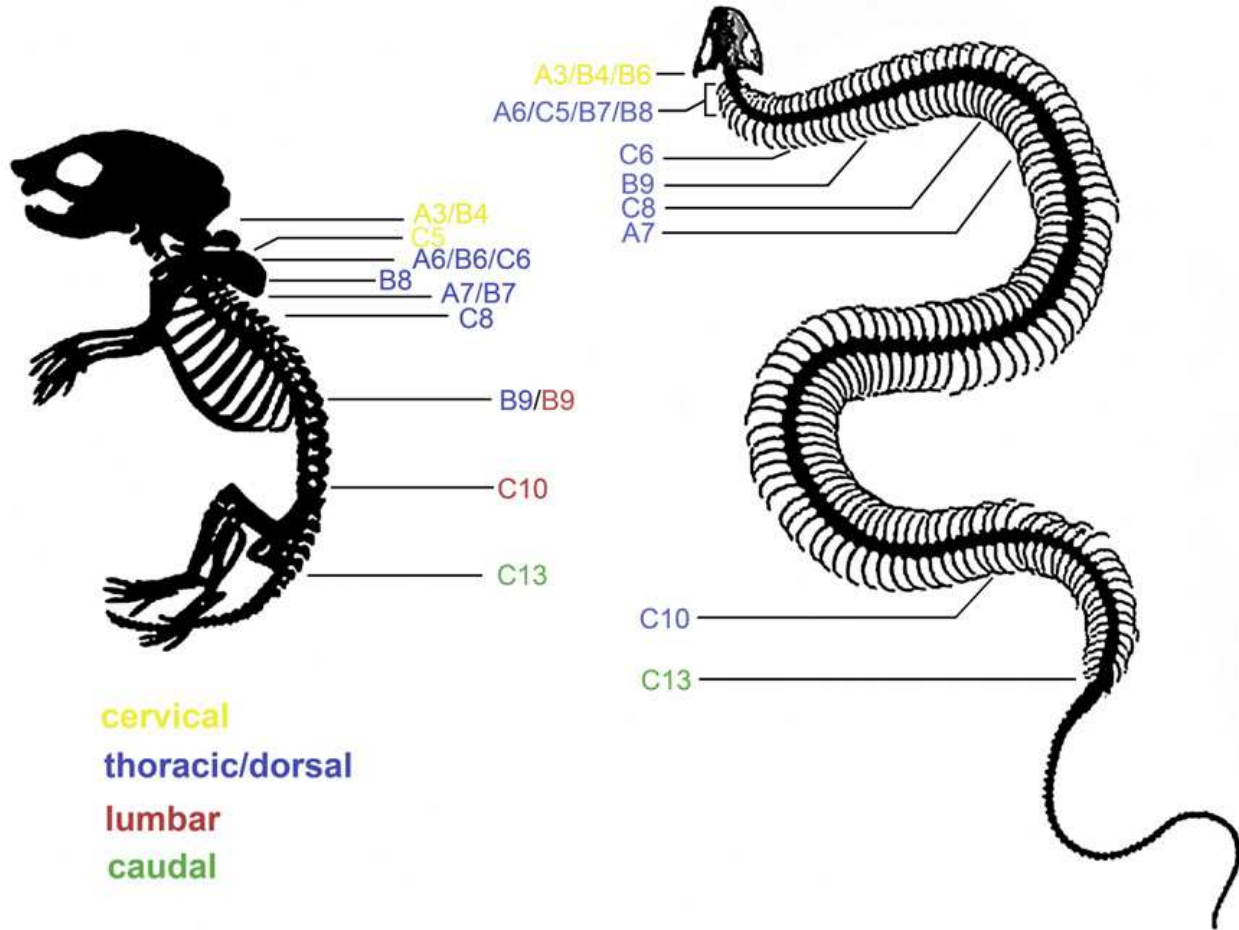
**Figure 13** | Schematic visualization of the approximate anterior mesodermal limits of *Hox* gene expression in the mouse and snake. Schematic drawing of a mouse and a snake skeleton, with indications of the approximate anterior boundaries of *Hox* gene expression, demonstrating a conservation of global collinear *Hox* expression zones in the snake despite absence of a correspondingly regionalized skeleton. Colours indicate the predominant anatomic region the genes are associated with. For precise levels of *Hox* expression zones in the snake, see Table 1, for the exact mouse expression data, refer to the cited literature. For the mouse, expression data from [138] were used for *HoxB4*, *HoxC5*, *HoxC6* and *HoxB7*. *HoxA7* is according to [154], *HoxA3* according to[139], *HoxC13* according to[155], *HoxA6* according to[153], *HoxB6* according to[152], *HoxB8* and *HoxC8* according to[148], and *HoxC10* according to[151]. The relative position of *HoxB9* is mapped according to its expression at the lumbar/thoracic transition in chicken[138]. In the mouse *HoxB9* expression appears however to be dynamically shifting along this axis, with at earlier stages more anterior and at later stages more posterior expression than indicated here [146,156]. Genetically the *Hox*-9 genes have been demonstrated to be involved in the patterning of the thorax as well as the suppression of anterior thoracic fate in the lumbar region [146,156], therefore the gene is indicated in the mouse as being both thoracic and lumbar.

## Conservation of global collinear *Hox* expression domains in a caecilian amphibian

Caecilian amphibians have independently evolved a snake like postcranial skeleton. We analysed *Hox* gene expression in embryos from a caecilian (*Ichthyophis* cf. *kohtaoensis*) which has a skeleton consisting of 126 vertebrae with the atlas being the only pre-cloacal rib-less vertebra (Fig. 15, axial formula given in Table 2). We examined expression of *HoxC5, HoxC6, HoxB8, HoxC8* and *HoxC13* and analyzed their expression in Brauer stage 26 embryos (Supplementary Material Fig. S1). The global expression patterns of these genes are comparable to snakes and they exhibit spatially collinear expression within the trunk (Fig. 16, Supplementary Material Fig. S5 & Table 2). Similarly to snakes, the expression domains of *HoxC5*, *HoxC6*, *HoxB8* and *HoxC8* are not linked to morphology in the axial skeleton in the dorsal region and *HoxC13* expression corresponds to the transition into the caudal region.
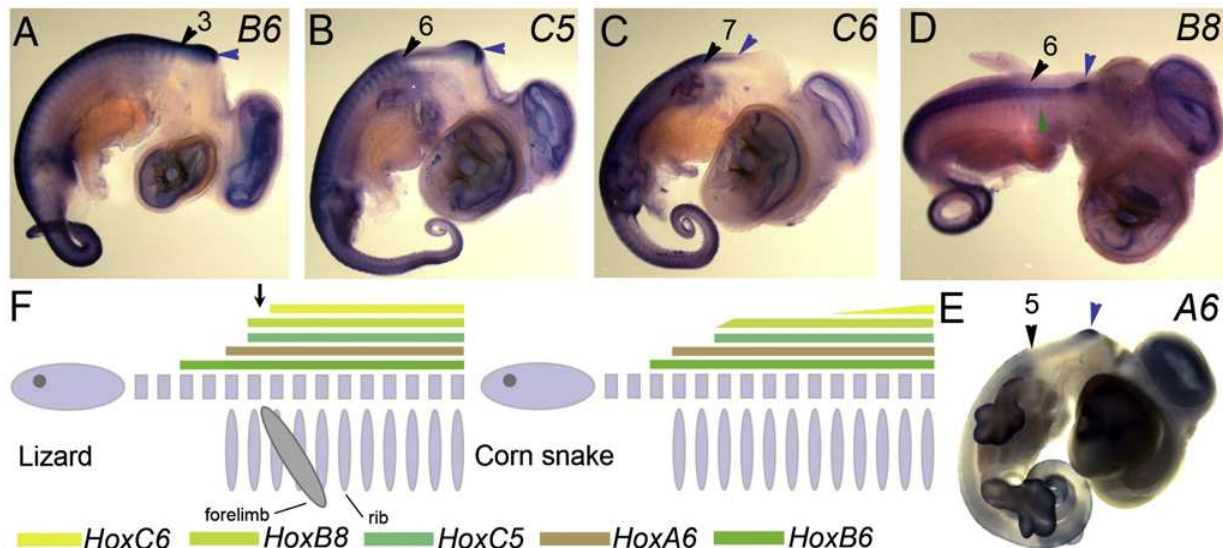


**Figure 14 |** *Hox* gene expression in lizard embryos and comparison of snake and lizard cervico-dorsal transition *Hox* patterning. (A-E) Expression of *HoxC6*, *HoxB6* and *HoxB8* in embryos of bearded dragon lizard and *HoxA6* in the green anole lizard. Expression in the somitic mesoderm is indicated with black arrowheads and numbers referring to the pre-vertebra levels. Anterior boundaries of neural expression are indicated with a blue arrowhead. *Hox* genes are indicated in the upper right corner of each image. In the panel for *HoxB8* (D) an example of spinal ganglion expression is indicated with a green arrowhead. (f) Schematic visualization of *Hox* expression patterns in the cervico-dorsal transition in lizard and corn snake. The arrow indicates the position of the transition between cervical and dorsal region in the lizard. Despite the absence of most of the morphological skeletal

features corresponding to the cervico-dorsal transition, a regionalized pattern of *Hox* expression is preserved. Note that *HoxA6* expression corresponds to the start of the rib-bearing region in both snakes and lizards.
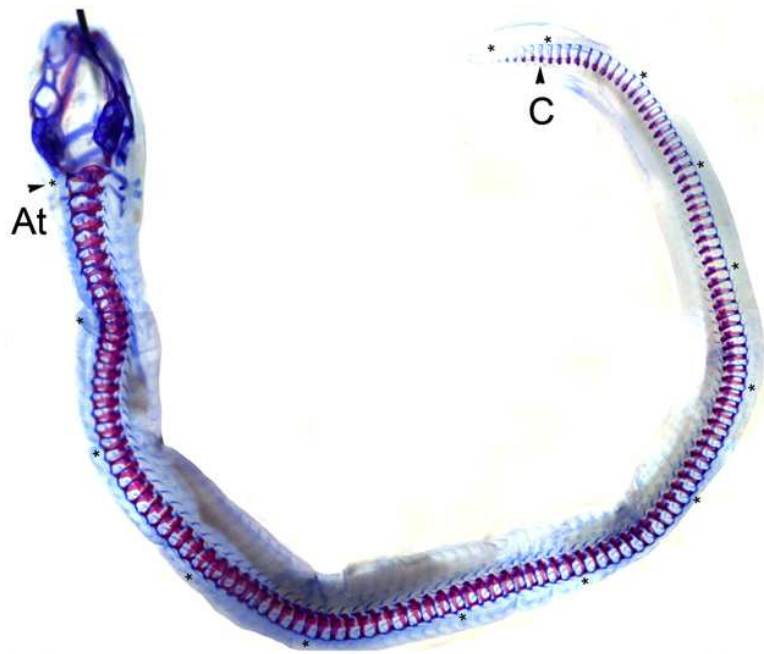


**Figure 15** | Alcian blue - Alazarin red staining of an advanced caecilian embryo to show the skeletal anatomy. The axial skeleton consists of 126 vertebrae, the atlas (At), 121 cervical / dorsal vertebrae and 5 caudal vertebrae. Each tenth vertebra is marked with an asterisk. C; cloaca.

**Table 2 |** Caecilian axial formula and anterior somitic *Hox* expression limits.

| Caecilian (*Ichtyophis* cf. *kohtaoensis*) axial formula | | | |
| --- | --- | --- | --- |
| Vertebra type | Position | Number | Rib |
| Atlas | 1 | 1 | No |
| Dorsal (Thoracic) | 2–121 | 120 | Yes |
| Caudal | 122–126 | 5 | No |
| | | | |
| Caecilian *Hox* expression boundaries | | | |
| Gene | Pre-vertebra level | Anatomical boundary | |
| *HoxC5* | 3 | – | |
| *HoxA6* | 3 | | |
| *HoxC6* | 5 | – | |
| *HoxB8* | 7 | – | |
| *HoxC8* | 21 | – | |
| *HoxC10* | 66 | | |
| *HoxC13* | ~120 | ~Cloaca/caudal vertebrae | |

Indicating here are the vertebral types, their position and number, and whether they are rib-bearing. Anterior somitic *Hox* expression boundaries are given below. Where *Hox* expression boundaries coincide with an anatomical transition this is indicated.



◄ somite ◄ neural ◄ spinal ganglion
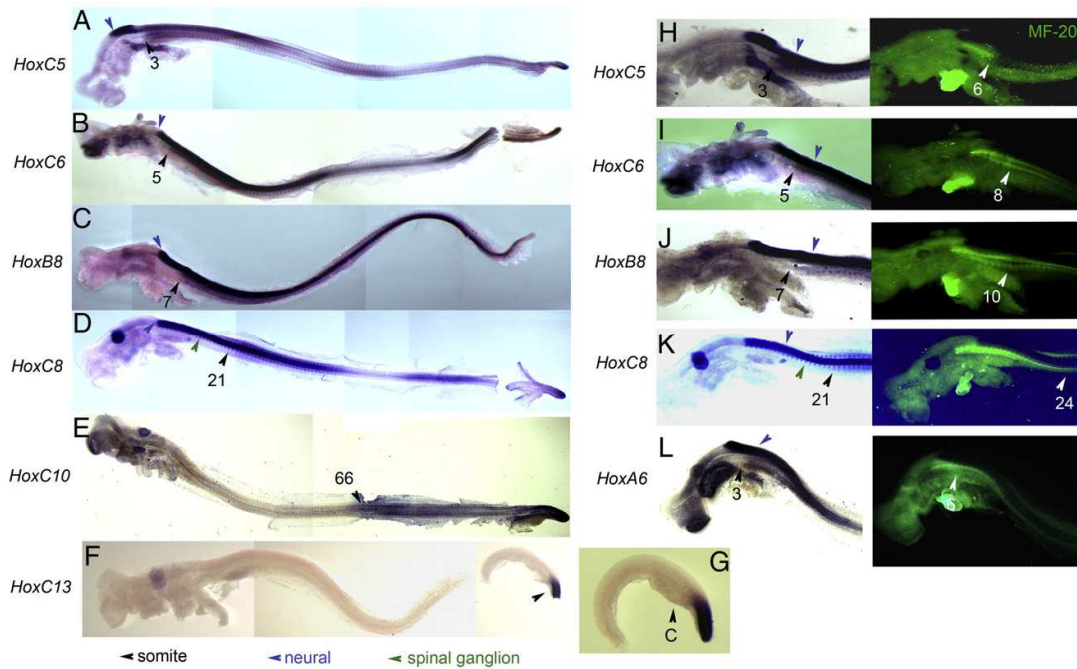
**Figure 16 |** (page 52) *Hox* gene expression in Brauer stage 26 caecilian embryos. (A-E) Shown are flat mounted embryos with the yolk removed (for whole mount unprocessed example see Supplementary Material Fig. S1). *Hox* genes are indicated on the left. Anterior boundaries *Hox* gene expression in the somitic mesoderm is indicated with black arrowheads and pre-vertebrae levels are indicated with numbers. The anterior boundary of neural expression is indicated with a blue arrowhead and in the panel for *HoxC8* (d) an example of spinal ganglia expression is indicated with a green arrowhead. (G) Close up of the tail plus posterior trunk region of *HoxC13* expression, "C" indicates the position of the cloaca. (H-L) Fluorescent immune staining with the somite marker MF-20 antibody together with *Hox* gene in situ hybridization. In the brightfield close ups of head plus anterior trunk, the anterior *Hox* gene expression boundaries in the somites are indicated with black arrowheads and numbers indicating the pre-vertebra level. In the fluorescent microscopy views of the same samples showing the somite specific staining of the MF-20 antibody, the anterior level of *Hox* expression is marked with a white arrowhead and a white number referring to the somite number.

## *Hox* and *Tbx5* expression in the snake lateral plate mesoderm

In addition to the homogenous, extended rib cage, another typical feature of snake-like species is the lack of limbs. *Hox* genes are involved in limb bud induction and it has been shown that the forelimb buds develop at the anterior boundary of *Hox-9* gene expression in the chick lateral plate mesoderm[157]. In the corn snake, we do observe collinear expression of *HoxA*, *HoxB* and *HoxC* genes within the lateral plate somatopleure (Fig. 12, Supplementary Material Fig. S2, 3 Fig. 16 A, D; lateral plate expression of all "thoracic genes" investigated, except *HoxB4, HoxB6* and *HoxA3*, was verified by cryosectioning: Supplementary Material Fig. S2, 3 and data not shown). In general the axial level of *Hox* gene expression coincides or is located very near the axial level of the somitic boundary. For *HoxC6* and *HoxC8,* which have very gradual anterior limits in the somites, we observe similar gradual anterior limits in the lateral plate (Supplementary Material Fig. S3 and data not shown).

*HoxB9* is expressed in the lateral plate mesoderm up to the same axial level as in the somitic mesoderm (Fig. 17 A, D). This position could thus correspond to a vestigial forelimb domain. We investigated the expression of the forelimb specific marker *Tbx5* [158,159] in the corn snake to see if any such domain could be detected at the molecular level. Unexpectedly however, *Tbx5* is expressed throughout the entire antero-posterior extent of the somatopleure (Fig. 17 B, E and Supplementary Material Fig. S6).

Thus, despite regionalized collinear *Hox* gene expression in the lateral plate mesoderm, no forelimb buds are induced in the snake and the forelimb gene *Tbx5* shows a very strong deregionalized expression that seems independent from the *Hox* expression pattern in the lateral plate (Fig. 17 F).
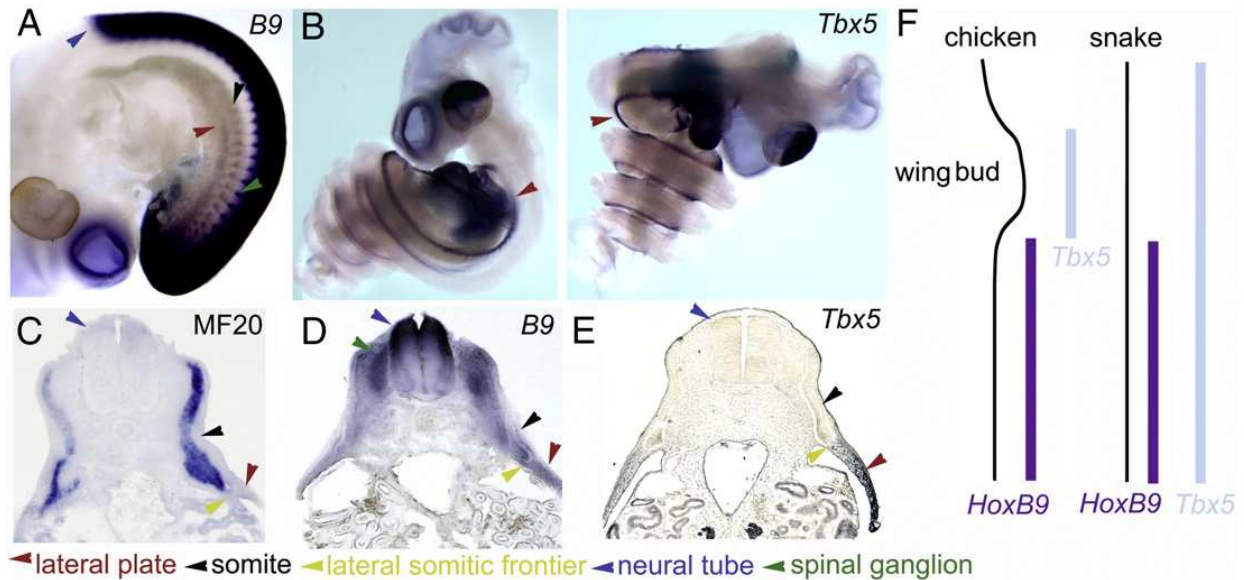


**Figure 17 |** Expression of *HoxB9* and *Tbx5* in the snake lateral plate mesoderm. All tissue shown is from embryos at the second day after oviposition. (A) *HoxB9* is expressed in the snake with an anterior limit in the lateral plate (red arrowhead) and somitic mesoderm (black arrowhead) around pre-vertebra 17. The anterior limit of neural tube expression is marked with a blue arrowhead and an example of spinal ganglion expression is marked with a green arrowhead. (B) The forelimb marker *Tbx5* is expressed throughout the antero-posterior extent of the pre-cloacal lateral plate somatopleure in the corn snake. Expression starts more anterior than *HoxB9* at the head-trunk transition (marked with a red asterisk). The general lateral plate expression domain is indicated with red arrowheads. (c) To determine the position of the lateral somitic frontier in the transversal sections immunostaining with MF20 antibody specific for somitic mesoderm was performed. The cryosection shown was taken in the anterior trunk. The lateral somitic frontier is indicated with a yellow arrowhead. (D) *HoxB9* expression in transversal cryosection taken around pre-vertebra 40 shows expression in lateral plate and somitic mesoderm. (E) *Tbx5* expression in transversal wax section taken at the mid-axial level shows *Tbx5* expression in the lateral plate mesoderm. (F) Schematic comparison of *HoxB9* and *Tbx5* expression in the lateral plate. In the chicken *HoxB9* and *Tbx5* are expressed mutually exclusive at these stages of development with *Tbx5* being anterior and *HoxB9* posterior. In the snake, *Tbx5* expression starts more anterior than *HoxB9* but overlaps posteriorly throughout the pre-cloacal region of the trunk with *HoxB9* expression. Thus, despite the "normal" collinear pattern of *HoxB9* expression, *Tbx5* is not regionalized to a forelimb domain but expressed more posterior than expected. Lateral plate, somitic, spinal ganglion and neural tube staining as well as the position of the lateral somitic frontier are indicated with

different colour arrowheads. Additional snake *HoxB9* expression data is provided in the Supplementary Material (Fig. S5) and additional chicken and snake *Tbx5* data in Fig.S8.

## Discussion

Snakes put their long bodies to good use when moving through dense vegetation, along tree branches or in water, which they do with incredible grace. Constricting snakes use their coils to suffocate prey, whereas venomous snakes are like a coiled spring that can strike prey at a distance and with terrifying speed. Snakes are among the most successful of vertebrate groups, both in terms of number of species and geographic distribution[160]. Elongated bodies have certainly contributed to this success. Model species (such as mice) will always be the mainstay of research because scientists can engineer changes in their genomes and look for effects on the body plan under controlled conditions[161]. Biologists are nevertheless increasingly considering non-model organisms. In snakes, natural selection has done the genetic engineering for us, and so we can study them as experiments performed by evolution. Investigating snakes hand-in-hand with model species will provide a holistic view of evolution and development. Snake embryos are not easy to work with in the lab because it is difficult to open their eggs without damaging the embryo, and obtaining eggs from snakes is not always easy. But if these technical challenges are overcome, and snake genome sequences become widely available, a new era of 'evo-devo' research may open up.

We have described a collinear pattern of *Hox* expression along the primary axis in snake and caecilian embryos. At the anterior and posterior regionalized parts of the axis we find that *Hox* gene expression coincides with expressed anatomical transition. Genes, such as *HoxA3*, *HoxB4*, *HoxA6*, and *HoxC13* have anterior expression boundaries corresponding to clear morphological transitions in the axial skeleton (atlas, axis, rib-bearing region and dorsal-caudal transition).

We however also detect collinear *Hox* expression boundaries within the dorsal homogenous rib-bearing part of the skeleton that are not correlated with transitions in the morphology of the axial skeleton. Genes such as *HoxC6*, *HoxA7*, *HoxC8*, *HoxB9* and *HoxC10*, that normally act to modify the

55

axial skeleton, display regionalized expression in "silent" domains (i.e. not reflected in anatomy), suggesting that they have lost their ancestral effects on vertebral regionalisation. The diffuse anterior expression boundaries of these genes in the snake also may reflect loss of involvement in the regionalization of the axial skeleton.

Although snakes lack a neck region, as defined in squamates by the positioning of the forelimb girdle, the comparison *Hox* gene expression in lizards indicates that a relative normal neck region is patterned. Interestingly, *HoxA6* expression coincides with the beginning of the rib-bearing region in both lizards and snake, and our observations do not rule out the possibility that a single *Hox* gene is responsible for rib formation. Apparently, though, the mechanism of posterior prevalence, according to which more posterior *Hox* genes modify these ribs or suppress rib formation completely, is not operational.

The snake lateral plate mesoderm shows a collinear pattern of *Hox* expression as has been implicated in the positioning of the limbs in chicken. No limb buds are present in corn snakes however, and neither is any molecular trace of a discrete forelimb. In contrast to the potential upstream regionalizing signal provided by differential *Hox* expression (especially *HoxB9*), the expression of *Tbx5* is highly deregionalized.

In summary, our data suggest that the axial expression patterns of *Hox* genes in the somitic and lateral plate mesoderm of the snake and caecilian do not induce the same downstream responses as they do in the mouse. And so, while some ancestral features of the '*Hox* code' are retained in the two long-bodied lineages studied here, they result in a remarkably different phenotype from that seen in the mouse.

We provide evidence that, besides changes in expression domains through alterations in *cis*-regulation, alterations in downstream gene interpretation may also play an important role in the evolution of a snake-like body plan. These changes in downstream response to the *Hox* pattern could, in principle, occur at two levels. First, the response of downstream genes could be altered through changes in *cis*-

56

regulatory sequences leading to a different response to *Hox* genes. Alternatively, changes in *Hox* coding sequences could lead to the differential activation of target genes. The contribution of coding sequence evolution to evolutionary modifications has received ample attention recently [162,163] and it has been shown that for instance changes in *Ubx* coding sequence were essential for changes in invertebrate bodyplans[164,165]. The possibility exists that in snakes and caecilians a similar mechanism is involved in changes in their body plan. Transgenic gain-of-function assays in mouse should reveal whether snake and caecilian *Hox* genes do indeed have different inducing properties from those of their mouse orthologs.

## Material & Methods

Corn snake and bearded dragon embryos were obtained from private breeders in the Netherlands and caecilian embryos were collected from the wild in Thailand, in accordance with local and international regulations. Alizarin red and alcian blue staining was carried out according to standard protocols, and embryos were cleared in methyl salicylate.

Genes were cloned by PCR from *Patherophis spiloides* cDNA or *Pantherophis guttatus* genomic DNA and caecilian genomic DNA and are deposited in genebank under accession numbers GQ176238-GQ176263.

Immunostaining and histology was carried out according to standard procedures. The MF-20 antibody was obtained from the Developmental Studies Hybridoma Bank, University of Iowa. *In situ* hybridization was carried out according to standard protocols. Pre-vertebral axial formulae were determined by somite counting and, if visible in the *in situ* hybridization wholemounts, the position of the first spinal ganglion. In corn snakes and bearded dragons, we used the amniote standard formula where the $1^{st}$ pre-vertebra forms from somite 5-6. In caecilian we determined (by Alcian blue staining and the position of the first spinal ganglion) that the $1^{st}$ vertebra forms from somites 3-4, which is the same as in *Xenopus laevis*.

# Chapter 5: Massive Evolutionary Expansion of Venom Genes in the King Cobra Genome

Vonk FJ, Henkel CV, Casewell NV, Kini RM, Kerkkamp HM, Wuster W, Castoe TA, Ribeiro JMC, Spaink HP, Jansen HJ, Hyder SA, Arntzen JW, Pollock DD, van den Thillart GEEJ, Boetzer M, Pirovano W, Dirks RP and MK Richardson.

Snake venom is a complex mixture of proteins and peptides evolved to immobilize prey and deter predators[166]. The rapid evolution of venom toxins is part of a predator-prey 'arms race' that represents a classic model for studying molecular evolution[13,25,27,167,168]. Snake toxins are thought to evolve from normal physiological proteins through gene duplication and recruitment to the venom gland[169-172]. However, in the absence of genomic resources, these hypotheses remain speculative. Using Illumina sequencing technology we have produced a draft genome of an adult male Indonesian king cobra (*Ophiophagus hannah*), and have deep-sequenced its venom gland transcriptome. Comparative genomics revealed evidence of tandem duplication of genes encoding physiological L-amino acid oxidase, cysteine-rich secretory proteins and metalloproteinases, followed by recruitment through selective expression in the venom gland. By contrast, nerve growth factor toxins appear to have evolved by duplication and dual recruitment, while hyaluronidase and phospholipase B evolved by recruitment of existing physiological genes without further duplication, similar to acetylcholinesterase[173]. We also identify 21 different three-finger toxin (3FTX) genes in the genome, suggesting a massive expansion of this family. We find a significant variation in the expression levels of these 3FTX genes in the venom. These data show that venom proteins originate and evolve through multiple distinct mechanisms. These sequences provide a valuable resource for studying rapid evolution of gene sequences and the evolution of recruitment of genes to different tissues.

Many advanced snakes (Caenophidia[174]) use venom, with or without constriction, to subdue their prey. Venom is produced in a post-orbital venom gland that probably evolved from an ancestral gland in the posterior part of the mouth[2]. There is evidence that the evolutionary history of venom in reptiles may be traced as far back as Triassic lizards[175]. The pharmacologically-active proteins and peptides in venom target a wide variety of proteins including receptors, ion channels and enzymes[176]. Their actions include disruption of the central and peripheral nervous systems, the blood coagulation cascade, the cardiovascular and neuromuscular systems, and homeostasis. Only in recent years has the remarkable variability of venom composition at the taxonomic, population, and individual levels been fully appreciated, and a start made to investigate the underlying ecological drivers and molecular mechanisms[176].

It is thought that toxin gene families have evolved through duplication of normal physiological genes[170-172,177], followed by recruitment and expression in the venom gland. However, this hypothesis has not been verified with genomic data. Therefore, we have produced a draft genome of an adult male Indonesian king cobra (*Ophiophagus hannah*) and deep-sequenced the transcriptome of its venom gland using Illumina technology. The sequence data were first assembled *de novo* into contigs, which were subsequently oriented and merged in scaffolds (see Methods summary). Haploid genome size was estimated using flow cytometry to be around 1.36-1.59 Gbp. Our assembled draft has an N50 contig size of 6533 bp, and an N50 scaffold size of 242 Kbp. The contigs sum to 1.34 Gbp, and the scaffolds (which contain gaps) to 1.59 Gbp.

Mitochondrial genome phylogeny confirms that the male specimen we used for genome sequencing clusters in the *Ophiophagus* group with other king cobras (Fig. 18). Using Augustus gene prediction[178], and our transcriptome data we estimate that the king cobra has approximately 22,183 protein-coding genes (data not shown). Although some of the predicted genes will be either part of a gene that spans multiple scaffolds, or will represent mispredictions, the values suggest that the total number of genes in snakes and other amniotes is similar[179-181].
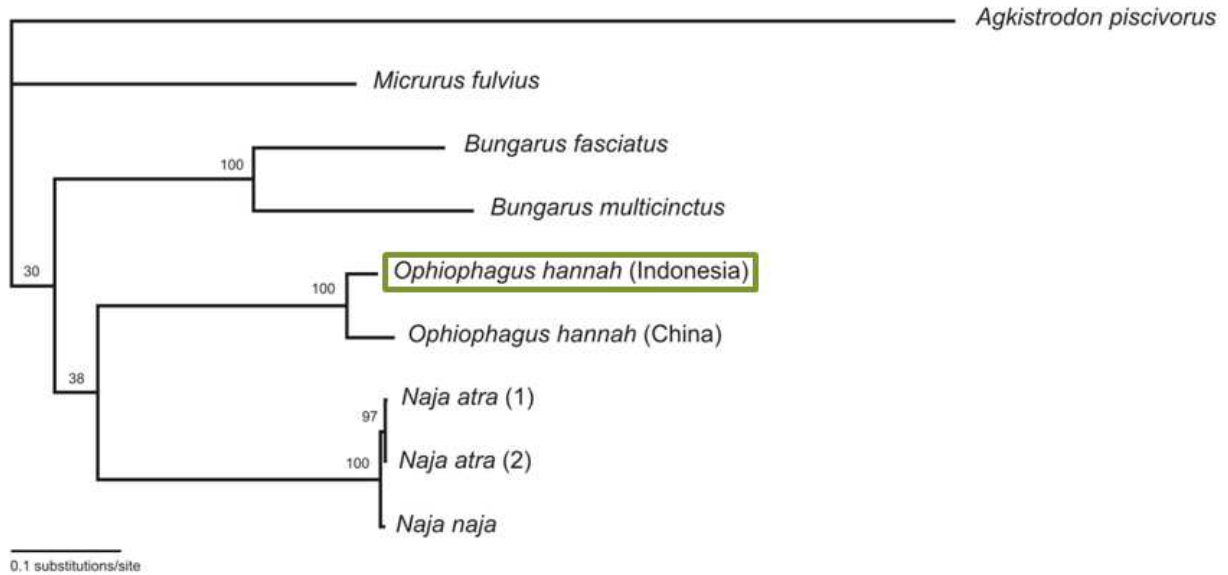
**Figure 18 |** Mitochondrial genome phylogeny showing that our male Indonesian king cobra (*Ophiophgus hannah*) groups with the king cobra from China (Yi Chuan. 2010 Jul;32(7):719-25). Note that the *Naja naja* is in fact a taxonomy error because the sample was taken from a *Naja atra* (Wolfgang Wuster, pers comm).

We identified 17 different toxin families in the venom gland transcriptome by blasting against reference sequences (from www.ncbi.nlm.nih.gov, and see Fig. 19) and annotated nine of them in the genome (Figs. 20-25). These nine include: three-finger toxins (3FTXs), L-amino acid oxidase (LAAO), phospholipase $A_2$ (PLA$_2$), phospholipase-B (PLB), cysteine-rich secretory protein (CRISP), metalloproteinases (ADAM), nerve growth factor (NGF), hyaluronidase (HYA), cobra venom factor (CVF). Three of these (NGF, PLB and CVF) have not previously been reported in king cobra venom.

**Figure 19 |** (page 64) Relative abundance of the venom toxins in the venom gland transcriptome. The percentages are calculated based on the expression value of the transcripts sequenced from the venom gland transcriptome. The most abundant family is the three finger toxins (31,54% of all toxin transcripts identified), represented in the genome by at least 21 different isoforms (see also Fig. 20).

ADAM; 5.42%

CRISP; 8.27%

CVF; 1.58%

Hya; 0.36%

Kunitz; 3.02%

LAAO; 6.51%

Lectin; 0.02%

NGF; 1.63%

NP; 4.70%

Peptidase; 5.53%

PLA2; 11.73%

PLB; 0.45%

Various; 0.19%

VEGF; 4.55%

Vespryn; 0.21%

Waprin; 0.26%

AchE; 1.79%

3Ftx; 43.78%

VENOM 19%

Unknown 9%

Housekeeping 72%
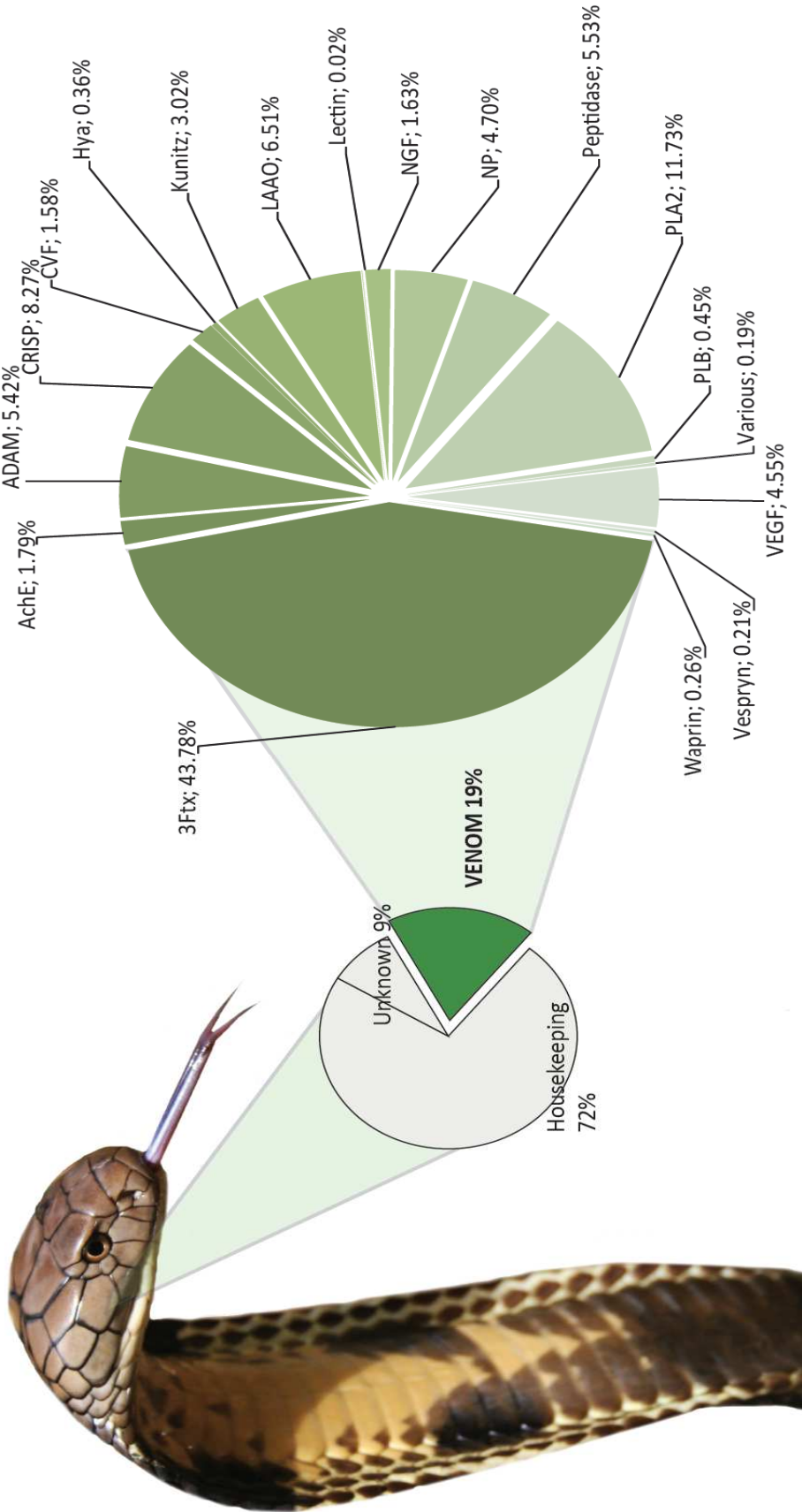
Proteins in two of these families (3FTX and PLA2), are known to exhibit a wide variety of toxic and pharmacological effects including neurotoxicity, cardiotoxicity and hemolysis[182,183]. We find evidence for massive expansion in the genome in both these families. We found seven different exons-2 that belong to PLA2 (Fig. 20). These genomic sequences do not contain premature stop codons or frameshifts indicating that they do not contain pseudogenes.

```
                  20                    40                    60                    80
                  |                     |                     |                     |
PLa-2 (O. hannah) MNPAHLLVLS AVCVSLLGAS SIPPQPLNLL QFNYMIQCTI PGSRPFLDYM DYGCYCGTGV AGHPVDELDR CCQTHDLCYS 80
   Gn_PLa-2_hit1  MNPAHLLVLS ---------- ---------- ---------- ---------- ---------- ---------- ---------- 10
   Gn_PLa-2_hit2  MNPAHLLVLS T--------- ---------- ---------- ---------- ---------- ---------- ---------- 11
   Gn_PLa-2_hit3  MNPAHLLVLS A--------- ---------- ---------- ---------- ---------- ---------- ---------- 11
   Gn_PLa-2_hit4  ---------- ---XXXLGAS SIPPQPLNLL QFNYMIQCTI PGSRPFLDYM DYGCYCGTGG RGTPVDELD- ---------- 56
   Gn_PLa-2_hit5  MNSAHLLVPA VVCVFLLGAS SIPPQSLNLY QFKNMIRCTI PRSIPWWDYS DYGCYCGAGG SGTAVDKLDR CCQVHDNCYT 80
   Gn_PLa-2_hit6  ---------- -VCVSLLGAS SIPPQPFDLY QFKYMIQCTI PGILSWLKYM NYGCYCGSGG SGTPVDKLD- ---------- 58
   Gn_PLa-2_hit7  ---------- ------LGAS SIPPQPLNLL QFNGMIECTI PGSVPWLDFS NYG------- ---------- ---------- 37
   Gn_PLa-2_hit8  ---------- -VCVSLLGAS SIPPQPLNLL QFNGMIECTI PGSIPWLDFS NYG------- ---------- ---------- 42
   Gn_PLa-2_hit9  ---------- -VCVSLLGAS SIPPQPLHLV QFNGMIRCTI PGSIPWWDYS DYGCYCG--- ---------- ---------- 46
   Gn_PLa-2_hit10 ---------- ---------- -------NLI QFSNMIKCTI PGSRPLLDYA DYGCYCGFGG SGTPVDQLD- ---------- 42
   Gn_PLa-2_hit11 ---------- ---------- ---------- ---------- ---------- ---------- ---------- CCQTHDLCYS 10
   Gn_PLa-2_hit12 ---------- ---------- ---------- ---------- ---------- ---------- ---------- CCQIHDNCYS 10
   Gn_PLa-2_hit13 ---------- ---------- ---------- ---------- ---------- ---------- ---------- CCQTHDLCYT 10
   Gn_PLa-2_hit14 ---------- ---------- ---------- ---------- ---------- ---------- ---------- CCQTHDLCYT 10
   Gn_PLa-2_hit15 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---VHDNCYT 7
   Gn_PLa-2_hit16 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- 
   Gn_PLa-2_hit17 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- 
   Gn_PLa-2_hit18 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- 
   Gn_PLa-2_hit19 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- 
   Gn_PLa-2_hit20 ---------- ---------- ---------- ---------- ---------- ---------- ---------- ---------- 

                  100                   120                   140
                  |                     |                     |
PLa-2 (O. hannah) KAEEQPKCSS LLNSPLMKKY SYTCSGGTLT CNDDNDECGA FICNCDRAAR ICFAGAPYNK ENKELDIATR CQ* 153
   Gn_PLa-2_hit1  ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 10
   Gn_PLa-2_hit2  ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 11
   Gn_PLa-2_hit3  ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 11
   Gn_PLa-2_hit4  ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 56
   Gn_PLa-2_hit5  QAKKISGCS- ----PYLKIY SYDCSGRTVT CK-------- ---------- ---------- ---------- --- 107
   Gn_PLa-2_hit6  ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 58
   Gn_PLa-2_hit7  ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 37
   Gn_PLa-2_hit8  ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 42
   Gn_PLa-2_hit9  ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 46
   Gn_PLa-2_hit10 ---------- ---------- ---------- ---------- ---------- ---------- ---------- --- 42
   Gn_PLa-2_hit11 KAEEQPKCSS LLNSPLMKKY SYTCSGGTLT CN-------- ---------- ---------- ---------- --- 42
   Gn_PLa-2_hit12 QAQQLSACSS IITDSPYIKFY SYDCSEGTL- ---------- ---------- ---------- ---------- --- 39
   Gn_PLa-2_hit13 QANKHPACKS LLD------- ---------- ---------- ---------- ---------- ---------- --- 23
   Gn_PLa-2_hit14 QAKKHPACKS LLD------- ---------- ---------- ---------- ---------- ---------- --- 23
   Gn_PLa-2_hit15 QAQKISGCSS MMETPYLKIY SYKCSERTVT CKDDNDECGA FICNCDRVAA HCFAASPYNN NNYNIDLKAR CQ* 80
   Gn_PLa-2_hit16 ---------- ---------- ---------- --DDNDECGA FICNCDRAAA ICFAGAPYNK ENKELNKSKY CK* 41
   Gn_PLa-2_hit17 ---------- ---------- ---------- ---NCDRAAA ICFAGAPYNK ENKELDITTR CQ* 30
   Gn_PLa-2_hit18 ---------- ---------- ---------- -------GA FICNCDRAAA ICFAASPYNR NNYKIDTTTR C*- 34
   Gn_PLa-2_hit19 ---------- ---------- ---------- --ADNDKCAA FVCNCDRVAA ICFAASPYNW NNYNIDTTTR C*- 40
   Gn_PLa-2_hit20 ---------- ---------- ---------- --------A FVCDCDRVAA ICFAGAPYNK DNINIDTTTR C*- 33
```
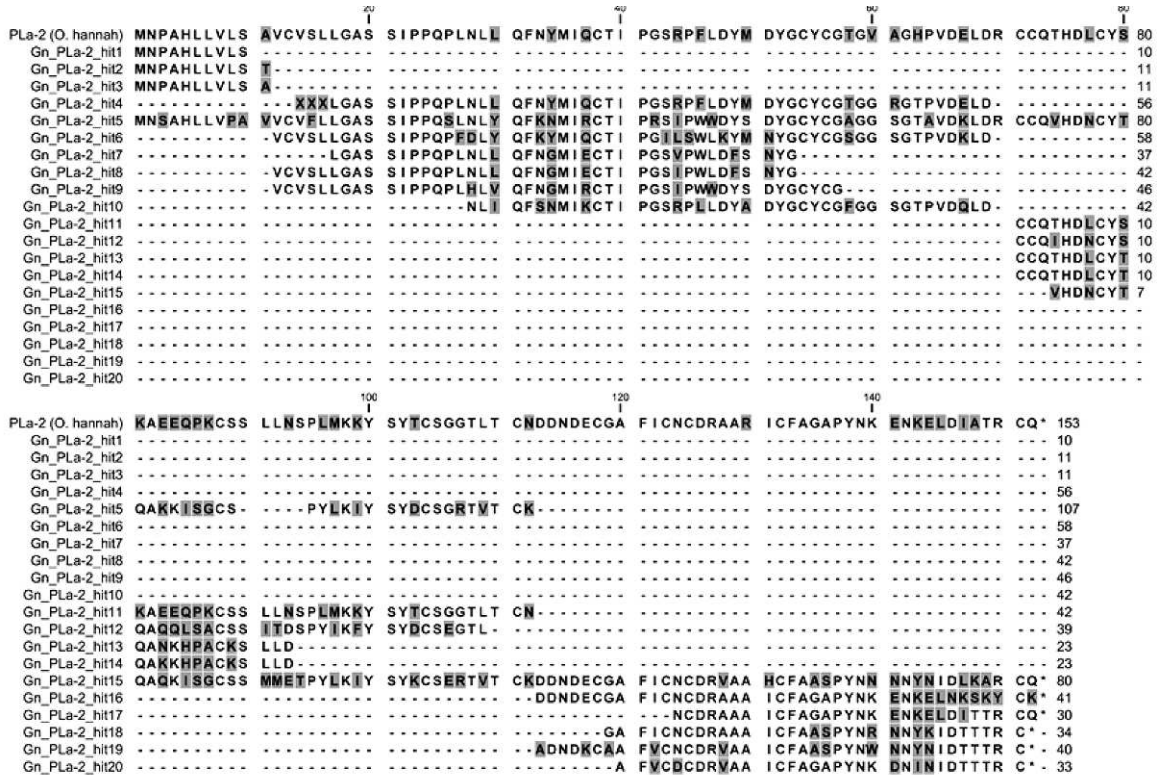
**Figure 20** | Massive expansion of PLA2 genes. Alignment of multiple PLA2 genomic hits. Note no stopcodons or frameshifts in the sequences.

3FTXs are three-exon genes, of which the second exon is most readily identified. We found 21 of these exons-2 in the genome (Fig. 21). However, some of these are on small contigs and covered by relatively many sequencing reads, indicative of high copy numbers. Therefore, the actual diversity of full-length 3FTX genes may be even higher. Most exons-2 are expressed in the venom gland, although the
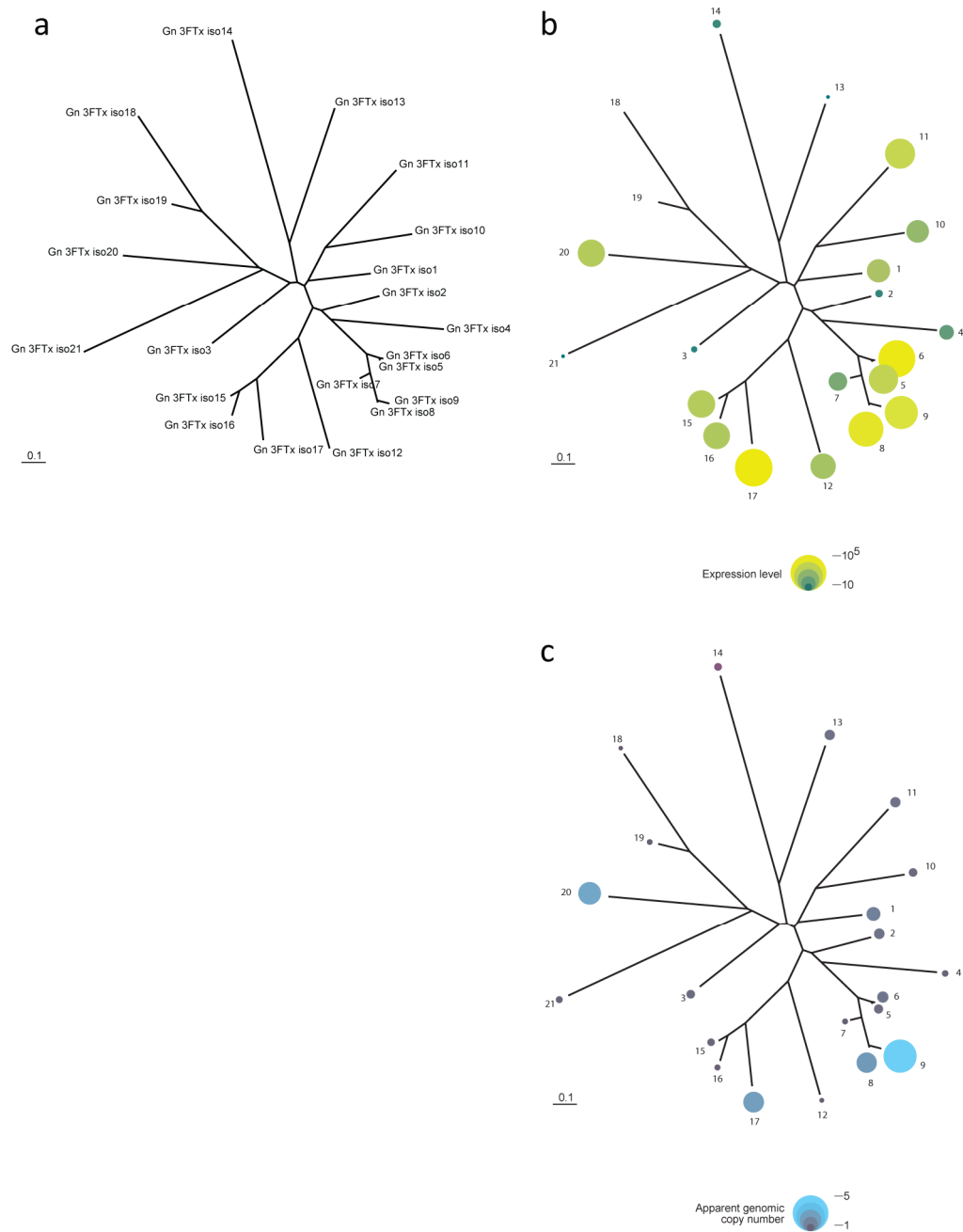
**Figure 21** | Unrooted phylogenetic tree constructed from all different exon-2 sequences of the three-finger toxin genes. Isoform 19 contains a premature stop codon, thus most likely is a pseudogene. Green circles indicate relative expression levels (on a logarithmic scale), blue circles apparent genomic copy numbers, both based on local coverage by venom.

expression levels differ by five orders of magnitude (Fig. 21). One non-expressed isoform (isoform 19) contains a premature stop codon and may be part of a pseudogene (Fig. 22). The presence of multi-copy and highly expressed exons is clustered in several 'successful' branches of the 3FTX gene family, and genomic copy number and expression level in the venom gland appear to be correlated (Fig. 20).



**Figure 22** | Alignment of multiple 3FTX isoforms. Note that only isoform19 (Gn_3FTx iso19) contains a stopcodon and may thus be a pseudogene.

There is a substantial difference in expression levels of each of the 3FTX isoforms (Fig. 20). Isoform diversity and toxin expression levels are thought to be important in optimization of the prey-specificity of the venom — more so than differences in the representation of entire toxin families and the recruitment of novel toxin families[184]. In general, a high genomic copy number is associated with a high relative expression value (Fig. 20). All relatively 'successful' branches of 3FTX genes (the ones that are expressed heavily) share sequence similarities (Fig. 22), indicating conservation of important functions.

Reptile venom CRISPs act as regulators of several types of ion channels[185]. We find three CRISP genes in tandem in the king cobra genome (Fig. 23), only two which are represented in our venom gland transcriptome (Fig. 23). Together with the comparative genomic data (Fig. 23), this is consistent with an evolutionary scenario in which the two venom genes have been derived by tandem duplication from the non-venom expressed (physiological) CRISP gene.
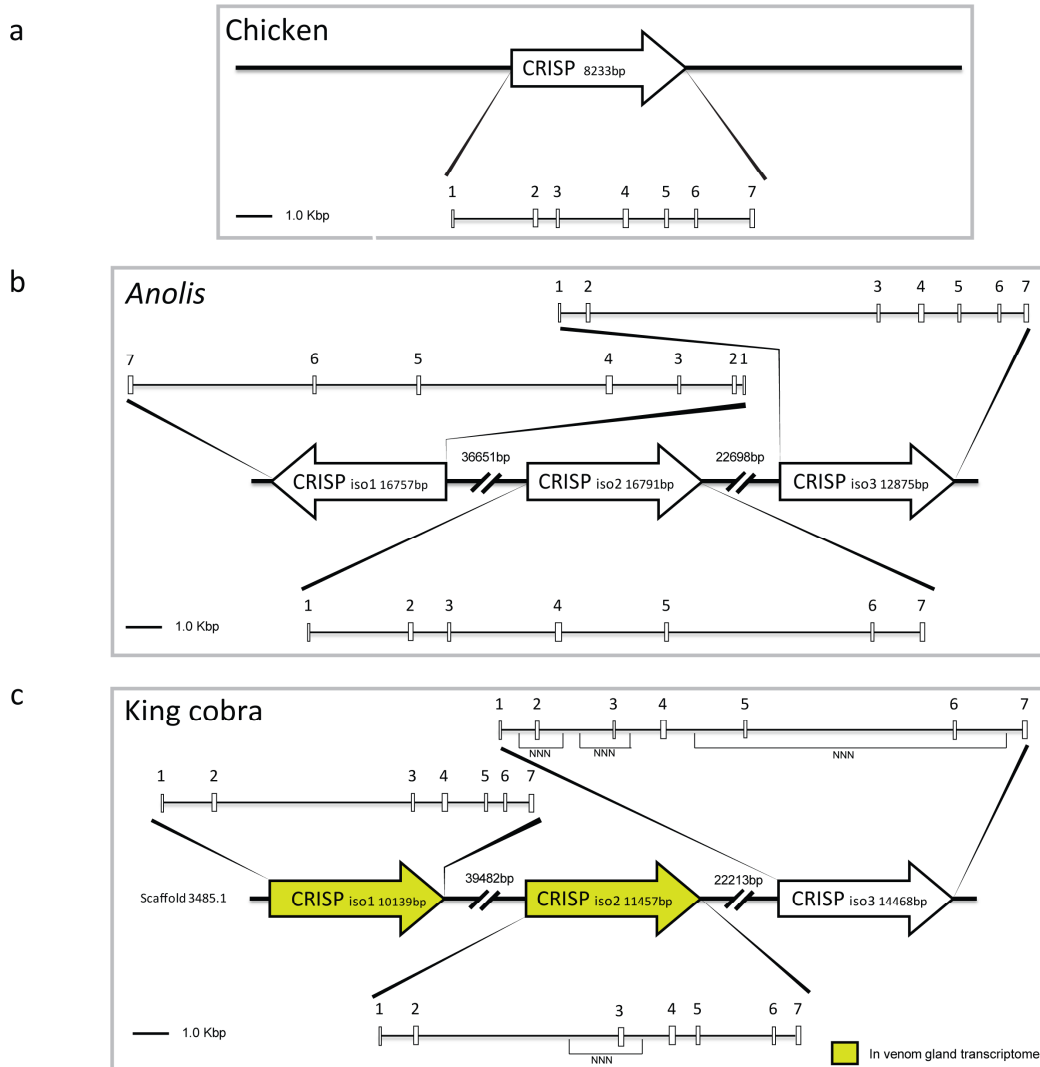


**Figure 23 |** Comparative genomic architecture of the CRISP genes. a, chicken (*Gallus gallus*); b, anole lizard (*Anolis carolinensis*); and c, king cobra (*Ophiophagus hannah*). Chick and Anolis sequences are from www.ensembl.org. The exploded views show scale diagrams of the exons and introns. Scale bar refers to the exploded views. NNN, unresolved sequence. In the

Anolis genome we annotated three CRISP genes with different orientations. Based on the relative sizes of the second introns the two 'venom' CRISP genes are comparable to isoform 3 in Anolis. In chicken we could only find one CRISP gene.
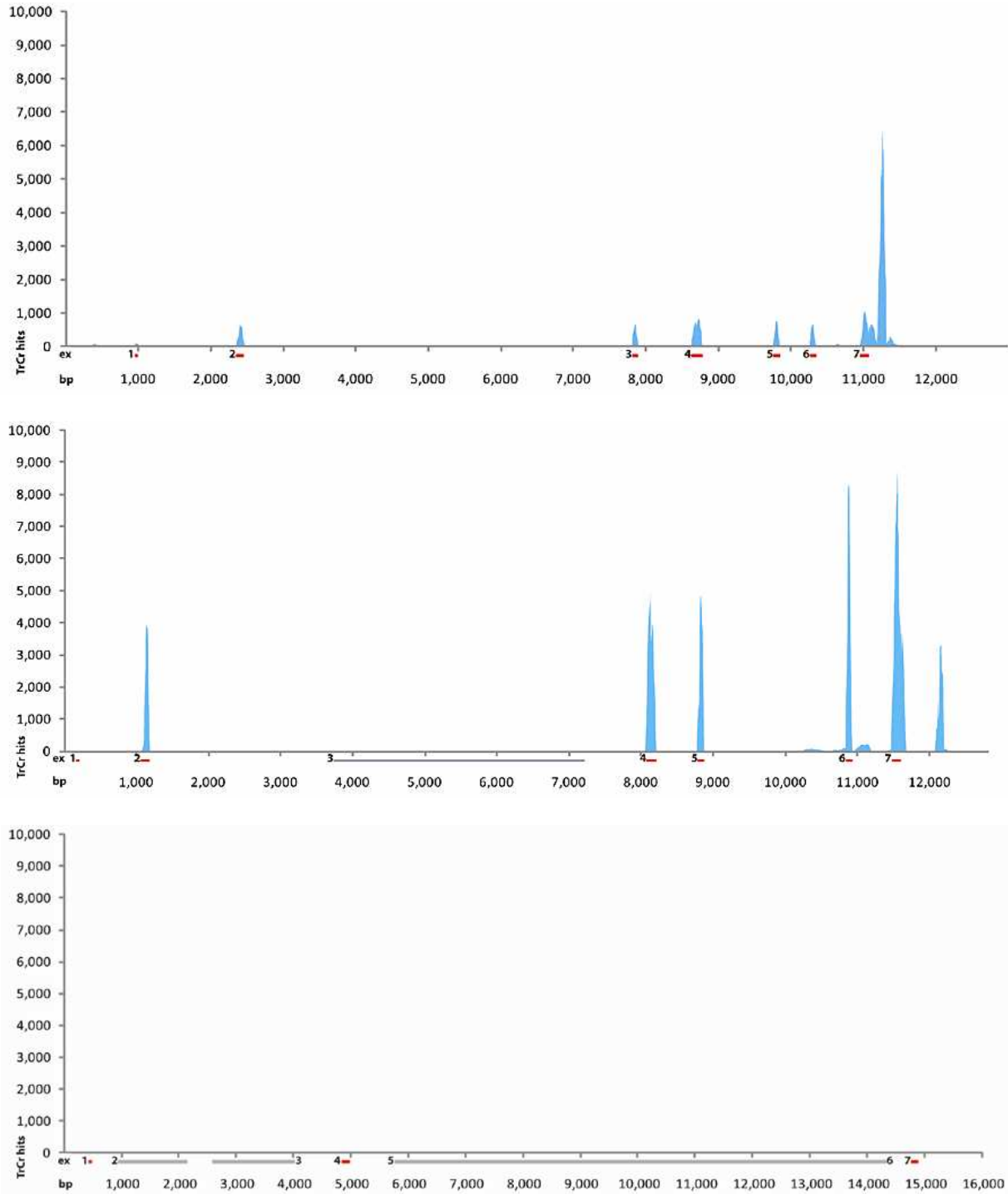


**Figure 24** | The scaffold containing three CRISP genes with different isoform transcripts (see Figure 23c) mapped on as follows: a) isoform 1; b) isoform 2; c) isoform 3. As can be seen, only the first two isoforms are expressed in the venom gland; d)

69

alignment of the three CRISP genes with reference sequences showing that our identified genes belong to the CRISP family. Isoform1 is opharin and isoform 2 is ophanin.

Venom metalloproteinases belong to the ADAM family and target various stages of blood coagulation and platelet aggregation and are responsible for hemorrhage[186]. We also find three ADAM genes in tandem, only one of which was expressed in the venom gland transcriptome (Fig. 25).
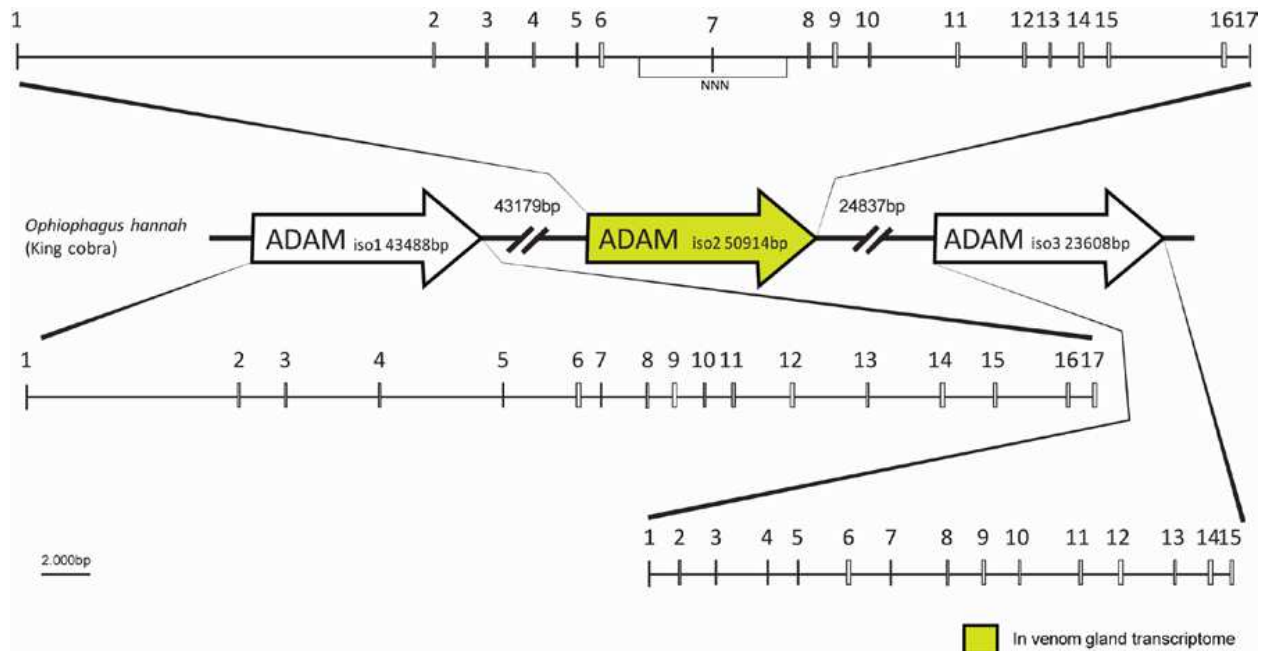


**Figure 25 |** Scaffold containing three ADAM genes; b) isoform 1; c) isoform 2; d) isoform 3. Only isoform 2 is expressed in the venom gland (data not shown). Amino acid alignments of these three metalloproteinase genes with the single transcriptome sequence (not shown) shows that one gene is identical and confirms its expression. Isoform 1 has a longer C-terminal tail. In *O. hannah* isoform 2 is expressed in the venom gland, while in *Naja atra* (a different elapid snake) isoform 3 appears to be expressed, since we find that *N. atra* metalloproteinase sequence is more similar to isoform 3 than isoform 2

LAAO produces $H_2O_2$ during oxidation of amino acids leading to cytotoxicity and inhibition of platelet aggregation (and is responsible for the yellow color of the venoms)[187]. We find two LAAO genes on two different scaffolds (Fig. 26 a). Based on the mapping of venom gland transcriptome reads (not shown), only one LAAO gene appears to be expressed in the venom gland; the other is presumably the

non-venom, physiological gene. To the best of our knowledge, non-venom LAAO proteins have not been found in reptiles before, although they are found widely among vertebrates.
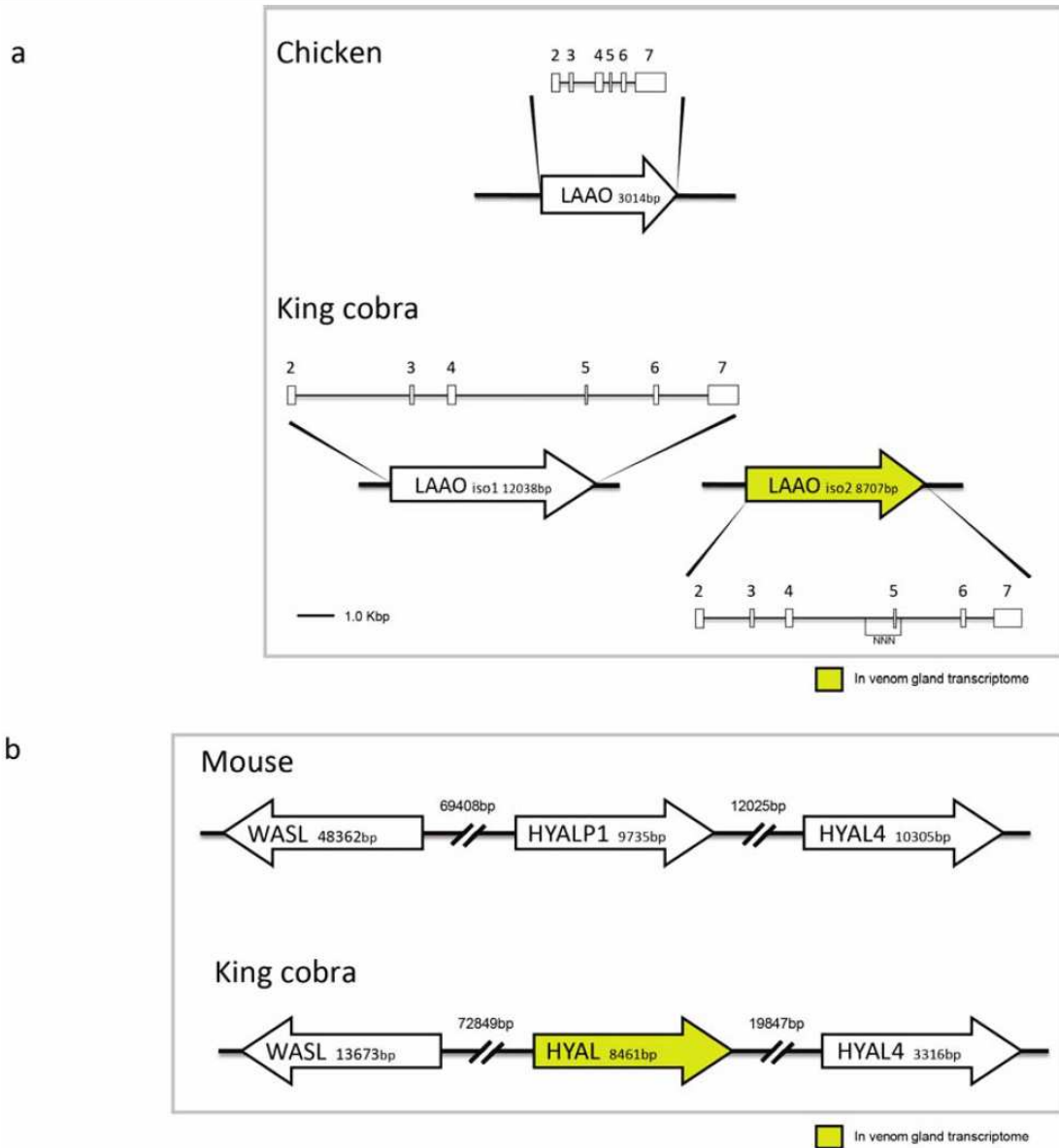


**Figure 26** | a, Genomic architecture of l–amino acid oxidase (LAAO) genes in the chicken and king cobra. b, scheme of the genomic context of the hyaluronidase genes in the mouse (*Mus musculus*) and king cobra. Mouse genomic sequences from www.ensembl.org. Scale bar refers to the exploded views. NNN, unresolved sequence.

The role of venom NGF is not clear[98]. We find two different NGF genes, both of which are encoded by a single exon; and both of them are expressed in the venom gland (Fig. 27). Presumably, one or both of these has duplicate functions (in both venom-gland and in other tissues). Venom hyaluronidase plays a key role as the venom 'spreading factor', making tissue more permeable[188]. We annotated two hyaluronidase genes in the king cobra genome, both lie downstream of the WASL gene, and we find the same arrangement in the mouse genome (Fig. 26 b). Only the gene corresponding to HYALP1 is expressed in the venom gland (data not shown), which is interesting because in the mouse this gene appears to be inactive[189]. This synteny is consistent with a scenario in which the duplication of the hyaluronidase gene took place long before one of the copies was recruited to the venom gland.
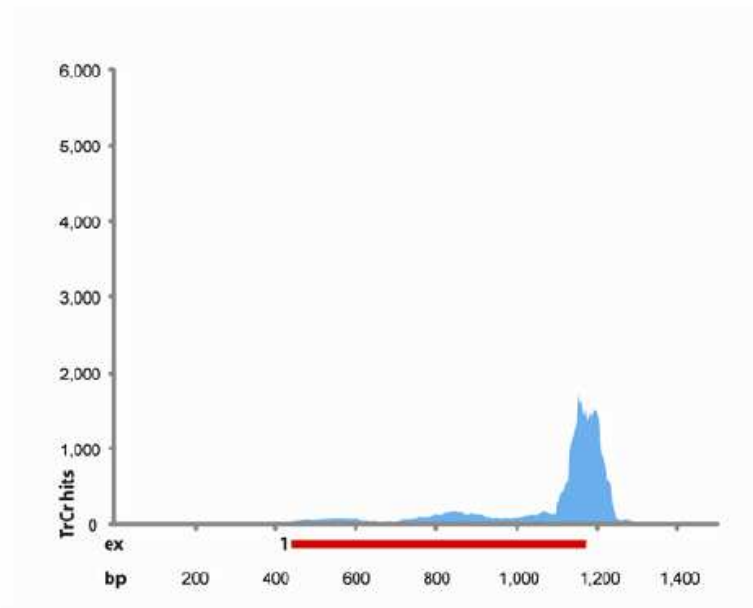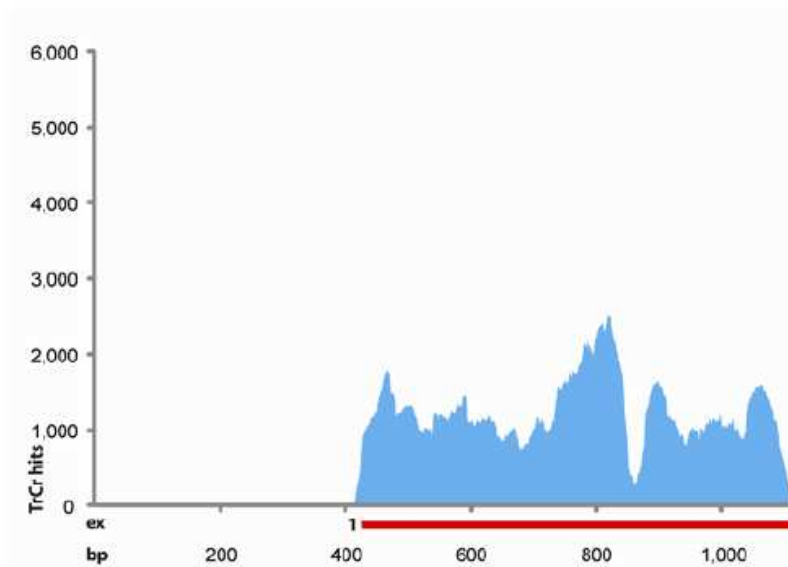
**Figure 27** | Mapping of the transcriptome reads onto the two scaffolds containing two NGF genes shows that both of these genes are expressed in the venom gland; a) isoform 1; b) isoform 2; c) Alignment of the two NGF genes with reference sequences showing that our identified genes belong to the NGF gene family (data not shown).

Recently, PL-B was also found to be expressed in the venom gland[190] but its role in toxicity is yet unclear. We could only find one PL-B gene (Figure 28). This indicates that an existing PL-B gene was recruited to the venom gland. Thus HYA, NGF and PL-B genes appear to be recruited for expression in the venom gland without gene duplication being involved. In the case of the Asian krait (*Bungarus fasciatus*) Acetylcholinesterase toxin, it was shown that both the neuronal and the venom enzymes are encoded by the same gene, although alternatively spliced[173].
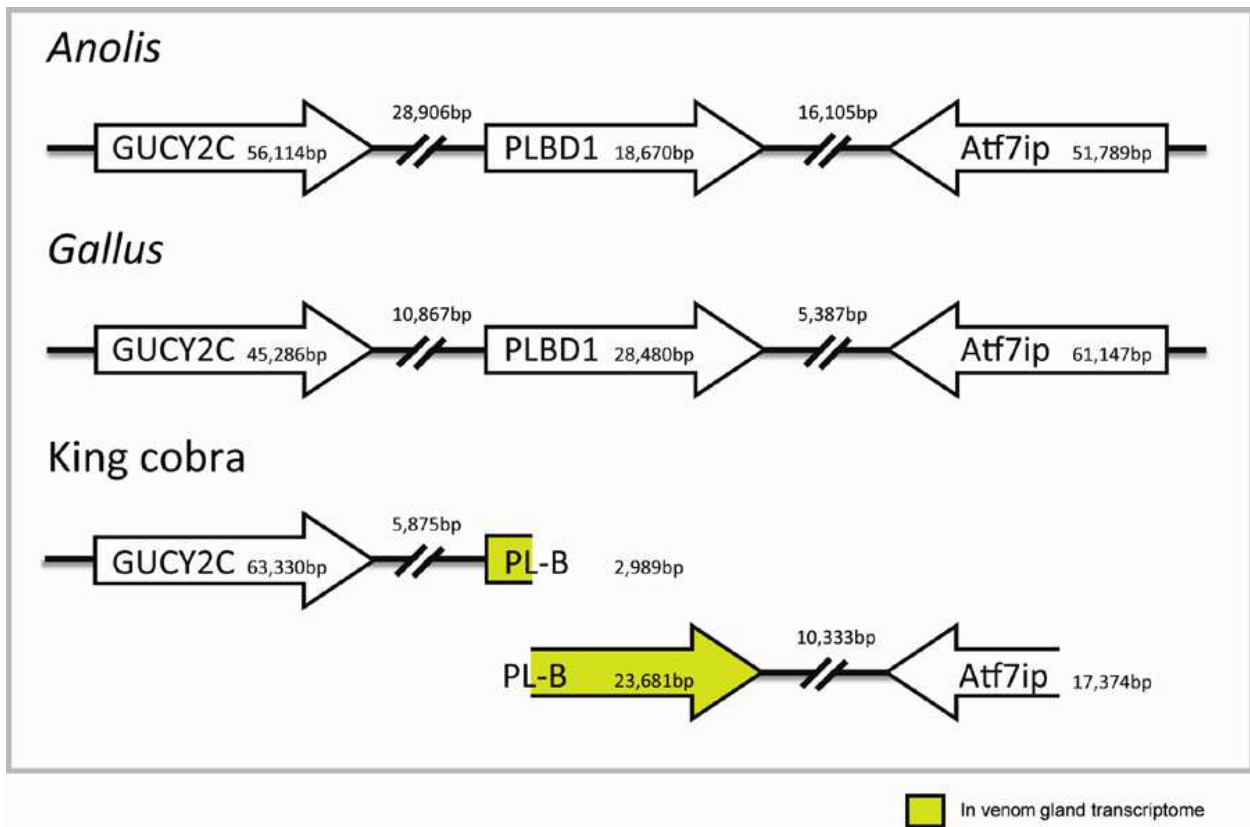


**Figure 28** | Scheme of the genomic synteny of the PL-B genes in the Anolis, Gallus and king cobra. Anolis, Gallus genomic sequences from www.ensembl.org. Mapping of the transcriptome reads onto one scaffolds containing the PL-B gene shows that this gene is expressed in the venom gland (data not shown). This synteny is consistent with the scenario of recruitment of the existing PLBD1 into the venom gland during snake evolution. The alignment of the PL-B gene with reference sequences showing that our identified genes belong to the PL-B gene family (data not shown).

It has been shown, in the case of factor X toxin in the rough-scaled snake (*Tropidechis carinatus*), that a specific insertion in the promoter region of the toxin was responsible for the selective recruitment to the venom gland[169]. We have scanned all our scaffolds for this sequence but could not find anything similar. This suggests that that the specific insertion is not a universal feature of toxin gene recruitment, and that several distinct mechanisms are responsible for the origin and recruitment of venom proteins.

## Conclusion and discussion

Comparative genomics has revealed flexible mechanisms of mutation and recruitment of venom genes. We found evidence of tandem duplication of genes encoding physiological L-amino acid oxidase, cysteine-rich secretory proteins and metalloproteinases, followed by recruitment through selective expression in the venom gland. By contrast, nerve growth factor toxins appear to have evolved by duplication and dual recruitment, while hyaluronidase and phospholipase B evolved by recruitment of existing physiological genes without further duplication. We also identify 21 different three-finger toxin (3FTX) genes in the genome, suggesting a massive expansion of this family. We find a significant variation in the expression levels of these 3FTX genes in the venom. Our data therefore shows that venom proteins originate and evolve through multiple distinct mechanisms.

The king cobra genome is an important resource for studying molecular evolution. The powerful combination of genomics with transcriptomics here used lead to the identification of toxin genes previously unknown in the king cobra. Because of the massive functional diversity known to exist not only between toxins but also their isoforms[176], functional studies of these new toxins could prove to be of great interest.

# Methods summary

**king cobra tissue acquisition and processing.** All animal procedure complied with local ethical approval. Genome sequencing was done on a blood sample obtained from an adult male king Cobra from a captive specimen that originated from Bali, Indonesia. Blood was obtained by caudal puncture and frozen in liquid nitrogen. The venom gland and other tissue samples were dissected from a freshly euthanized second adult male specimen and stored in RNAlater.

**Sequencing and assembly.** We used a whole-genome shotgun sequencing strategy and Illumina Genome Analyser sequencing technology. Two paired-end and four mate pair libraries were constructed with insert sizes of up to 15K nucleotides. In total, we generated 41.2 Gbp (approximately 28x genome coverage) of sequence data for contig building, 21.1 Gbp for scaffolding, and 1.7 Gbp for the transcriptome. We built contigs from the short reads using the CLC bio *de novo* assembler (CLC bio, Aarhus, Denmark) and oriented these contigs using SSPACE. A more extensive methods section is included in the supplementary information.

**Annotation and gene prediction.** Gene prediction was carried out automatically using Augustus software[178], using venom gland transcripts as hints. Further extensive manual annotation was performed to establish the intron-exon boundaries.

**Author contributions** F.J.V., study concept and design, tissue preparation, flow cytometry, sequence analysis, writing of manuscript. C.V.H., genome assembly and analysis, transcriptome analysis, preparation of figures, writing of manuscript. R.M.K., analysis of venom genes, comparative genomics. H.M.IJ.K, sequence analysis, comparative genomics, drawing of figures. H.P.S,, study concept and design, genome and transcriptome sequencing, sequence analysis. H.J.J. genome and transcriptome

# Chapter 6: An Efficient Analytical Platform for On-line Microfluidic Profiling of Neurotoxic Snake Venoms Towards Nicotinic Receptor Like Affinity

Heus F, Vonk FJ, Bruyneel B, Smit AB, Lingeman H, Richardson MK, Niessen W and J Kool.

The highly variable toxic peptides in snake venoms target a myriad of neurotoxic and hemotoxic receptors and enzymes and comprise highly interesting candidates for drug discovery. Venom profiling studies are normally directed either at identification of venom peptides or at biological activity using elaborate and complicated workflows employing relatively large amounts of precious venom. We have developed a microfluidic on-line screening approach utilizing nano-LC–MS with parallel bioaffinity detection, which enables efficient correlation of bioaffinity to identity of the venom peptides. Using the acetylcholine binding protein (AChBP) as biological target, we evaluated our technology with high-affinity α-bungarotoxin and hemotoxic/proteolytic *Vipera ammodytes* venom spiked with α-bungarotoxin. The methodology was applied to profile the venoms of *Dendroaspis jamesoni kaimosae*, *Naja annulifera* and *N. nivea*, which took less than two hours per venom. The data yielded over 20 AChBP ligands of which the corresponding accurate masses were used to retrieve information from literature regarding their function and targeting specificity. We found that from these 20 ligands, 13 were previously reported on, while information on the others could not be found. Our data demonstrated that nano-LC-MS coupled to an on-line biochemical assay is a promising analytical approach for rapid identification of bioactive peptides in snake venoms.

## Introduction

Recent advances in scientific understanding of diseases have raised hopes of an increase in the number of novel drugs reaching the market. Further, many potentially disease-related orphan receptors are now being investigated. It is therefore disappointing to realise that the number of drugs reaching the market is actually dwindling[191]. One of the solutions to this problem is the exploration of natural sources, such as animal venoms, as to increase the numbers of potentially new lead compounds. Why? Possible reasons: venoms are highly potent because they are under extreme strong natural selection[176,192]. Venoms often contain multiple novel compounds, free from IP restrictions. They have also provided many new drugs or even new classes of drugs, in the past.

There is huge variability in snake venom composition[15,184,193] due to an accelerated form of mutation and evolution[25,33,192,194,195] that is currently not understood at the molecular level. Venomous snakes differ in their venom composition not only between species (interspecific) but also within a species (intraspecific). In order to exploit this massive natural resource, several challenges have to be resolved. These include bio-analytical challenges such as: 1) the complex nature of the samples and the intrinsically time-consuming process of elucidation, 2) the occurrence of synergistic actions between components in a complex natural extract, and 3) sourcing adequate quantities. For these and other reasons, pharmaceutical companies scaled-down or even abandoned their natural extract drug discovery pipelines. Nevertheless, many drugs approved during the last decades are based on natural products [196]. Another potential reason for abandoning natural extract product pipelines is that most of them focused on plant and fungal derived bioactives[197], because these mostly small molecule compounds fitted well in their small molecule drug discovery pipelines. Traditional problems with peptide and protein derived drugs are amongst others due to issues with (oral) drug administration, bodily distribution (e.g. limited crossing of the blood-brain barrier) and immunogenicity issues. Advanced formulation methodologies, administration routes[198-201], protein engineering and production techniques[202-204,204], as well as current knowledge and available technologies dealing with immunogenicity and immunotoxicity[205-209] have now brought these biopharmaceuticals back into consideration as candidate drugs.

Great advances have been made in protein-based pharmaceuticals during the last ten years[210-213]. In this regard, venom peptides also gained interest. A couple successful examples include: (i) Prialt (Ziconotide), a non-opioid painkiller derived from the neurotoxic venom of a predatory cone snail (*Conus Textilus*) and used to treat severe chronic pain by injection directly into the spinal fluid [214]; (ii) Byetta (Exenatide), a drug against diabetes TypeII derived from the venom of the Gila monster (*Heloderma suspectum*)[215]; ACE inhibitors used for the treatment of hypertension and derived from a compound in the venom of the South American Lancehead snake (*Bothrops jararaca*)[76]; and the antiplatelet drug Integrillin (Eptifibatide) derived from a compound the venom of the Saw-scaled viper (*Echis carinatus*) and used to reduce the risk of acute cardiac ischemic events in patients with unstable angina. In addition to the drugs that have been derived from animal venoms, there have also been some successful applications in diagnostics like the detection of lupus anticoagulants (abnormally high risk of blood clotting ) using Taipan venom (*Oxyuranus scutulatus*)[216]. Some snake toxins have also been used to probe receptors. The deciphering of the molecular structure of neurons was made possible by the use of bungarotoxin from the venom of the banded krait (*Bungarus multicinctus*) and led to the discovery of a receptor for acetylcholine[217]: the acetylcholine receptor.

One of the specific difficulties of effective natural extract screening is the accurate and rapid bioactivity/identity correlation of bioactives with current traditional effect directed analysis (EDA) methodologies[218]. EDA approaches are today's standard for bioactive mixture analysis and are widely applied in environmental[219-221] and natural product screening[222-224], as well as animal venom profiling[225]. In this approach, bioactive compounds are separated and subsequently fractionated, e.g., with liquid chromatography (LC), into a number of fractions, after which each fraction is tested for a specific bioactivity. Ultimately, the bioactives are analysed with mass spectrometry (MS) to identify them. However, EDA studies often fail to identify the bioactive compounds, because even after repeated fractionation, biologically active fractions often remain too complex for chemical identification. This might result in wrong interpretation and erroneous identification of bioactives, especially for those bioactives with high affinity but low abundance. These are usually present in the same fraction with high-

abundance non-bioactive compounds, which impairs their detection. Some snake toxins are also only present in very low amounts. Nerve Growth Factor toxins (NGF) – for example – are only present in 0.1-0.5% in snake venom and are easily missed out. But NGF from the venom from cottonmouth vipers (*Agkistrodon piscivorus*) allowed the authors to study the mechanisms regulating cell and organ growth[99], for which they were awarded the Nobel prize in 1986. Hence it is crucial to take into account these low abundance bioactives in order to make sure the full potential of venoms are realized.

Furthermore, the freeze-drying (lyophilisation) of the chromatographic carrier fluid from the fractionated sample may deactivate toxins, by altering the tertiary structure of proteins by (for instance) unfolding, which might lead to false negatives. Additionally, the EDA workflow requires large amounts of sample, which is difficult in the case of the very precious venoms. Even though some snake species produce relatively large amounts of venom, most don't[166]. Most snakes produce less than a milligram of venom per milking[226], severely hampering efficient venom profiling.

Even though some snake venom proteomes have been completely elucidated and sequenced, many toxins have unknown functions and protein targets. Of these "orphan toxins', only a hypothesised functionality is known by sequence homologies with toxins having known bioactivities. And based on the large functional diversification followed my only minor mutation of some of the toxin families[47], it is difficult to estimate function based on sequence similarity. These toxins are often only classified broadly as cytotoxic, cardiotoxic or weak-neurotoxic[227], including a myriad of different (unknown) receptors and proteins recognized or bound.

Due to current advances in analytical strategies and microfluidics technologies, effective screening for some pharmaceutical targets is currently possible with so-called on-line high-resolution screening (HRS) approaches [228-232]: chemical analysis and biological screening are integrated in a single instrument platform by post-LC continuous-flow biochemical detection and parallel MS analysis[232].

Analyzing natural extracts with a HRS platform can reveal in a single measurement the number and chemical nature of the majority of bioactive compounds, with their affinity towards the target receptor analysed [233-236]. Unfortunately, as stated above, the often yields in animal venoms obtained during a single milking are not compatible with the larger amounts required for traditional on-line HRS approaches. We recently developed a miniaturized HRS methodology for analysis of small compound libraries where nano-LC was hyphenated on-line to a microfluidic biochemical detector[237]. This methodology however, although it would in theory allow for analysis of intrinsically low venom amounts, is unsuited for profiling snake venoms as the complexity in this case overwhelms the analytical system. The main issue here is that the system is only able to assess bioaffinity in one run therefore necessitating a second nano-LC-MS analysis for peptide identification. Now, the main bottleneck becomes apparent as repeatability of nano-LC systems is by far not sufficient for the demanded precision in retention time of eluting peptides. Therefore, the bioactivity of peptides can never be correlated to their identity. For identification purposes, a nano-electrospray MS (nano-ESI-MS) has to be placed in parallel via a post-column split as is traditionally done in on-line HRS approaches[232]. Performing post-column splits after nano-LC, however, is very tedious, error prone and unstable. Furthermore, often the split and the flows cannot be operated in a controlled manner. Therefore, this endeavour has been avoided by the analytical community. Nano-LC-MS is also the analytical technique of choice in proteomics approaches due to its mass-sensitive detection and low sample consumption, and well suited as most snake venoms comprise predominantly peptides and proteins. Many (snake) venomics studies nowadays comprise nano-LC-MS in their analytical strategies[7].

During the last ten years, drug discovery directed at the α7 nicotinic acetylcholine receptor (α7 nAChR) experienced a great leap by using the acetylcholine binding protein[238,239]. Ls-AChBP is a stable structural homologue of the extracellular ligand binding domain of the α7 nAChR. It was first crystallized and validated as model for nAChRs, especially the α7 nAChR, by us[240] in 2001 and has often been used as nAChR model since then. In this regard, it is used as mimicking drug target for example to identify nAChR ligands[241,242], and very well suited for HRS approaches, as was recently demonstrated by us[243].

Subsequently, we developed a miniaturized screening format for bioaffinity profiling of small molecular ligands towards the Ls-AChBP[237]. This system, however, only allowed bioaffinity profiling and did not involve parallel compound identification by MS.

In the current work, the miniaturized HRS approach was applied to efficient venom profiling of mainly snake venoms from species belonging to the Elapidae (cobras, mambas and relatives). The HRS system comprises a post-column nano-flow split which directs the chromatographic effluent to mass spectrometric and on-line biochemical detection allowing for simultaneous correlation of separated bioactive venom toxins with their corresponding accurate mass. This approach may direct subsequent targeted purification (with straightforward LC-MS). The analytical approach led to the identification of at least 20 bioactive peptides with Ls-AChBP affinity and assigned accurate mass, in the venoms of the species examined: the Mozambique spitting cobra (*Naja mossambica*); the black-tailed Jameson's mamba (*Dendroaspis jamesoni kaimosae*); the banded cobra (*Naja annulifera*); and the Cape cobra *(Naja nivea)*. From horned viper (*Vipera ammodytes*) venom - a member of the Viperidae family known for having mainly hemotoxic/proteolytic peptides -  also a low-affinity bioactive peptide was discovered.

## Experimental

### Chemical and biological reagents

Ls-AChBP (from snail species *Lymnaea stagnalis*) was expressed from *Baculovirus* using the pFastbac I vector in Sf9 insect cells and purified from the medium as described by (ref [244]). The fluorescent-tracer ligand DAHBA, (E)-3-(3-(4-diethylamino-2-hydroxybenzylidene)-3,4,5,6-tetrahydropyridin-2-yl)pyridine, was synthesized in-house[243]. NaCl, trizma base, HCl, formic acid (FA), DL–dithiothreitol, guanidine-HCl, rhodamine-110, 5% dimethyldichlorosilane (DMDCS) and α-bungarotoxin (α-BTX) were purchased from Sigma-Aldrich, (Zwijndrecht, The Netherlands). $KH_2PO_4$, $Na_2HPO_4$ and $NH_4HCO_3$ were obtained from Riedel-de Haën (Seelze, Germany). I-Nicotine (99.0%) was purchased from Janssen Chimica (Beerse, Belgium). Enzyme-linked immunosorbent assay (ELISA) blocking reagent (ELISA BR) was purchased from Hoffmann-La Roche (Mannheim, Germany). ULC/MS grade methanol (MeOH;

99.98%) and acetonitrile (ACN; 99.97%) were purchased from Biosolve (Valkenswaard, The Netherlands). HPLC grade water was produced using a Milli-Q purification system from Millipore (Amsterdam, The Netherlands).

## Instrumentation

*Microfluidic confocal fluorescence detection system:* The complete system consisted of a nano-LC unit, an MS, an on-line chip functioning as biochemical reactor, where the LC flow and a bioassay solution were mixed, and a microfluidic LED-based LIF detector as described by (ref [237]). The complete HRS platform is schematically shown in Fig. 29.

*Nano-LC and solvent delivery system:* The Ultimate nano-LC system with a Famos autosampler was from LC Packings (Amsterdam, The Netherlands). The system was operated at 400 nL/min. Mobile phase solvent A consisted of water and 0.1% Tri Fluoro Acetic acid (TFA) and solvent B of ACN and 0.1% TFA. Sample volumes of 500 nL were injected into the analytical capillary column (150 mm × 75 μm internal diameter (i.d.)) packed in-house with Aqua C18 particles (particle size 3 μm, 200 Å pore diameter; Phenomenex, Torrance, CA, USA). For venom analysis, a 90-minute gradient elution was applied running 5 min isocratic at 5% solvent B, then rising to 15% solvent B in 10 minutes, followed by a rise towards 45% solvent B in 60 min, where after solvent B was set to reach 70% during 10 minutes, running isocratically at 70% for 5 minutes. To prevent particulate matter to clog the capillary column, the column inlet was connected to a low dead volume union containing a filter capsule (type M-572 and M-128 respectively, Upchurch Scientific, Oak Harbor, WA, USA). The column exit was connected to a low void volume T-connector where the effluent was split to the MS and the on-line bioassay by two pieces of bare fused-silica (1,000 mm × 10 μm i.d.). All fused-silica capillaries used for connections to and from the chip were purchased from BGB Analytik AG, Schloßböckelheim, Germany.

*Microfluidic chip:* The microfluidic chip and the chip holder (type 4515), produced by Micronit Microfluidics, (Enschede, The Netherlands), was described in detail elsewhere[237]. The open tubular microreactor had a total volume of 4 μL. The reactor had two inlets and one outlet. One inlet was used to

connect the nano-LC carrier flow; the other inlet to infuse the AChBP and tracer ligand DAHBA[237,243] at a flow rate of 5 µL/min by a Model 980532 syringe pump (Harvard Apparatus, Holliston, MA, USA). The chip outlet was connected via a deactivated fused-silica to the microfluidic LIF detector, as described[237].

*Microfluidic LED-based LIF detector:* The flow cell of the detector consisted of a 150 µm i.d. extended light path "bubble-cell" with 50 µm i.d. connecting capillaries (G1600-64232, Agilent Technologies, Amstelveen, The Netherlands). This bubble cell served as the actual flow-through detector cell. Light emanating from a high-intensity LED passed a 465 nm single band pass filter, was collimated by a lens, reflected by a dichroic mirror under 90° and focused into the centre of the bubble cell by 20× quartz microscope objective. Emitted light passed the same dichroic mirror, a focusing lens, and a 520 nm single band pass filter, after which it was detected by photomultiplier tube. A detailed description of this detector can be found elsewhere[237].

*Nano-LC/MS:* A Shimadzu ('s Hertogenbosch, The Netherlands) ion trap – time-of-flight (IT-TOF) hybrid mass spectrometer equipped with a Picoview nano-ESI source from New Objective (Woburn, MA, USA) was operated in positive-ion mode. A 40 mm × 180 µm o.d. × 30 µm i.d. stainless steel emitter was used (ES522, Proxeon/Thermo Scientific, Waltham, MA, USA), functioning as the spray needle. The spray needle was connected to the nano-LC system via a 1,000 mm × 10 µm i.d. bare fused-silica capillary by a low void volume connector (type P-720, Upchurch Scientific) which was integrated in the nano-ESI source. The temperature of the heating block and curved desolvation line were set to 200°C. The interface voltage was set at 1.7 kV, resulting in a current of ~32 µA.

## Biochemical assay and samples

*Biochemical assay:* Fresh solutions of 5 nM Ls-AChBP and 15 nM DAHBA were made every day by dissolving in a bioassay solution containing 1 mM $KH_2PO_4$, 3 mM $Na_2HPO_4$, 0.16 mM NaCl, 20 mM trizma base/HCl at pH 7.5 and 400 µg/mL ELISA BR. The bioassay solution was housed in a 2.5 mL

syringe (type 1002LTN, Hamilton, Bonaduz, Switzerland) which was kept at 4 °C. The on-line bioassay itself was performed at 22 °C.

*Snake venoms:* Lyophilized venoms samples of *Naja mossambica, N. annulifera, Dendroaspis jamesoni kaimosae* and *N. nivea* were acquired as described by (ref 166). *Vipera ammodytes* and *N. pallida* venom were purchased from Sigma Aldrich. About 2 to 3 mg of lyophilized crude venom sample was dissolved in water/ACN/TFA 95:5:0.1% at a concentration of 10 mg/mL and subsequently centrifuged at 13,400 rpm for 10 min to remove particulate matter. Aliquots of these samples were stored at -20°C until further use. Before analysis, 40 µM internal standard nicotine was added together with three peptides at 2 µM each. The resulting venom solutions were directly injected in duplicate onto the nano-LC system for parallel bioaffinity screening and MS identification. Samples were re-analyzed at lower or higher concentrations whenever necessary. The nicotine served as an internal standard and for alignment between the MS chromatograms and the bioassay readout. In the same way, the three reference peptides served to align MS data between different runs, whenever necessary. These peptides were Met-Enkephalin, Mca-Lys-Pro-Leu-Gly-Leu-Dap(Dnp)-Ala-Arg-NH2 (Bachem, Bubendorf, Switzerland) and Enkephalin acetate salt hydrate (Sigma Aldrich).

## Results and Discussion

This works describes the use of an efficient analytical platform for the profiling of neurotoxic snake venoms. The technology is based on an on-line HRS approach in which the on-line biochemical assay assesses bioactivity towards Ls-AChBP, mimicking predominantly neuronal nAChR binding. Parallel MS analysis provides accurate masses of the bioactive peptides for identification purposes. We developed a miniaturized nano-LC setup to overcome the main challenge of the limitation of the amount of venom available per snake species, and to provide sufficient sensitivity. This miniaturized setup was developed previously by using an on-line microfluidic chip (4 µL) and nano-LC[237]. In the current setup (Fig. 29), an additional post-column split is introduced directing half of the 400 nL/min effluent from the nano-LC to

the nano-ESI interface. The other half is directed to the microfluidic chip where it is mixed continuously with the bioassay solution, which is infused at 5 µL/min by a syringe pump. The outlet of the chip is connected to the confocal LED-induced fluorescence detector, which monitors bioaffinity of the eluting toxins after in-flow incubation. The bioassay is based on fluorescence enhancement in which the tracer ligand DAHBA shows increased fluorescence yield upon binding to the AChBP. As an eluting peptide competes with DABHA for Ls-AChBP binding, a negative peak in the signal is observed. Bioactives can subsequently be purified by normal-bore LC(-MS) in a straightforward fashion for further pharmacological studies on nAChRs.
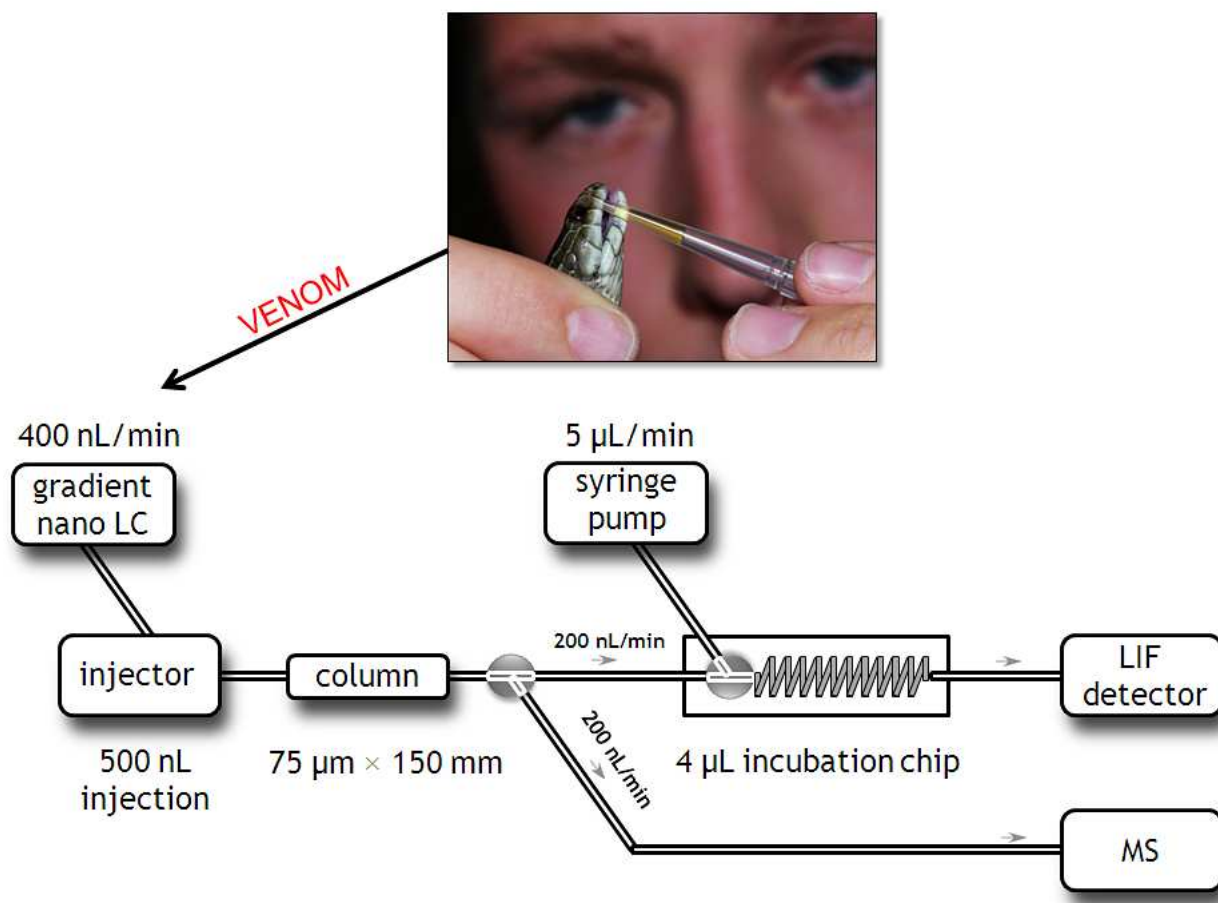
**Figure 29** │ Schematic view of the complete analytical setup from 'milking' venom to our developed method. The bioassay solution containing AChBP and the tracer molecule DAHBA were infused by a syringe pump at 5 µL/min and mixed with the nano-LC effluent running at 400 nL/min, in the 4-µL open tubular reaction chamber of the microfluidic chip. The chip's outlet was hyphenated to the bubble cell capillary wherein light emanating from the high intensity LED was focused. Emitted light from the AChBP/DABHA complex was subsequently detected by the photomultiplier tube. The setup is described in detail in the "Experimental section".

## Method evaluation for the screening of snake venoms

In order to evaluate the applicability and performance of the miniaturized methodology for the screening of bioactive peptides in snake venom, a standard addition study was performed first, to assess sensitivity, reproducibility, correlation of peak shapes observed in MS and the biochemical assay, and (post-column) band broadening. This was done by injecting 5 μg venom from the sand viper (*Vipera ammodytes*) in a 500 nL sample containing increasing amounts of α-bungarotoxin (α-BTX). *Vipera ammodytes* venom is primarily hemotoxic, thus not containing abundant, highly active neurotoxic peptides (with probable affinity towards Ls-AChBP). In this way, the highly active spiked α-BTX was studied in a complex venom matrix without other significant bioactives. This allowed determination of whether the HRS system is capable of pinpointing a single neurotoxic peptide amongst hundreds of other (mainly non-neurotoxic) peptides. Furthermore, this part of the study gives information on the expected limits of detection (LODs) for the venom toxins with respect to their biochemical and MS detection.
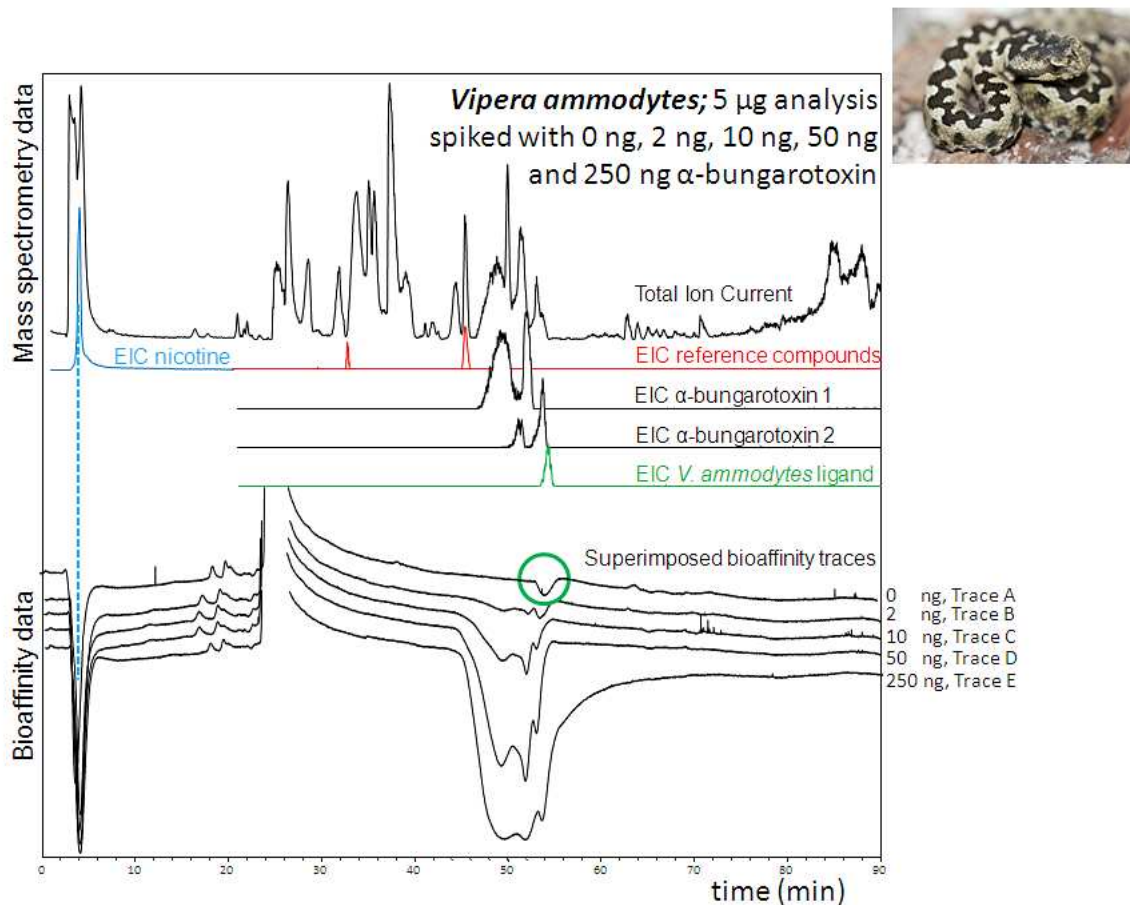
**Figure 30** | The applicability and performance of the analytical setup was evaluated by injecting 5 μg of *Vipera ammodytes* (sand viper) venom, spiked with increasing amounts of α-Bungarotoxin (α-BTX). The mass spectrometric data represents the data obtained for the Vipera venom spiked with 250 ng α-BTX. Shown are the Total Ion Current (TIC) of the spiked *V. ammodytes* venom and the Extracted Ion Currents (EICs) of α-BTX (isoforms 1 and 2 which both elute as two peaks, as also previously reported in literature). Reference peptides allow for alignment with other chromatographic systems. The superimposed bioaffinity data represent 5 injections of the *V. ammodytes* venom with 0 ng, 2 ng, 10 ng, 50 ng and 250 ng α-BTX. The venom analysis without α-BTX (trace A) obtained a previously unreported *V. ammodytes* ligand as is marked by the green circle (corresponding EIC is also shown). Nicotine was added for alignment of the mass spectral data with the bioaffinity data. This experiment was repeated without *V. ammodytes* venom (data not shown) where 1 ng, 2 ng, 10 ng, 50 ng and 250 ng α-BTX were injected. Chromatographic and mass spectrometric performance where comparable to this figure, while the bioaffinity data clearly showed the absence of the *V. ammodytes* venom ligand binding to Ls-AChBP.

91

As can be seen in Fig. 30, α-BTX elutes from the reversed-phase LC column as two iso-forms in 4 peaks, a separation profile earlier reported by[245,246]. This profile can be easily mistaken for being 4 different peptide species, but due to correlating peak shapes between the biochemical and the MS detection, misinterpretations like this can be avoided. The results depicted in Fig. 30 clearly emphasize one of the main advantages of this methodology. In contrast to reversed-phase LC separations of reduced and unfolded peptides in proteomics studies, the LC elution profiles of intact small proteins often do not give single sharp peaks. Nevertheless, the bioaffinity can be correlated to corresponding accurate masses. With the highest amount of spiked α-BTX (250 ng, absolute amount injected), the biochemical assay became saturated, indicating that practically all DAHBA tracer ligand was displaced from Ls-AChBP (Fig. 30, trace E). It was found that the LOD of the high-affinity ligand α-BTX ($K_d$ of 8.59)[240] was ~2 ng (Fig. 30, Trace B) in the biochemical assay and ~25 times higher with MS.

Although the system showed to be very reproducible, sometimes, and especially when samples were repeatedly measured over several days, retention times shifted slightly (data not shown). This is a known issue with nano-LC systems and in our case did not impose serious problems as the biochemical data and MS data were measured simultaneously and despite the shifting retention times, the correlation remained accurate. Additionally, three reference peptides were introduced to enable alignment of chromatograms from different runs, and to allow for convenient transfer of the methodology and samples analyzed to other laboratories, if necessary. In addition, α-BTX was also analyzed separately without the viper venom. This resulted in similar biochemical detection signals, while only α-BTX peaks were observed in the LC-MS data (data not shown). These data confirm that the biochemical assay gives similar results for α-BTX, whether present in a complex venom sample or injected separately, thus also ruling out matrix effects. Finally, in these standard addition experiments, a viper venom peptide was found with affinity for Ls-AChBP, as is observed in the in Fig. 30 trace A (encircled green), as discussed below.

## Snake venom analysis

Snake-venom proteomes consist of a complex cocktail of peptides, which during millions of years of evolution, evolved into highly active ligands towards enzymes and receptors. In general, the *Viperidae* family geared its venom to affect the function of various blood components, while the family Elapidae (including several species from the Colubrid colubrinae subfamily) mainly developed a venom system with neurotoxic and cardiotoxic peptides[107,176] - although this is by no means a golden rule and there are many exceptions to be found in both families. The three-finger toxins (3FTX) – a major and diverse family of low-molecular weight (<10kDa) neurotoxins - share a common tertiary structure. This main body is shaped by four disulfide bridges, truncating the amino acid sequence in such a way that three loops are formed (hence the three "fingers"). All extant members have retained their peculiar "three-fingered" molecular scaffold whilst exhibiting a diversity of different functions[47]. However, despite retaining the common three-finger motif, members of this family have since evolved a myriad of different functions[182]. These include effects on: platelet function[247], different receptors associated with neurotransmission[248-254]; ion channels[255-257,257], viability of cardimyocytes[258] and red bloods cells[259], mitosis and apoptosis[260]; and effects on the cell membrane[261,262]. The different 3FTX members also vary considerably in binding affinity for the different receptors[263] and are important research ligands for studying receptors[264-266]. Homologous 3FTXs are found in all advanced snakes: Elapidae[35] (cobras and relatives), Viperidae[194] (vipers and pitvipers), and others[4,227,267]. Not only their tertiary structure, but also their chain length is similar. Type I (or short-chain) neurotoxins and cardiotoxins usually contain 60 to 62 amino acid residues, while the Type II (or long-chain) neurotoxins comprise 70 to 80 amino acids[182]. Furthermore, these peptides show 40 – 50% amino acid sequence homology. The target affinities of these peptide toxins are 'tuned' by mutations in the extremities of the three loops, as this is where interactions with protein targets take place. For the neurotoxic peptides, the amino acids in the 'fingertips' of Loops I and II are most determining. Type I and Type II neurotoxins bind competitively to (postsynaptic) nicotinic acetylcholine receptors (nAChRs). These neurotoxins are commonly known as α-neurotoxins (α-NTXs)[49,225]. Besides an elongated c-terminus, the Type II α-NTXs differ from Type I by an additional

disulfide bridge at the extremity of Loop II. This difference causes the Type I α-NTXs to mainly target muscle α1 nAChRs, while Type II α-NTXs also target α7 nAChRs. The cardiotoxic (CTX) peptides are essentially cytotoxic peptides as they disturb membrane structures of a range of excitable and non-excitable cells[268]. Most likely, similar to the neurotoxins, there is a wide range of receptors and (ion)channels targeted by the CTX peptides[268-270]. Although mainly α-neurotoxins from Elapidae have been associated with nAChR binding, toxins from the Viperidae have also been reported to show (low) nAChR affinity. In these venoms, phospholipases A2 A (PLA2s) rather than 3Ftxs are responsible for the neurotoxic effects. This family of toxins typically contains seven disulfide bridges and a peptide-chain length of 115 to 139 amino acids. It was reported that crude viperid venoms also block nAChRs (and voltage-gated $Ca^{2+}$ channels) in isolated identified neurons of *L. stagnalis*[271]. Subsequently, they showed that an isolated PLA2 called Bitanarin has affinity for human α7 and *Torpedo californica* nAChRs, as well as for the AChBP (from *L. stagnalis*). This novel ligand is structurally similar to (smaller) PLA2s and shows PLA2 like activity, although it has 14 disulfide bridges and a peptide-chain length of 138 amino acids[271]. Table 3 provides a summary of all peptides, which were mass-correlated with literature via Uniprot. Most toxins have cysteine bridges and we measured the intact peptides. Since Uniprot and most literature give nominal masses from the reduced peptides, the known difference in molecular mass (2 Da per cysteine bridge) was taken into account when the number of cysteine bridges was known. In our case, most of the known bioactive peptides found had 4 or 5 cysteine bridges, implying a molecular mass difference of 8 or 10 Da, respectively. The toxin nomenclature used for the identified toxins, by correlation of nominal masses of bioactive peptides found per snake with literature, is done according to (ref [227]), and/or given as Uniprot identifier for easy reference.

**Table 3** | The table on the next page shows the accurate mass of the most abundant charge state of each bioactive peptide, the corresponding nominal peptide mass, the toxin family the peptide belongs to (when identified and known), the corresponding accession number and their possible mechanism of action. All peptide masses reported in the table and in the following discussion are nominal masses and are correlated with masses found in Uniprot by searching the venom proteomes of the respective snakes profiled. Peptides not correlated in this way are referred to as 'peptide #'.

| | | most abundant isotope pattern | | ~ nominal mass | | |
|---|---|---|---|---|---|---|
| | snake | m/z | charge state | | protein family; sub-family | citable name; uniprot entry name |
| | | | | | | |
| low affinity | *Vipera Ammodytes* | 1540,701 | 9 | 13857,309 | Phospholipase A2 | *peptide 1* |
| very low affinity | *Naja Mossambica* | 1342,665 | 5 | 6708,325 | 3Ftx; Type IA CTX | P01452; CX4_NAJMO |
| **high affinity** | *Naja Mossambica* | 1359,467 | 5 | 6792,335 | 3Ftx; Type IA CTX | P01469; CX2_NAJMO |
| **high affinity** | *Naja Mossambica* | 1378,293 | 5 | 6886,465 | 3Ftx; Type IA CTX | P01470; CX3_NAJMO |
| **high affinity** | *Naja Mossambica* | 1364,674 | 5 | 6818,370 | 3Ftx; Type IA CTX | P01467; CX1_NAJMO |
| low affinity | *Naja Mossambica* | 1368,491 | 5 | 6837,455 | 3Ftx; Type IA CTX | P25517; CX5_NAJMO |
| low affinity | *Naja Mossambica* | 1899,262 | 8 | 15186,096 | *unknown* | *peptide 2* |
| very low affinity | *Dendroaspis jamesoni kaimosae* | 1121,021 | 6 | 6720,126 | 3Ftx; Type IA NTX | P01417; NXS1_DENJA |
| **high affinity** | *Dendroaspis jamesoni kaimosae* | 1159,085 | 6 | 6948,510 | unknown (3Ftx) | peptide 3 |
| **high affinity** | *Dendroaspis jamesoni kaimosae* | 1124,405 | 6 | 6740,430 | unkown (3Ftx) | peptide 4 |
| **co-elution high affinity** | *Dendroaspis jamesoni kaimosae* | 1232,229 | 6 | 7387,374 | 3Ftx; Orphan group XIX | P25682; TX66_DENJA |
| **co-elution high affinity** | *Dendroaspis jamesoni kaimosae* | 1766,309 | 8 | 14122,472 | unknown | peptide 5 |
| co-elution low affinity | *Dendroaspis jamesoni kaimosae* | 1416,428 | 6 | 8492,568 | unknown (3Ftx) | peptide 6 |
| low affinity | *Dendroaspis jamesoni kaimosae* | 1357,036 | 5 | 6780,180 | unknown (3Ftx) | peptide 7 |
| low affinity | *Dendroaspis jamesoni kaimosae* | 1353,840 | 5 | 6764,200 | 3Ftx; Orphan group XI | P01406; TX54_DENJA |
| low affinity | *Dendroaspis jamesoni kaimosae* | 1395,485 | 5 | 6972,425 | unknown | peptide 8 |
| low affinity | *Dendroaspis jamesoni kaimosae* | 1328,074 | 5 | 6635,370 | 3Ftx; Type IA | P01418; NXS1_DENVI |
| **very high affinity** | *Naja Haje annulifera* | 1563,300 | 5 | 7811,500 | 3Ftx; Type IIA NTX | P25674; NXL_NAJHH |
| very low affinity | *Naja Haje annulifera* | 1517,420 | 5 | 7582,100 | 3Ftx; Orphan group II | P01399; TXW3B_NAJHA |
| low affinity | *Naja Haje annulifera* | 1367,719 | 5 | 6833,595 | 3Ftx; Orphan group XV | P62390; CX13A_NAJHA |
| low affinity | *Naja Haje annulifera* | 1359,890 | 5 | 6794,450 | 3Ftx; Type IA CTX | P01461; CX4_NAJHA |
| **very high affinity** | *Naja Nivea* | 1316,441 | 6 | 7892,646 | 3Ftx; Type IIA NTX | P01390; NXL1_NAJNI |
| very low affinity | *Naja Nivea* | 1256,781 | 6 | 7534,686 | 3Ftx; Orphan group II | P25680; TXW10_NAJNI |
| **high affinity** | *Naja Nivea* | 1338,750 | 5 | 6688,750 | 3Ftx; Type IA CTX | P01456; CX1_NAJNI |
| low affinity | *Naja Nivea* | 1370,864 | 5 | 6849,320 | unknown (3Ftx; type) | peptide 9 |
| low affinity | *Naja Nivea* | 1373,651 | 5 | 6863,255 | 3Ftx; Type IA CTX | P01463; CX2_NAJNI |
| co-elution low affinity | *Naja Nivea* | 1380,115 | 5 | 6895,575 | unknown (3Ftx) | *peptide 10* |
| co-elution low affinity | *Naja Nivea* | 1367,715 | 5 | 6833,575 | unknown (3Ftx; Type IA CTX) | *P68419 (NXS1_NAJNI)* |

| snake | # disulfide bonds; reduced mass | reported molecular function |
| --- | --- | --- |
| | | |
| | | |
| *Vipera Ammodytes* | 7; 13871 | catalysis of calcium-dependent hydrolysis of the 2-acyl groups in 3-sn-phosphoglycerides |
| *Naja Mossambica* | 4; 6715 | cardiotoxin |
| *Naja Mossambica* | 4; 6800 | cardiotoxin |
| *Naja Mossambica* | 4; 6894 | cardiotoxin |
| *Naja Mossambica* | 4; 6826 | cardiotoxin |
| *Naja Mossambica* | 4; 6845 | cardiotoxin |
| *Naja Mossambica* | *unknown* | unknown |
| *Dendroaspis jamesoni kaimosae* | 4; 6729 | α1 AChR inhibitor |
| *Dendroaspis jamesoni kaimosae* | unknown | unknown |
| *Dendroaspis jamesoni kaimosae* | unkown | unkown |
| *Dendroaspis jamesoni kaimosae* | 5; 7298 | *unknown/toxin* |
| *Dendroaspis jamesoni kaimosae* | unknown | *unknown* |
| *Dendroaspis jamesoni kaimosae* | unknown | *unknown* |
| *Dendroaspis jamesoni kaimosae* | unknown (4; 6788) | *unknown* |
| *Dendroaspis jamesoni kaimosae* | 4; 6772 | unknown/toxin |
| *Dendroaspis jamesoni kaimosae* | unknown (4; 6980.425) | *unknown* |
| *Dendroaspis jamesoni kaimosae* | 4; 6743 | α1 AChR inhibitor |
| *Naja Haje annulifera* | 5; 7821 | α7 AChR inhibitor |
| *Naja Haje annulifera* | 5; 7592 | *unknown/(weak)toxin* |
| *Naja Haje annulifera* | 4; 6842 | cardiotoxin |
| *Naja Haje annulifera* | 4; 6802 | cardiotoxin |
| *Naja Nivea* | 5; 7902 | α7 AChR antagonist |
| *Naja Nivea* | 5; 7545 | unkown/(weak)toxin |
| *Naja Nivea* | 4; 6697 | cardiotoxin |
| *Naja Nivea* | unknown (4; 6857.320) | uknown |
| *Naja Nivea* | 4; 6871 | cardiotoxin |
| *Naja Nivea* | unknown (4; 6904) | *unknown* |
| *Naja Nivea* | (4; 6842) | *α1 AChR inhibitor* |

# Results

## Sand viper (*Vipera ammodytes*)

As discussed in the previous section, a viper ligand was detected from a 5 µg total venom sample. Figure S1 shows both the mass spectrum and the biochemical detection signal of this analysis. Upon increasing concentrations of α-BTX, this viper bioactive is quickly masked by the highly active α-BTX standard. Table 1 additionally shows mass spectrometric information of this ligand, which could not been identified and is called *peptide 1*.
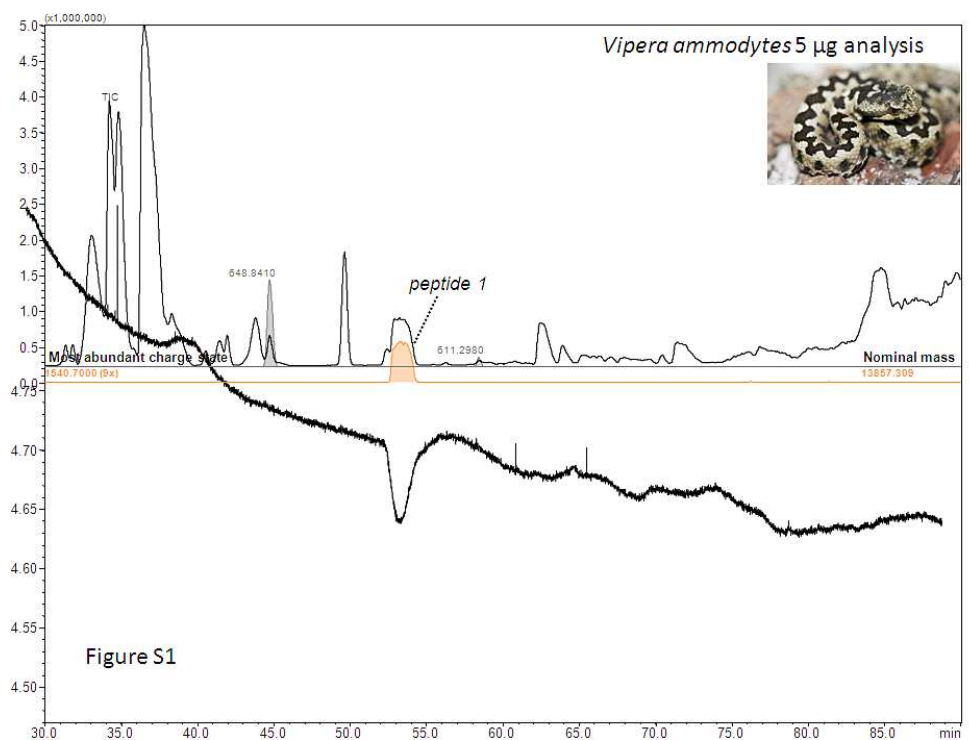


**Figure 31** │ The ligand (peptide 1) obtained from the 5 µg *Vipera ammodytes* (sand viper) venom analysis. The most abundant charge state corresponds to total peptide mass of 13857 Da. Reported Viperidae neurotoxic phospholipases in this mass range typically contain 7 disulphide bridges. Hence the nominal mass of this ligand is probably 13871.

## Mozambique spitting cobra (*Naja mossambica*)

The analysis of *Naja mossambica* venom is a good example how neurotoxic ligands can be screened by the miniaturized on-line HRS methodology, effectively narrowing down the number of probable and possible ligands. When correlating the extracted ion chromatograms (EIC) with the biochemical detection signal of a 5 µg *Naja mossambica* venom analysis, three strong binders and one weaker binder were detected. Their nominal masses correlate with peptides P01469, P01470, P01467 and P25517 respectively (Fig. 32).
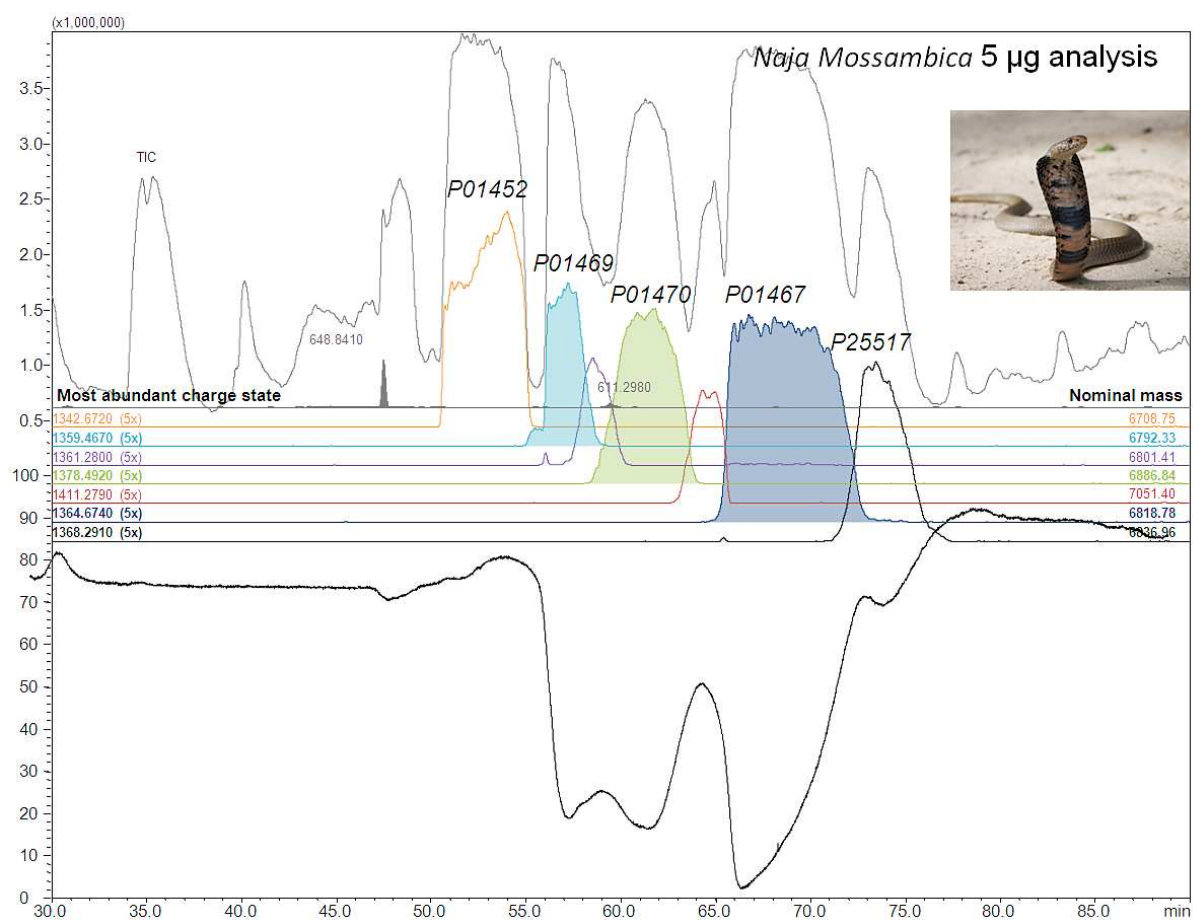


**Figure 32 |** 5 µg injection of *Naja mossambica* (Mozambique spitting cobra) venom. Three strong binders and one weaker binder were detected (P01469, P01470, P01467 and P25517 respectively) which were preciously reported as Type IA cytotoxins, a group consisting of cytotoxins and cardiotoxins.

The first three of these peptides (P01469, P01470 and P01467) share a >90% homology, whereas P25517 has a ~80% homology with the very high affinity ligands as deduced from a Blast search in Uniprot. These toxins have been classified by (ref [227]) by homology to the group Type IA cytotoxins, a group of cardiotoxic and cytotoxic toxins. Interestingly, these four particular toxins have also been found to form a clade separate from the Type IA cytotoxins[227], including being separated from other African cobra species because of their amino acid sequence divergence level. Furthermore, these authors suggested comparative assaying to determine differences in potency or specificity between these toxins and the Type IA group of cytotoxins[227] – that normally exhibit cadiotoxic and cytotoxic effects. Indeed, this proved to be a valid suggestion as we found these toxins showed relatively high affinities to the model Ls-AChBP target and are thus likely to be neurotoxic. This, to our knowledge, has not been reported before. These toxins, together with the highly homologous toxin P01468 from the venom of the *Naja pallida* (closely related to *N. mossambica*) that is the only other found in the clade of (ref [227]) could form a new group of so-called Type IV α-neurotoxins because of their affinity/functionality but being distinctly different in their sequence than, for instance, the Type I α-neurotoxins. So far, predominantly 3FTX-type cardiotoxins have been reported in comprehensive venomics studies on the composition of the *Naja mossambica* venom[272]. In addition, two Type I α1 nAChR ligands have been reported (NMM I and NMM III), one of which was also observed in our LC-MS analysis. It has been shown that this peptide ('NMM I' or 'P01431' 62 amino acids (reduced) molar mass 7081 Da, most abundant charge state *m/z* 1415,635 ), binds exclusively to muscle type α1-nAChRs[273,274]. In our analysis, this peptide did not show affinity towards the AChBP (Fig. 32).

Initially, 5 µg of venom was injected for analysis to a) also observe low-affinity and/or low-abundant binders, and b) to provide sufficient sensitivity in MS. This, however, resulted in severe peak broadening for the high-affinity and/or high-abundant bioactive peptides, as can be seen in Fig. 32. Subsequently, lower venom concentrations were injected and analysed. In this case, sharper peaks were

observed and correlations were more straightforward, which assisted in the total elucidation process (Fig. 33).
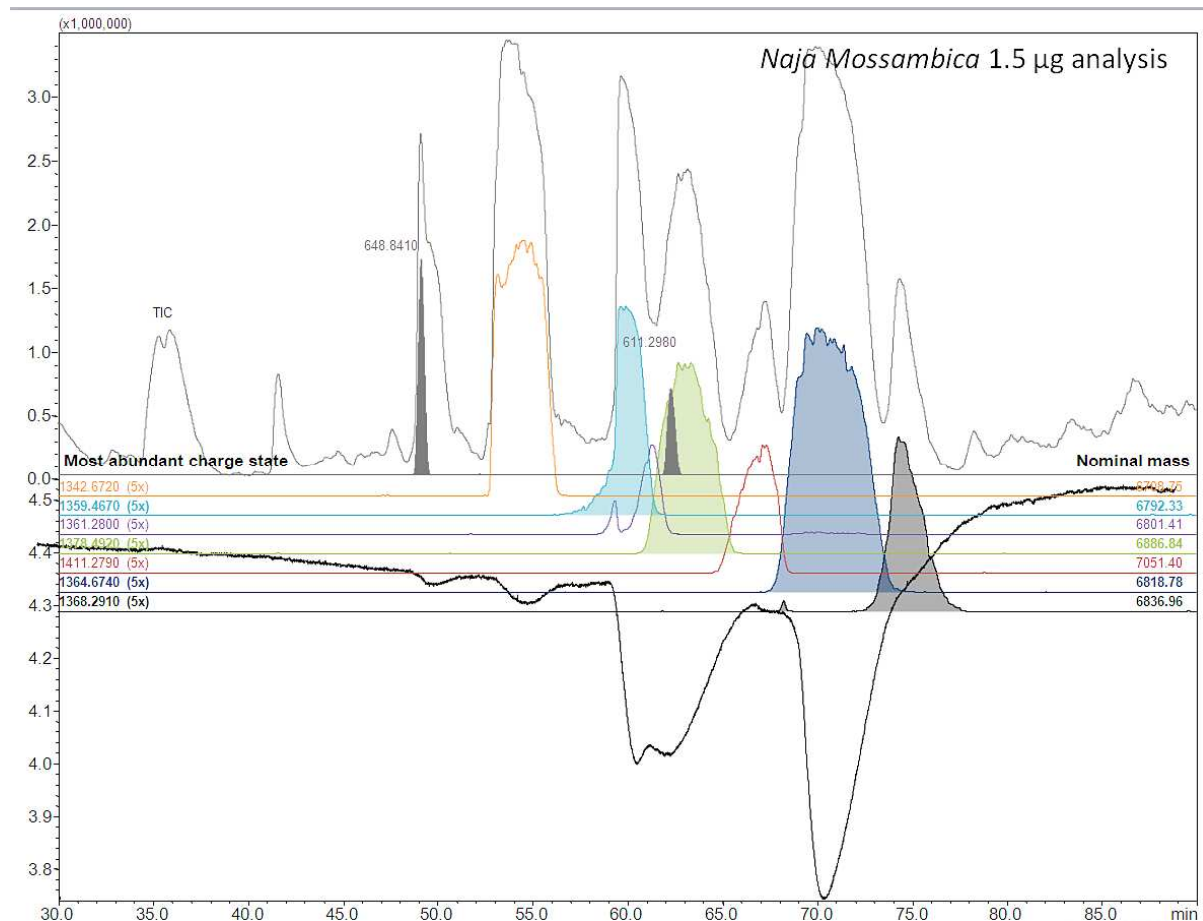


**Figure 33 |** For more straightforward correlation between the mass spectrometric data and the bioaffinity data, 1.5 μg *Naja mossambica* (Mozambique spitting cobra) venom was injected, which obtained sharper peaks and less co-elution of toxins than the 5 μg injection (Fig. 32). To demonstrate the effect of a chromatographic run in terms of selectivity, the pH of the reversed-phase LC run was increased from pH 2.5 to pH 7.5 (data not shown). This resulted in an elution order change where P01469, P01470, P01467 (at pH 2.5) shifted to P01469, P01467, P01470. As a result of shifting retention times another probable ligand could be detected (peptide 2 (Table 3)) with a mass of 15.186.

In order to test P01469, P01470, P01467 and P25517 for α7-nAChR affinity in an off-line radioligand assay, these toxins were separated and fractionated on a normal bore LC-column (4,6mm x 100mm column Symmetry Shield, Waters Etten-Leur, the Netherlands). For the 'large bore' LC measurements, 2.5 mg venom was injected at 2.5 mL/min under the same gradient conditions. The fractions were subsequently lyophilized, re-dissolved in water/ACN/TFA 95:5:0.1% and injected in the on-line miniaturized assay to check for purity and activity. It was found that after fractionation, lyophilisation and re-desolvation most activity was diminished (data not shown). Experiments were carried out such as collecting the fractions in 50 mM ammonium bicarbonate to neutralize the TFA in the LC carrier fluid, but this did not significantly increase activity.

## Jameson's mamba (*Dendroaspis jamesoni kaimosae*)

From a 5 µg venom injection of the *Dendroaspis jamesoni kaimosae* (Jameson's mamba), eight ligands were detected in the biochemical detection signal (Fig. 34). The first (weak) binder, eluting at 33 min, can be mass-correlated to a peptide (P01417) first reported by (ref [275]) as a (weak) nAChR neurotoxin based on anti-venom affinity for the homologous cobra venom. This toxin was then classified as a Type I α-neurotoxin[227]. The massive bioactive peak at 39 min is caused by two peptides (*peptide 3* and *peptide 4*) which could not be mass-correlated with literature. The fourth affinity peak at 44 min was attributed to either co-eluting P25682 or *peptide 5*. P25682 has been classified an orphan toxin XIX (a group of *Dendroaspis* and *Bungarus* toxins of unknown function / protein target)[227]. The onset of binding signal at 48 min is probably caused by *peptide 6*, whereas the remaining signal is caused by P01406, which has been classified as an orphan group XI toxin (of unknown function/target)[227]. P01406 elutes in two forms, also causing another binding signal at 55 min. The small binding signal at 62 min is caused by *peptide 7*, whereas the last signal at 77 min can be attributed to P01418, which has been classified as a Type I α-neurotoxin[227].
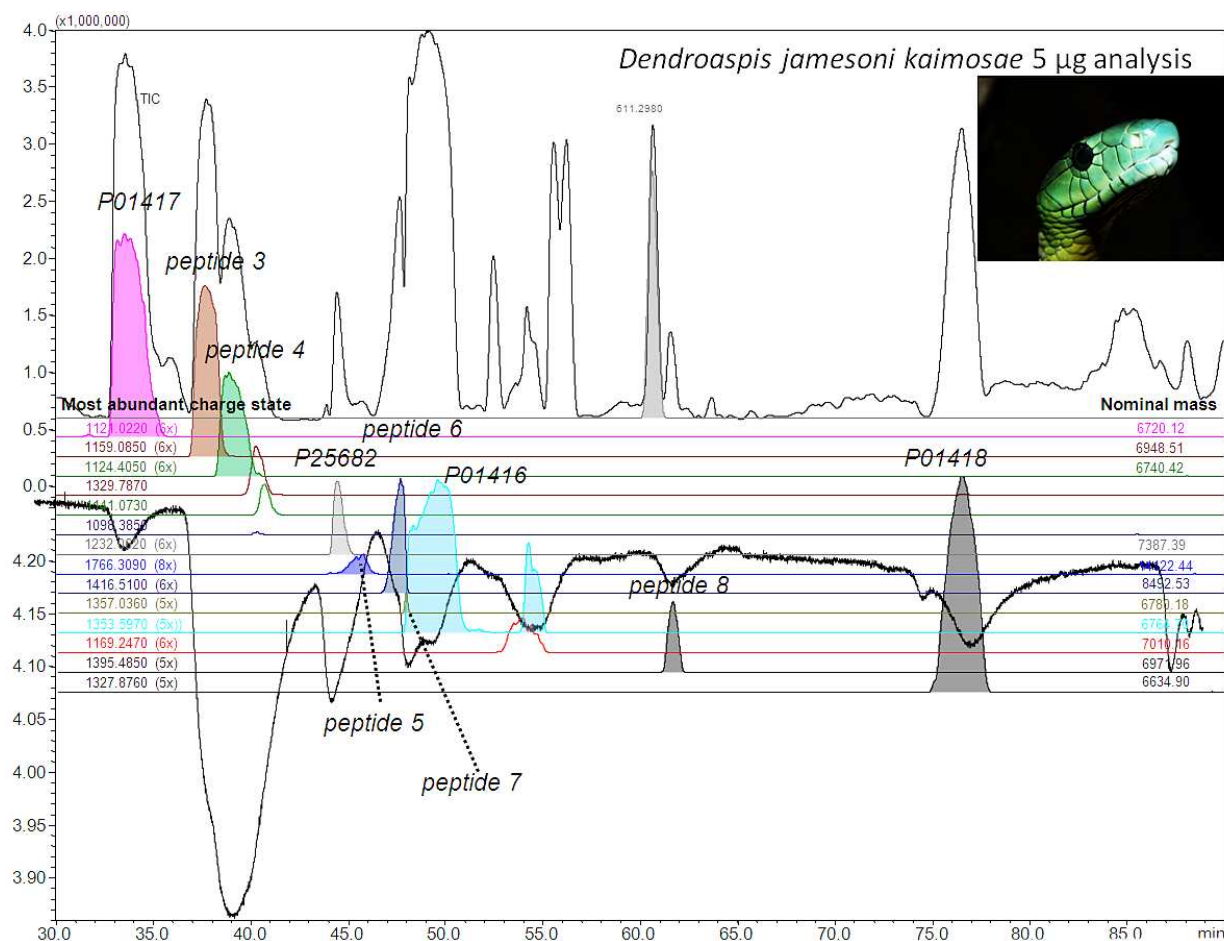
**Figure 34 |** Typical result of a crude snake venom analysis where 5 μg of *Dendroaspis jamesoni kaimosae* (Jameson's mamba) was injected. This analytical run obtained 10 ligands of which 3 are previously reported on in literature; classified as either muscle type neurotoxin or orphan ligand. The venom analysis was repeated with a 1.67 μg injection (Figure 4b) to provide for better chromatographic peak shapes enabling more straightforward correlations. Consequently, less ligands were detected at this venom concentration as some lower affinity neurotoxins reached their level of (biochemical) detection.

A lower, 1.67 μg venom injection provided better chromatographic peak shapes enabling more straightforward correlations (data not shown). In conclusion, the majority of the peptides found in the *Dendroaspis jamesoni kaimosae* venom require subsequent purification of the unknown co-eluting peptides followed by comprehensive functionality assaying and sequencing. This, however, is beyond the scope of the current work.

## African cobra (*Naja annulifera*)

From a 5 µg *Naja annulifera* venom injection, clearly three binding signals can be observed (Fig. 35). The two binding signals eluting at ~ 65 min can be attributed to P62390 and P01461, first reported by (ref [276]). The latter has been classified a Type IA cytotoxin[227]. P01461 sequence-wise is not part of the African Spitting cobra's 'divergence clade' mentioned above, although it shares 81% homology with *Naja mossambica*'s P01467. The binding peak shows that possibly more Type IA cytotoxins exhibit neuronal α-neurotoxicity. P6290 has not been classified by (ref [227]), but by sequence similarity it mostly resembles the 'orphan group XV' toxins (via Blast search in Uniprot), a group coined with only limited known biochemical function besides probable 'low-level cytotoxicity' for *Naja* toxins[227]. P6290 shares, for instance, 90% homology with P1541, a *Naja kaouthia* toxin classified as orphan group XV type[227].

As the MS signals of these peptides are quite high, while the affinity signals are (relatively) low, P62390 and P01461 are probably low-affinity binders. Based on the same criteria, the main ligand eluting between 40 and 50 min is a low-abundant, high-affinity ligand. From the initial analysis at pH 2.5, two co-eluting ligand peptides were identified, which could be mass-correlated with P25674 and P01399 from *Naja haje haje* (Egyptian cobra). When repeated at pH 7.5, P25674 and P01399 showed improved resolution (data not shown). In addition, the MS intensity of peptide P25674 increased, making correlation with the bioaffinity signal more straightforward. P01399 was first reported by (ref [277]) and classified as a weak cytotoxin. A homologue of this peptide has also been found in the *Naja sputatrix* (Indonesian spitting cobra (Q802B2); with 82% homology) and is reported as a weak (muscle type) nAChR ligand[278].
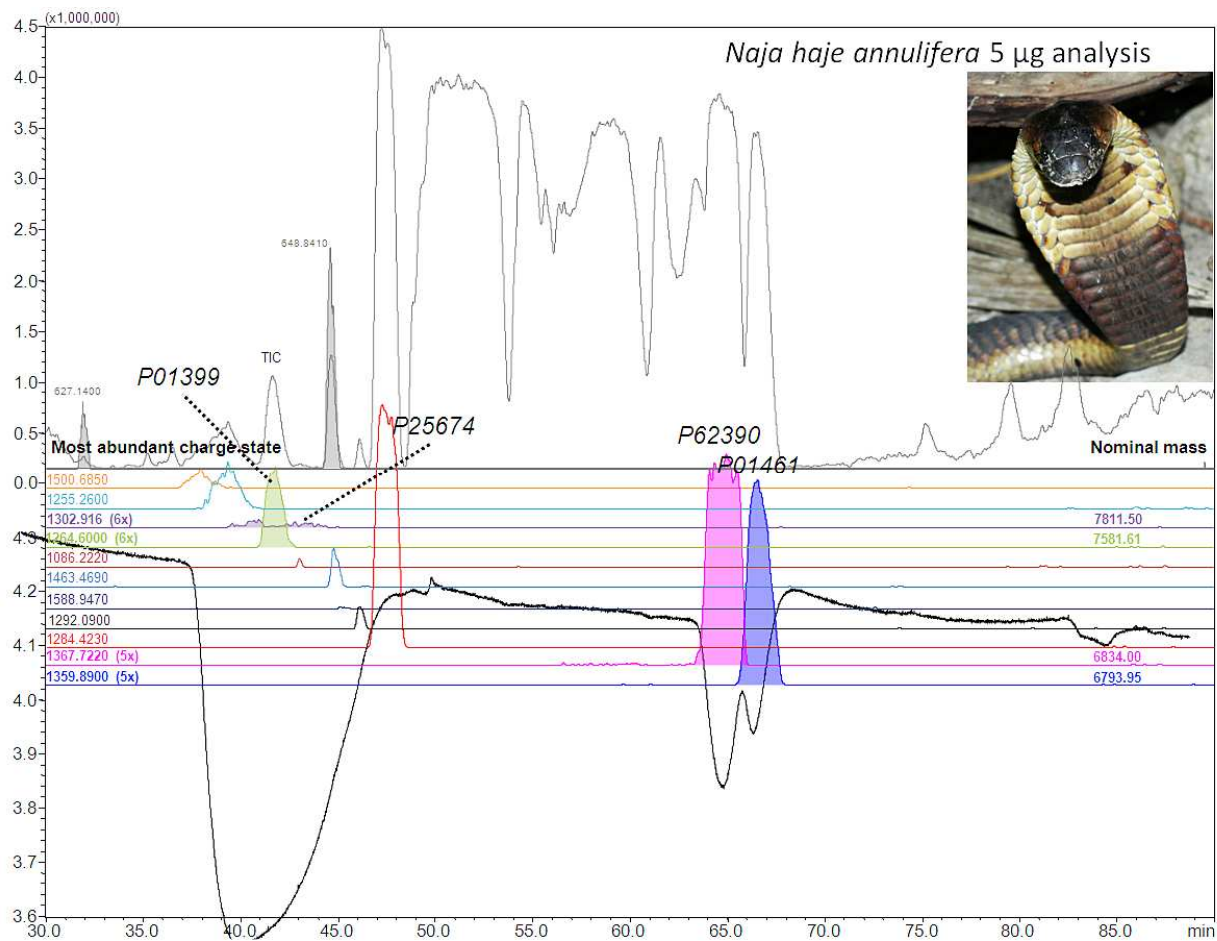
**Figure 35** | Analysis of 5 µg *Naja annulifera* (African cobra) venom obtained three binding signals . The first, large binding signal can be attributed to a highly active α-cobratoxin ligand P25674 (a toxin which has been reported in the venom of *N. haje*). The two latter binding signals were correlated to P62390 and P01461 (previously classified as Type IA cytotoxins). As the α7 ligand P25674 co-elutes with toxin P01399, the analysis was repeated at pH 7.5 (data not shown). P25674 and P01399 showed improved resolution and the MS intensity of peptide P25674 increased. Again this clearly shows the value of the higher pH analysis 'mode'. The analysis was thereafter repeated by injecting 1 µg and 0.2 µg venom (data not shown), which not only allows for more precise correlation between the MS and bioaffinity data sets, but also demonstrates the repeatability of the system.

The main neurotoxic ligand in the *Naja annulifera* venom, P25674, was identified by its homologue reported in the *Naja haje haje* venom, which contains a peptide with an identical mass and has

been reported as a highly active α7 ligand[279]. As these cobra subspecies share many homologues toxin peptides, it can be assumed that this particularly high-affinity peptide has been 100% conserved in sequence in both species. Although it has to be taken into account that in the 70s and 80s many venoms came from species that were mis-identified[280], so it cannot be ruled out that *the Naja haje haje* venom used in (ref [279]) was in fact *Naja annulifera* venom. Lower amounts of venom injected (1 µg and 0.2 µg) provided better chromatographic signals allowing for more reliable biochemical correlations (data not shown). The caveat mentioned earlier and also the rationale for injecting a high amount of venom for analysis: the major bioactive was not observed in MS anymore when looking at the amount of 0.2 µg injected (data not shown). This can easily lead to misinterpretations. Therefore, in some cases like the one described here, different concentrations need to be analysed in order to reach a high confidence regarding correlation of bioactives to their identity. As one additional analytical run with a lower concentration injected takes less than two hours, this is not a significant problem.

### Cape cobra (*Naja nivea*)

*Naja nivea* is closely related to other African cobras such as the *Naja haje* and *Naja annulifera*, as can be traced back to its venom, having a high-affinity α-cobratoxin and several low-affinity Ls-AChBP binders (Fig. 36). At least five ligands could be detected from analyzing 5, 1 and 0.2 µg crude venom samples. The high-affinity binder at 39 min, P01390, is a Type II α-neurotoxin[227]. As due to column overloading a small percentage of the total bioactive P01390 slowly elutes from the column in a tailing fashion, the Ls-AChBP affinity of the remaining neurotoxins is somewhat harder to distinguish (see Fig. 36). This again shows the potential problem of seemingly severe peak broadening when very high-abundant and/or high-affinity ligands are present in a venom sample. It has to be noted that the bioaffinity signals are sigmodial in their response and not linear. This results in seemingly very severe peak broadening in the bioaffinity trace, while this is not observed in the parallel MS (or UV) response. This indicates that it is in fact not severe peak broadening but due to an intrinsic characteristic of biochemical detection in general.
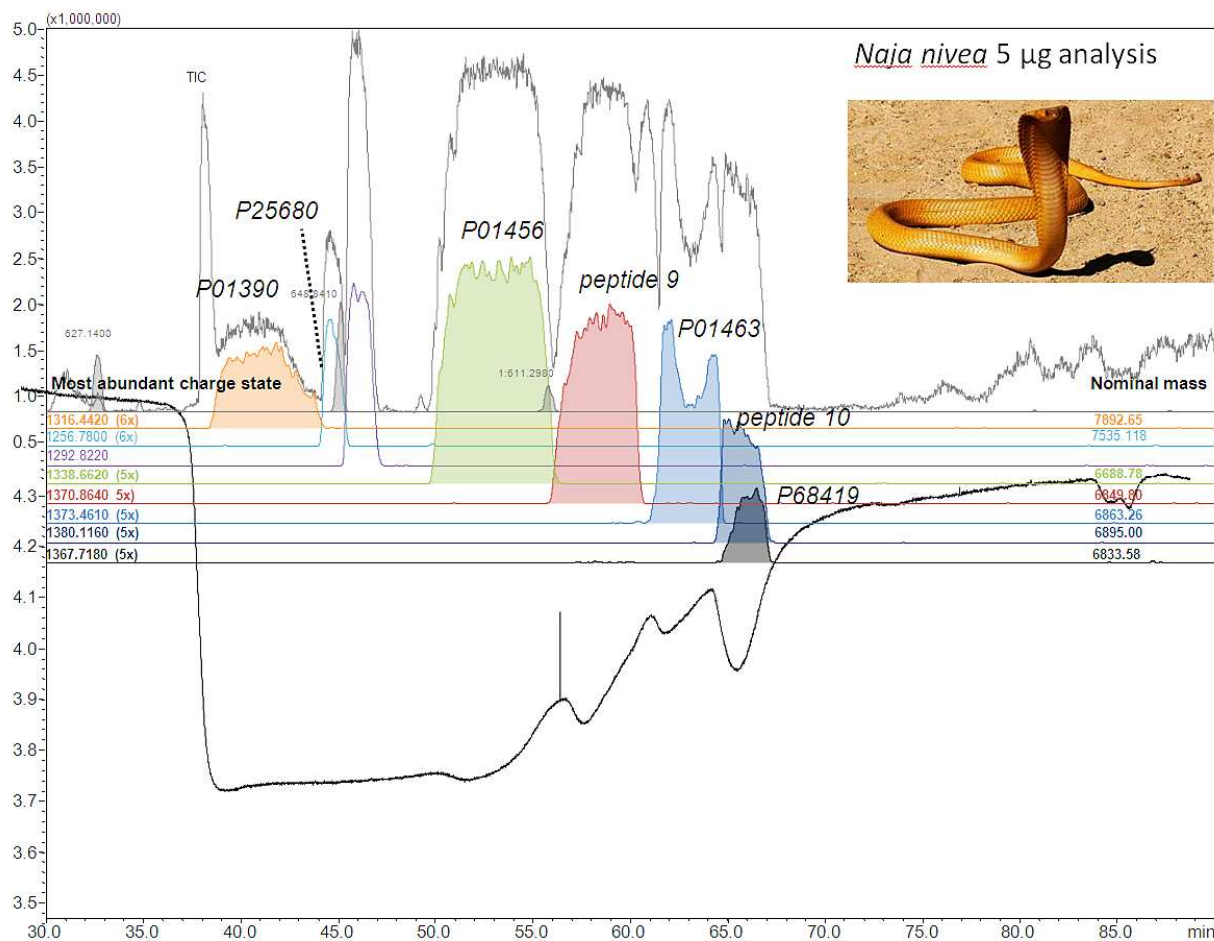
**Figure 36 |** Injection of 5 μg of crude *Naja nivea* (Cape cobra) venom obtained at least 5 ligands. The binding signal at 39 minutes can be attributed to the high-affinity α-cobratoxin P01390. As this strong binder elutes from the analytical column in a very broad peak (and thereby overloading the bioaffinity signal) the remaining toxins are hard to distinguish (such as toxin P25680). This was resolved by injecting smaller concentrations of venom, as seen in Figs. 37 and 38.

Closely eluting to the high affinity binder P01390 is P25680, an orphan group II toxin[227]. It could be a low-affinity binder, but due to the high-binding signal of P01390 no conclusions can be drawn. The bioactivity signal at 57 min can be attributed to P01456, classified as a Type IA cytotoxin[227]. The binding profile at 64 min is caused by a peptide (*peptide 9*) could not be mass-correlated with literature data. At 67 min, the peptide is most probably binding to AChBP is P01463, which is classified as an α1 nAChR ligand. But AChBP bioaffinity of the co-eluting *peptide 8* cannot be ruled out. As also shown for the *Naja mossambica*, *Dendroaspis jamesoni* and *Naja annulifera* venom analysis, analyzing lower concentrations

106

of the *Naja nivea* venom assisted us in more accurate mass-correlations of bioactives, thus facilitating
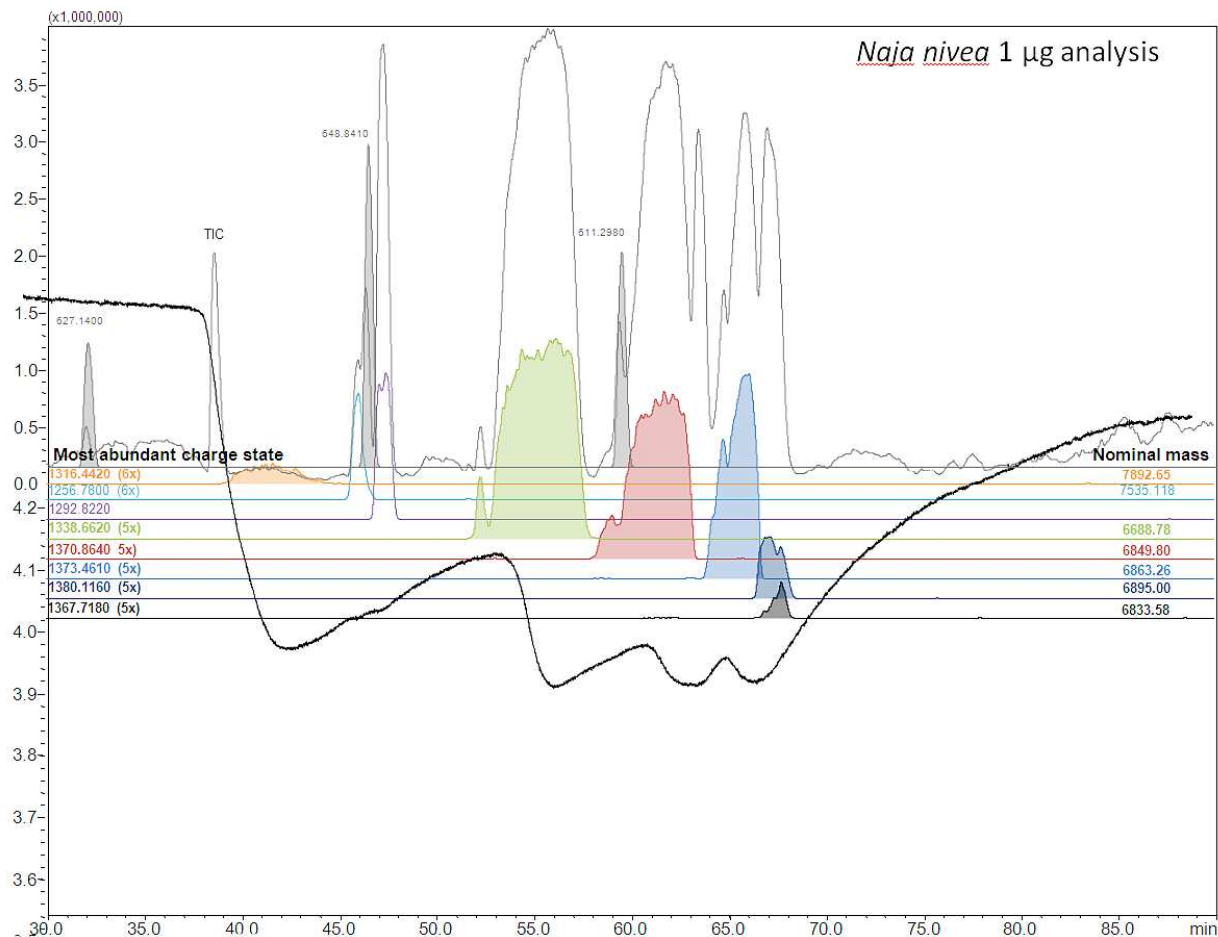their identification.



**Figure 37 |** Injection of 5 μg of crude *Naja nivea* (Cape cobra) venom. From the lower concentration venom analysis it was
possible to correlate at least four other toxins, of which three are classified as Type IA cytotoxins (P01456, P01463 and P68419)
while the other two toxins are not reported on.

The data in these analyses are complementary: while the major bioactive peptides are readily
detected by MS upon a 5 μg injection, their detection becomes more difficult or even impossible at 1 μg
(Fig. 37) and 0.2 μg analysis (Fig. 38). However, in the biochemical detection, the smaller sample
injections result in increased chromatographic resolution and allow the minor bioactives to be analyzed
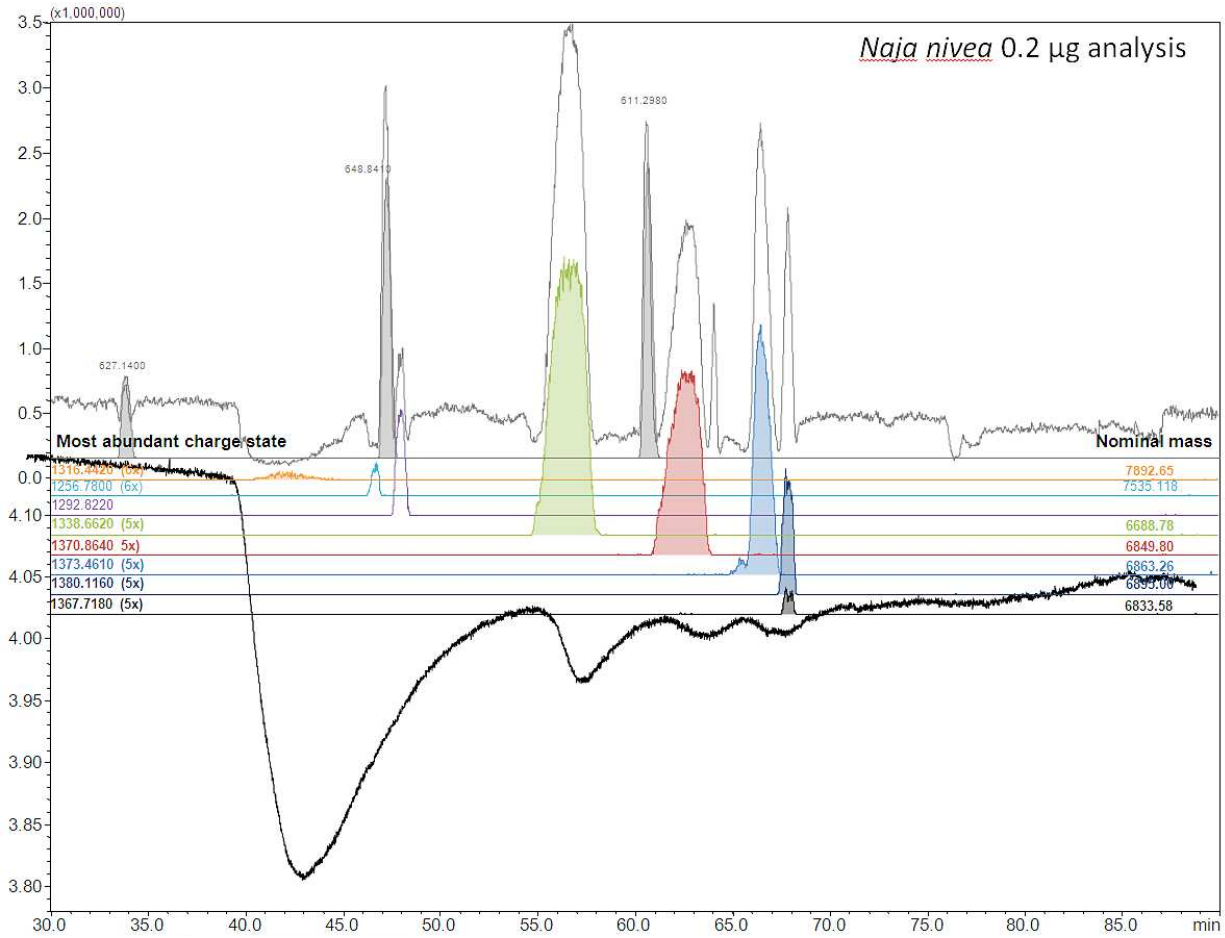more clearly.

**Figure 38 |** Injection of 0.2 μg of crude *Naja nivea* (Cape cobra) venom increased chromatographic resolution and allowed minor bioactives to be analyzed more clearly.

## Discussion and perspectives

The AChBP is an established structural homologue of the α7 nAChR, a (potential) drug target for pain syndromes, cognitive and neurodegenrative diseases. As Type II α-NTXs are high-affinity binders to the α7 nAChR, AChBP is a valid and very useful water-soluble target model to screen for high-affinity binders in these venoms. Snake venoms might contain potential biopharmaceuticals. In case of neurotoxin venoms, particular interest is in targeting nAChRs, e.g., the α7 nAChR. Besides discovery of such potential biopharmaceuticals, screening of venoms may also yield valuable tool compounds with high affinity towards specific ion channels, which can be used as tools for biological and pharmacological studies[281,282]. Typical examples of such snake toxin peptides are the α-bungarotoxins and α-cobratoxins[240,283,284]. This study showed that screening of snake venoms using the HRS approach with the Ls-AChBP yields not only many high-affinity peptide/protein ligands, but also low-affinity Ls-AChBP ligands, which were reported upon in the literature as Type I α-NTXs, cardiotoxins and PLA s. In addition, several other bioactive peptide ligands were found that were previously unreported or reported as orphan-ligands, of which the biological function is presumed, only based on sequence homology.

The data obtained from screening *Naja mossambica* venom clearly shows that ligands still can be identified even though the peptides are closely eluting in relatively broad peaks. For the screening of *Dendroaspis jamesoni kaimosae* venom, this is shown again, but here the number of ligands found is substantially higher, which makes ligand identification more complex. Here, low-affinity binders are first identified by injecting a higher concentration sample, followed by an analysis of more diluted sample(s) to distinguish ligands with high affinity from low-affinity ligands. In the third snake venom profiled, a strategy is shown in which co-eluting peptides from *Naja haje annulifera* venom (of which one has a very high affinity towards Ls-AChBP) are correlated with the biochemical detection signal with higher confidence by means of changing the selectivity of the reversed-phase LC separation followed by reanalysis of the sample. This strategy, where an additional chromatographic run was performed at a high

pH, narrowed possible ligands significantly. The last example showed how a serial dilution analysis of *Naja Nivea* venom enabled the detection of 5 possible AChR ligands.

As described in the method evaluation, the structure of venom peptides is highly conserved in their evolutionary trail causing their chromatographic behaviour also to be similar. Besides poor separation because of many iso-toxins[166], also bad chromatographic peak-shapes pose an analytical challenge. Even though these characteristics result in often challenging chromatographic separations, peptide ligands could be identified by direct correlation of the biochemical detection signal with the EICs from MS as obtained from the IT-TOF MS measurements. Once correlated, these masses give an idea of the identity of peptides (e.g. Type 1 or Type 2). As the correlated masses found are exact (measurements are performed by an accurate and high resolution MS), these masses furthermore provide for a straightforward and targeted fractionation approach. After chromatographic purification, a validation can be performed by re-analyzing the purified sample, followed by further structural elucidation by e.g. tryptic digestion and novo-sequencing, Edman degradation and receptor binding and activation studies.

This study described the evaluation, validation and demonstration of the potential of our miniaturized on-line HRS approach for comprehensive profiling of snake venoms with respect to nAChR affinity. Snake venoms of *Vipera ammodytes*, *Naja mossambica*, *Dendroaspis jamesoni kaimosae*, *N. annulifera* and *N. nivea* were analysed yielding data on both the accurate molecular mass and identity of the peptides present and their individual bioaffinity towards the AChBP.

## Conclusion

As is well known, animal venoms provide for numerous ligands towards a myriad of receptors, enzymes and ion channels and therefore have the intrinsic potential to modulate biological and pharmaceutically relevant signal cascades. This study describes the application of a miniaturized on-line analytical methodology where neurotoxic (Elapidae) snake venoms are screened for toxins with Ls-AChBP affinity, mimicking binding to neuronal nAChRs, especially α7 nAChR. Prior to snake profiling, a post-column nano-split was developed and implemented followed by evaluation and validation of the methodology.

The methodology proved to be suitable for profiling of snake venoms. By correlating bioactive peptides with their mass identified we were able to identify bioactives. In many cases, toxins described in literature are only correlated to a possible function by amino acid homology only.

We found 11 peptides that were not previously correlated with neuronal nAChR affinity and therefore are interesting molecules for further studies on their biological interactions with nAChR. As discussed, many of the toxins sequenced and described in literature are classified by function only by their homologies towards toxins of which the protein target / biological function is known. Although subsequent in-depth studies of the toxins described in this study for nAChR binding are needed, their affinities towards the Ls-AChBP give a good hint of (one of) their function(s): i.e. targeting the nAChRs to causing immobilization of the prey animal. Besides these orphan toxins described, also 10 unknown proteins were found, which are in need of additional sequencing and functional assaying. Even though only the long neurotoxins (Type II) have been correlated with the neuronal α7 nAChR, most of the peptides that were found in this study were short neurotoxins (60 to 64) amino acids. All short neurotoxins described in literature are 6700-6900 Da. Four previously sequenced toxins were found in *Naja mossambica* venom. Although these toxins were classified as cardiotoxins, they did show significant binding to Ls-AChBP. However, further functionality assaying is required to validate human (neuronal) activity. Attempts to do this by fractionation failed as most activity was lost in the lyophilisation procedure. This shows the value of our 'on-line' method as no additional lyophilisation process is required. Therefore, probably, our methodology suffers less from false negatives. Venom analysis from *Dendroaspis jamesoni kaimosae* yielded two ligands known as α1 nAChR ligands, while two other binders are classified as orphan ligands. Further studies are required to validate these results, as for the six other (possible) neuronal nicotinic acetylcholine receptor ligands found in this venom proteome, which, to this date, are not described in literature. The main binder in the *Naja annulifera* proteome was attributed to a known α7 nAChR binder, of a closely related *Naja haje haje*, while two other distinct ligands are correlated to toxins described in literature as cardiotoxin only. The analysis of this particular venom shows the benefit of analysing at both a high and low pH, as this caused selectivity changes in the

reversed phase LC separation leading to more straightforward correlation of the bioactivity and the mass spectrometric detection of a low abundant, high affinity ligand. The *Naja nivea* venom proteome shows four distinct binding 'curves' which can be attributed to five ligands; the largest signal is caused by the a known α7 ligand, while the three other curves can be attributed to tree ligands previously described as cardiotoxin and two unknown peptides, of which one co-elutes with one of the cardiotoxins.

The rapid and high resolution screening of these five venom proteomes shows the potential of the here described microfluidic on-line screening technique. Not only did the application of this technique show the analysis speed, it also demonstrates the advantage of the direct post-column analysis, as fractionation and lyophilisation causes the deactivation of toxins, which in previous runs did show significant signals in the on-line Ls-AChBP assay. Furthermore, by the coupling of a miniaturized separation technique (nano-LC) to a microfluidic on-line assay and sensitive fluorescence detection, we showed that multiple, sensitive and robust analyses are possible with only a minimal amount of (the very precious) venom required. In the future, it may be possible to combine these technologies with the emerging microfluidic whole-animal assay systems such as those being developed in our lab for zebrafish embryos[285].

## Acknowledgements

# Chapter 7: Summary and Discussion

Snakes are a fascinating and intriguing group of vertebrates that have gone to extremes in adaptation. Their venom system belongs to one of the most sophisticated natural weapons systems in the natural world. It is composed of a venom gland, duct and fang, and in some phylogenetic groups an accessory venom gland. In Chapter 2 we have elucidated the evolution of the different types of snake fangs. There are three distinct groups of snake fangs found. The first is a rear-fanged type, where the fangs are positioned in the rear of the upper jaw and this fang-type is found in most snake species. Then there are two types of front-fangs found. The first is found in vipers and pitvipers, where the fangs are often large and lie recumbent in the upper jaw. They are mobile and can move back and forth. The second type is the one found in cobras and mambas and relatives and is smaller and immobile. Because these two groups do not form a monophyletic group, the evolution of these different fang types has always been subject to debate. We have collected many different snake embryos of all three groups and looked at the development of the dentition using the *Sonic hedgehog* gene as a marker. Using *in situ* hybridization followed by careful 3D reconstruction we could follow the development of the fang. Our results showed a great amount of similarity in morphogenesis of the fang of the three different groups of snakes. The two front-fangs start right at the base of the rear of the upper jaw in development, just as the rear-fang. The only difference is that the two front-fangs move forward in development by ontogenetic allometry (pushing the fang forward due to rapid growth of the upper jaw behind the fang) to its adult front position. Also, we showed that the rear-fang develops from its own dental lamina instead of with the same dental lamina as the front 'normal' teeth. It was this posterior dental lamina that showed great similarity to the dental laminae in the two front-fang snake groups. Our results thus suggest a scenario in which all three fang-types are homologous and descendent from a rear-fanged ancestor (by means of uncoupling the posterior from the anterior teeth), after which the two front-fanged groups moved their rear-fang forward during evolution. This probably allowed them to feed on larger and more dangerous prey items as a quick

'snap-and-release' was made possible by the front-fangs compared to the 'bite-and-hold-on' that is only possible with rear-fangs.

Another extreme of adaptation in snakes is their body elongation. Snakes have elongated bodies with numerous vertebrae, some even up to 500 of them. They have lost most clear morphological boundaries in their adult bodies. This has been of major importance in the evolution of the venom-delivery system, as snakes use their elongated bodies as a coiled spring that can strike prey at a distance and with terrifying speed, injecting the immobilizing venom. Without elongated bodies able to strike and bite, their venom-delivery systems (especially the fangs) would probably have evolved in a completely different manner. In chapter 3 we have analyzed and examined the expression of *Hox* genes in different snake embryos from different stages to examine the evolution of s serpentine body form. *Hox* genes determine the basis structure of animals, and it was assumed that the species with a snake-like body evolved such a body by a homogenization of the *Hox* expression domains along the primary axis. However, in our extensive analysis using *in situ* hybridizations we have found - on the contrary - that snakes do retain a collinear expression of the *Hox* genes in the developing embryo. Some *Hox* genes boundaries correspond to expected anatomical boundaries (the most anterior and posterior *Hox* genes), but many do not. The dorsal (thoracic), homogenous rib-bearing region of trunk has also retained regionalized *Hox* expression that is not obviously reflected in the anatomy. Our results suggest that the evolution of a deregionalized, serpentine body involved not only alterations in *Hox* gene cis-regulation but also a different downstream interpretation of the *Hox* code.

The snake venom secretions themselves consist of a complex mixture of proteins and pepties that have evolved to target important physiological pathways and immobilize prey items. These toxins are of major evolutionary and biomedical interest. From an evolutionary perspective, they undergo an accelerated form of evolution. The genes undergo accelerated evolution due to a larger amount of mutation in coding regions than in non-coding regions, while in almost all other known genes in the natural world this is the other way around - except the genes in the immune system. This allows fast

evolution and adaptation towards ever changing prey and prey physiology (evolution of resistance). In an analogous manner, our immune systems must fight ever changing pathogens. This arms-race is the driving force behind diversity. From a biomedical perspective, this diversity is right at the base of why snake venoms form such a potential goldmine for new pharmaceuticals. In chapter 4 we have sequenced the king cobra (*Ophiophagus hannah*) genome to look at the evolution (and recruitment) of the venom genes. They are thought to have evolved from normal, physiological genes by means of gene duplication followed by selective expression in the venom gland. However, in the absence of genomic resources these hypotheses remain mainly speculative. Our results show that venom genes evolve through distinct mechanisms. L-amino acid oxidase, cysteine-rich secretory proteins and metalloproteinases, evolved by tandem-duplication of ancestral physiological genes, followed by recruitment through selective expression in the venom gland. By contrast, nerve growth factor toxins appear to have evolved by duplication and dual recruitment, while hyaluronidase and phospholipase B evolved by recruitment of existing physiological genes without further duplication. The massive toxin families the 3FTXs and the PLAs have undergone extensive duplications that have resulted in (at least) 21 copies in the genome. These results show a much more dynamic evolutionary history in venom genes as previously assumed. It is yet unclear what is causing the recruitment into the venom gland, and we are currently working on answering this question.

The enormous diversity in venom genes, caused by the different ways of evolution and recruitment as shown in chapter 5, makes it a challenge to identify bioactives for a particular receptor in order to either study the venom for evolutionary studies, or to find new pharmaceuticals. In addition to this, the extreme low venom yield that most snakes have make it a very precious sample. For most traditional methods of evaluating a sample with mass spec and bioactivity studies you need a relatively large amount of venom, which is a major challenge for many species. We have developed in chapter 6 a new method to study the make-up of the venom and - at the same time and in the same run - look for bioactives that bind to acetylcholine binding protein. One test used up less than 5 ug of venom and took

less than 2 hours. We found 21 bioactives of which 7 were never reported on before. This technique includes a microfluidic on-line screening approach utilizing nano-LC–MS with parallel bioaffinity detection using the acetylcholine binding protein (AChBP). The system comprises a post-column nano-flow split which directs the chromatographic effluent to mass spectrometric and on-line biochemical detection allowing for simultaneous correlation of separated bioactive venom toxins with their corresponding accurate mass.

In conclusion, in this thesis I have shown that snakes have undergone multiple changes in their genome and embryonic development that has provided them with the variation to which natural selection could act. This thesis provides evidence for the variable mechanisms of venom gene evolution, which presumably is much more flexible than previously thought. But it also underscores the potential use of the many different types of snake venom toxins that could be screened for use against human disorders. And most of all, I hope I have contributed towards the fact that snakes are just an incredibly interesting group of vertebrates from both the perspective of ecology and life-style, as well as from a genomic and molecular perspective. They are, and will always be, my first and true love.

## Samenvatting en Discussie

Slangen zijn een fascinerende en intrigerende groep gewervelde dieren die tot extremen zijn gegaan in evolutionaire adaptatie. Het gifsysteem is bijvoorbeeld een van de meest geavanceerde biowapensystemen uit de natuur. Het omvat een gifklier, een gifkanaaltje, een giftand, en in sommige fylogenetische groepen een 'accessory gland'. In hoofdstuk 3 hebben we een hypothese opgesteld voor de evolutionaire oorsprong van de verschillende typen giftanden. We kunnen drie typen van elkaar onderscheiden. De eerste is een kleine, onbeweeglijke giftand die achter in de bek staat, en dit type giftand vinden we bij de meeste slangensoorten. Dan zijn er twee verschillende typen van giftanden die voorin de bek staan. Het eerste type is te vinden bij de familie Viperidae, dat bestaat uit adders en ratelslangen (groefkopadders). Hierbij zijn de giftanden vaak groot en liggen ze horizontaal in de bovenkaak. Ze zijn beweeglijk en kunnen op en neer bewegen door een scharnier. Het tweede type is te vinden in de familie van de Elapidae, de cobra's, mamba's en verwanten. Hier is de giftand vaak kleiner en onbeweeglijk. Omdat deze twee families geen monofyletische groep vormen, is de evolutie van deze giftanden onderwerp geweest van discussie. Om dit te onderzoeken hebben wij een grote hoeveelheid slangen embryo's verzameld van verschillende soorten, behorend tot alle drie de giftand typen. Hierbij hebben we de ontwikkeling van het gebit gevolgd met het gen *sonic hedgehog* als marker. Door middel van *in situ* hybridisatie en 3D reconstructie konden we de algehele ontwikkeling van de giftand in kaart brengen. De resultaten waren verbluffend: er was een grote mate van gelijkenis te vinden in de ontwikkeling van alle drie de typen. Bij de Elapidae en de Viperidae begon de ontwikkeling van de giftanden in de embryonale ontwikkeling bij beiden achter in de bek, net als bij de overige slangen. Het grote verschil is echter dat de giftanden van de Elapidae en de Viperidae tijdens de ontwikkeling naar voren werden geduwd door ontogenetische allometrie, doordat de kaak achter de giftand sneller groeide dan het gedeelte ervoor. Bij achtertand giftige slangen ontstaat de giftand vanuit zijn eigen onafhankelijke dental laminae (embryonaal primordium), en deze vertoont grote gelijkenis met de ontwikkeling van de dental laminea van de giftanden van de Elapidae en de Viperidae.  Onze resultaten suggereren een scenario waarin alle drie de

117

giftanden homoloog zijn aan elkaar, afkomstig van een giftand die achterin de bek stond. Hierna hebben de Elapidae en de Viperidae onafhankelijk van elkaar de giftand naar voren geschoven in de evolutie. Vermoedelijk kwam dit door het bijkomstig evolutionair voordeel dat deze slangen grotere en gevaarlijkere prooien aan konden door middel van een snelle "bijt en laat los" methode. De slangen met hun giftand achter in de bek kunnen dit niet, en moeten hun prooi vasthouden om gif in de wond te kauwen. Dit is onmogelijk bij gevaarlijke prooien zoals grote ratten.

Slangen hebben verlengde lichamen met tot 500 wervels. Ze hebben de duidelijke morfologische grenzen van hun lichaam verloren die veel andere dieren (inclusief mensen) wel hebben. Het verlenging van het lichaam is van groot belang geweest voor de evolutie van het gif systeem, omdat slangen hun lange lichamen kunnen gebruiken als een springveer om prooien met grote snelheid te bijten, en gif te injecteren. Zonder deze lange lichamen om te bijten, zal het gifsysteem – en zeker de giftanden - waarschijnlijk op een hele andere manier zijn geëvolueerd. In hoofdstuk 4 hebben we de expressie van *Hox* genen bekeken in verschillende slangen embryo's van verschillende stadia om de evolutie van een 'slangachtig lichaam' te bestuderen. *Hox* genen bepalen de lichaamsvorm van dieren. Het was algemeen aangenomen dat dieren met een 'slangachtig lichaam' dit geëvolueerd hadden door een homogene expressie van *Hox* genen langs de primaire lichaamsas, in tegenstelling tot de altijd zo karakteristieke colineaire (regionaliserende) expressie. Echter, in onze experimenten met behulp van *in situ* hybridisatie hebben wij juist het tegenovergestelde gevonden. Slangen hebben wel de colineaire expressie van *Hox* genen behouden in hun embryonale ontwikkeling. De grenzen van de expressie van sommige *Hox* genen corresponderen met de te verwachtte anatomische grenzen (de most anteriore en posteriore *Hox* genen), maar bij veel andere *Hox* genen is dat niet het geval. Het dorsale (thoracische), homogene rib-dragende deel van de slang heeft ook de regionaliserende *Hox* expressie patronen behouden. Deze grenzen zijn echter niet terug te vinden in de anatomie van een slang. Onze resultaten suggereren daarom dat de evolutie van een gederegionaliseerde, 'slangachtig lichaam' niet alleen veranderingen in de expressie van *Hox* genen zelf verreist maar ook een andere downstream interpretatie van de *Hox* code.

Slangengif bestaat uit een complex mengsel van verschillende eiwitten en peptiden die geëvolueerd zijn om prooien te immobiliseren door bepaalde belangrijke fysiologische doelen aan te vallen. Deze gifstoffen zijn zeer interessant voor de evolutionaire en biomedische wetenschap. Ze ondergaan bijvoorbeeld een vorm van 'versnelde evolutie', door een groter aantal mutaties in de coderende stukken DNA (exons), in relatie tot de mutaties in de niet-coderende stukken DNA (introns). In de meeste genen is dit juist andersom. Eenzelfde patroon vinden we in de genen die betrokken zijn bij immuunsystemen. Wij moeten ons wapenen tegen alsmaar veranderende pathogenen, maar op een analoge wijze moeten slangen zich wapenen tegen alsmaar veranderende prooien die resistentie ontwikkelen. Deze wapenwedloop is de drijvende kracht achter de diversiteit in gif. Vanuit een biomedisch oogpunt, zijn deze stoffen interessant omdat ze een grote variatie vertonen en daarmee een potentiële goudmijn vormen voor nieuwe farmaceutische stoffen. In hoofdstuk 5 hebben we het genoom van de koningscobra (*Ophiophagus hannah*) in kaart gebracht. We hebben gekeken naar de oorsprong, evolutie en werving van gifgenen in de gifklier. Er werd gedacht dat gifgenen volledig evolueren door middel van de duplicatie van normale, fysiologische genen gevolgd door werving van de gifklier. Maar omdat er nog geen slangengenoom beschikbaar was, bleef dit speculatie. Onze resultaten laten zien dat gifgenen evolueren door meerder verschillende mechanismen, en dat de evolutie van gifgenen dus flexibeler is dan dat voorheen werd gedacht. L-amino acid oxidase, cysteine-rich secretory proteins en metalloproteinases zijn bijvoorbeeld geëvolueerd door tandem duplicatie van fysiologische genen, gevolgd door werving door de gifklier. Maar, nerve growth factor toxins zijn geëvolueerd door duplicatie waarbij beide kopieën door de gifklier zijn gerekruteerd, terwijl bij hyaluronidase en phsopholipase B geëvolueerd zijn door werving van al bestaande genen zonder duplicatie. De three-finger toxins en phospholipases A2 hebben beiden veel duplicaties ondergaan dat heeft geresulteerd in meerdere kopieën in het genoom. De resultaten laten zien dat de evolutionaire geschiedenis van gifgenen dynamischer is dan tot nu toe werd aangenomen. Hoe de gifgenen door de gifklier worden gerekruteerd is nog onbekend, maar deze vraag zal hopelijk in de nabije toekomst beantwoord kunnen worden door zorgvuldige analyse van de promotor regio's in de gifgenen.

De enorme variatie binnen gifstoffen maken het een uitdaging om stoffen te vinden die aan een bepaalde interessante receptor aanhechten. Dit kan zowel belangrijk zijn voor evolutionaire studies als voor de zoektocht naar nieuwe farmaceutische interessante stoffen. Het feit dat slangen vaak maar een hele kleine hoeveelheid gif produceren maakt dit niet makkelijker. Bij de meeste traditionele methodes waarbij een monster wordt geanalyseerd met massaspectrometrie gevolgd door een zoektocht naar een bepaalde bioactieve stoffen is een relatieve grote hoeveelheid van het monster nodig, wat een uitdaging is bij veel soorten slangen. In hoofdstuk 6 hebben we een nieuwe methode ontwikkeld waarbij we in een run zowel de samenstelling van het gif via massaspectrometrie kunnen bepalen, als ook bioactieve stoffen kunnen vinden die hechten aan de zogenaamde acetylcholine binding protein (een analoog van de acetylcholine receptor, die vermoedelijk betrokken is bij verscheidene menselijke aandoeningen). Voor het uitvoeren van een enkele test hebben we minder dan 5ug gif nodig, en minder dan twee uur tijd. We hebben 21 bioactieve stoffen gevonden, waarvan 7 nog niet waren gerapporteerd. Deze techniek omvat een microfluidics on-line screening dat een nano-LC-MS gebruikt, met daaraan gekoppeld een bioaffinity detectie met de acteylcholine binding protein. Hierdoor kan het systeem (dat gebruik maakt van een post-column nano-flow split) gelijktijdig de gescheiden bioactieve gifstoffen correleren aan hun bepaalde massa.

Ik heb in dit proefschrift laten zien dat slangen verscheidene veranderingen in hun genoom en embryonale ontwikkelingen hebben ondergaan die ze de variatie hebben gegeven waarop natuurlijke selectie druk kon uitoefenen. Ik heb in dit proefschrift bewijs geleverd dat gifgenen op verschillende manieren evolueren, en daarin meer flexibel zijn dat gedacht werd. Ik onderstreep ook de potentie die de verschillende stoffen in het slangengif hebben  voor de farmaceutische wetenschap. Maar bovenal, heb ik hopelijk bijgedragen aan het feit slangen een waanzinnig interessante groep gewervelde dieren zijn vanuit zowel het perspectief van ecologie, alsook genomics en evolutie. Ze zullen altijd mijn eerste en grote liefde blijven!

Reference List

1 Vidal,N. *et al.* (2007) The phylogeny and classification of caenophidian snakes inferred from seven nuclear protein-coding genes. *C. R. Biol.* 330, 182-187

2 Vonk,F.J. *et al.* (2008) Evolutionary origin and development of snake fangs. *Nature* 454, 630-633

3 Lawson,R. *et al.* (2005) Phylogeny of the Colubroidea (Serpentes): new evidence from mitochondrial and nuclear genes. *Mol. Phylogenet. Evol.* 37, 581-601

4 Fry,B.G. *et al.* (2003) Isolation of a neurotoxin (alpha-colubritoxin) from a nonvenomous colubrid: evidence for early origin of venom in snakes. *Journal of molecular evolution* 57, 446-452

5 Fry,B.G. *et al.* (2008) Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol. Cell Proteomics.* 7, 215-246

6 Calvete,J.J. (2009) Venomics: digging into the evolution of venomous systems and learning to twist nature to fight pathology. *J Proteomics.* 72, 121-126

7 Fox,J.W. and Serrano,S.M. (2008) Exploring snake venom proteomes: multifaceted analyses for complex toxin mixtures. *Proteomics.* 8, 909-920

8 Georgieva,D. *et al.* (2008) Proteome analysis of snake venom toxins: pharmacological insights. *Expert. Rev. Proteomics.* 5, 787-797

9 Gibbs,H.L. *et al.* (2009) Snake population venomics: proteomics-based analyses of individual variation reveals significant gene regulation effects on venom protein expression in Sistrurus rattlesnakes. *J. Mol. Evol.* 68, 113-125

10 Barlow,A. *et al.* (2009) Coevolution of diet and prey-specific venom activity supports the role of selection in snake venom evolution. *Proc. Biol. Sci.* 276, 2443-2449

11 Olamendi-Portugal,T. *et al.* (2008) Proteomic analysis of the venom from the fish eating coral snake Micrurus surinamensis: novel toxins, their function and phylogeny. *Proteomics.* 8, 1919-1932

12 Oguiura,N. *et al.* (2009) Intraspecific variation of the crotamine and crotasin genes in Crotalus durissus rattlesnakes. *Gene* 446, 35-40

13 Gibbs,H.L. and Mackessy,S.P. (2009) Functional basis of a molecular adaptation: prey-specific toxic effects of venom from Sistrurus rattlesnakes. *Toxicon* 53, 672-679

14 Salazar,A.M. *et al.* (2009) Venom variation in hemostasis of the southern Pacific rattlesnake (Crotalus oreganus helleri): isolation of hellerase. *Comp Biochem. Physiol C. Toxicol. Pharmacol.* 149, 307-316

15 Daltry,J.C. *et al.* (1996) Diet and snake venom evolution. *Nature* 379, 537-540

121

16 Gutierrez,J.M. *et al.* (2010) Impact of regional variation in Bothrops asper snake venom on the design of antivenoms: integrating antivenomics and neutralization approaches. *J. Proteome. Res.* 9, 564-577

17 Salazar,A.M. *et al.* (2007) A comparative analysis of the clotting and fibrinolytic activities of the snake venom (Bothrops atrox) from different geographical areas in Venezuela. *Thrombosis research* 120, 95-104

18 Currier,R.B. *et al.* (2010) Intra-specific variation in venom of the African Puff Adder (Bitis arietans): Differential expression and activity of snake venom metalloproteinases (SVMPs). *Toxicon* 55, 864-873

19 Sanz,L. *et al.* (2006) Venom proteomes of closely related Sistrurus rattlesnakes with divergent diets. *J. Proteome. Res.* 5, 2098-2112

20 Williams,V. *et al.* (1988) Variation in venom proteins from isolated populations of tiger snakes (Notechis ater niger, N. scutatus) in South Australia. *Toxicon* 26, 1067-1075

21 Alape-Giron,A. *et al.* (2009) Studies on the venom proteome of Bothrops asper: perspectives and applications. *Toxicon* 54, 938-948

22 Antunes,T.C. *et al.* (2010) Comparative analysis of newborn and adult Bothrops jararaca snake venoms. *Toxicon*

23 Calvete,J.J. *et al.* (2010) Snake venomics of the Central American rattlesnake Crotalus simus and the South American Crotalus durissus complex points to neurotoxicity as an adaptive paedomorphic trend along Crotalus dispersal in South America. *J Proteome. Res.* 9, 528-544

24 Menezes,M.C. *et al.* (2006) Sex-based individual variation of snake venom proteome among eighteen Bothrops jararaca siblings. *Toxicon* 47, 304-312

25 Kini,R.M. and Chan,Y.M. (1999) Accelerated evolution and molecular surface of venom phospholipase A2 enzymes. *Journal of molecular evolution* 48, 125-132

26 Calvete,J.J. (2009) Venomics: digging into the evolution of venomous systems and learning to twist nature to fight pathology. *J. Proteomics.* 72, 121-126

27 Kini,R.M. and Chinnasamy,A. (2010) Nucleotide sequence determines the accelerated rate of point mutations. *Toxicon* 56, 295-304

28 Doley,R. and Kini,R.M. (2009) Protein complexes in snake venom. *Cell Mol. Life Sci.* 66, 2851-2871

29 Fujimi,T.J. *et al.* (2003) Molecular evolution and diversification of snake toxin genes, revealed by analysis of intron sequences. *Gene* 313, 111-118

30 Tamiya,T. *et al.* (1999) Complete nucleotide sequences of cDNAs encoding long chain alpha-neurotoxins from sea krait, Laticauda semifasciata. *Toxicon* 37, 181-185

31 Servent,D. *et al.* (1997) Only snake curaremimetic toxins with a fifth disulfide bond have high affinity for the neuronal alpha7 nicotinic receptor. *J. Biol. Chem.* 272, 24279-24286

32 Pawlak,J. *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

33 Doley,R. *et al.* (2009) Role of accelerated segment switch in exons to alter targeting (ASSET) in the molecular evolution of snake venom proteins. *BMC. Evol. Biol.* 9, 146

34 Chang,C.C. *et al.* (1973) Studies of the presynaptic effect of -bungarotoxin on neuromuscular transmission. *J. Pharmacol. Exp. Ther.* 184, 339-345

35 Habermann,E. and Breithaupt,H. (1978) Mini-review. The crotoxin complex--an example of biochemical and pharmacological protein complementation. *Toxicon* 16, 19-30

36 Atoda,H. and Morita,T. (1989) A novel blood coagulation factor IX/factor X-binding protein with anticoagulant activity from the venom of Trimeresurus flavoviridis (Habu snake): isolation and characterization. *J. Biochem. (Tokyo)* 106, 808-813

37 Kini,R.M. (1997) Phospholipase A2 A Complex Multifunctional Protein Puzzle. In *Venom Phospholipase A2 Enzymes: Structure, Function and Mechanism* (Kini,R.M., ed), pp. 1-28, John Wiley & Sons

38 Bonfim,V.L. *et al.* (2009) Toxicity of phospholipases A2 D49 (6-1 and 6-2) and K49 (Bj-VII) from Bothrops jararacussu venom. *Cell Biol. Toxicol.* 25, 523-532

39 de Paula,R.C. *et al.* (2009) Structural and pharmacological features of phospholipases A2 from snake venoms. *Protein Pept. Lett.* 16, 899-907

40 Benishin,C.G. (1990) Potassium channel blockade by the B subunit of beta-bungarotoxin. *Mol. Pharmacol.* 38, 164-169

41 Hite,L.A. *et al.* (1994) cDNA sequences for four snake venom metalloproteinases: structure, classification, and their relationship to mammalian reproductive proteins. *Arch. Biochem. Biophys.* 308, 182-191

42 Pawlak,J. *et al.* (2009) Irditoxin, a novel covalently linked heterodimeric three-finger toxin with high taxon-specific neurotoxicity. *FASEB J.* 23, 534-545

43 Jiang,M. *et al.* (1987) Isolation and pharmacological characterization of a new alpha-neurotoxin (alpha-AgTx) from venom of the viper Agkistrodon halys (Pallas). *Toxicon* 25, 1019-1022

44 Pahari,S. *et al.* (2007) The venom gland transcriptome of the Desert Massasauga Rattlesnake (Sistrurus catenatus edwardsii): towards an understanding of venom composition among advanced snakes (Superfamily Colubroidea). *BMC. Mol. Biol.* 8, 115

45 Menez,A. (1998) Functional architectures of animal toxins: a clue to drug design? *Toxicon* 36, 1557-1572

46 Tsetlin,V. (1999) Snake venom alpha-neurotoxins and other 'three-finger' proteins. *Eur. J. Biochem.* 264, 281-286

47 Kini,R.M. (2002) Molecular moulds with multiple missions: functional sites in three-finger toxins. *Clin. Exp. Pharmacol. Physiol* 29, 815-822

48 Kini,R.M. and Doley,R. (2010) Structure, function and evolution of three-finger toxins - Mini proteins with multiple targets. *Toxicon*

49 Nirthanan,S. and Gwee,M.C. (2004) Three-finger alpha-neurotoxins and the nicotinic acetylcholine receptor, forty years on. *J. Pharmacol. Sci.* 94, 1-17

50 Chu,N.S. (2005) Contribution of a snake venom toxin to myasthenia gravis: the discovery of alpha-bungarotoxin in Taiwan. *J. Hist Neurosci.* 14, 138-148

51 Mulugeta,E. *et al.* (2003) Loss of muscarinic M4 receptors in hippocampus of Alzheimer patients. *Brain Res.* 960, 259-262

52 Sekine-Aizawa,Y. and Huganir,R.L. (2004) Imaging of receptor trafficking by using alpha-bungarotoxin-binding-site-tagged receptors. *Proc. Natl. Acad. Sci. U. S. A* 101, 17114-17119

53 Tsetlin,V. *et al.* (2007) Detection of alpha7 nicotinic acetylcholine receptors with the aid of antibodies and toxins. *Life Sci.* 80, 2202-2205

54 Reid,P.F. (2007) Alpha-cobratoxin as a possible therapy for multiple sclerosis: a review of the literature leading to its development for this application. *Crit Rev. Immunol.* 27, 291-302

55 Gilquin,B. *et al.* (1999) Conformational and functional variability supported by the BPTI fold: solution structure of the Ca2+ channel blocker calcicludine. *Proteins* 34, 520-532

56 Harvey,A.L. and Karlsson,E. (1980) Dendrotoxin from the venom of the green mamba, Dendroaspis angusticeps. A neurotoxin that enhances acetylcholine release at neuromuscular junction. *Naunyn Schmiedebergs Arch. Pharmacol.* 312, 1-6

57 Joubert,F.J. and Taljaard,N. (1979) Snake venoms. The amino-acid sequence of protein S2C4 from Dendroaspis jamesoni kaimosae (Jameson's mamba) venom. *Hoppe Seylers. Z. Physiol Chem.* 360, 571-580

58 Millers,E.K. *et al.* (2009) Crystal structure of textilinin-1, a Kunitz-type serine protease inhibitor from the venom of the Australian common brown snake (Pseudonaja textilis). *The FEBS journal* 276, 3163-3175

59 Torres,A.M. *et al.* (2003) Identification of a novel family of proteins in snake venoms. Purification and structural characterization of nawaprin from Naja nigricollis snake venom. *J. Biol. Chem.* 278, 40097-40104

60 Nair,D.G. *et al.* (2007) Antimicrobial activity of omwaprin, a new member of the waprin family of snake venom proteins. *Biochem. J.* 402, 93-104

61 Shikamoto,Y. *et al.* (2005) Crystal structure of a CRISP family Ca2+ -channel blocker derived from snake venom. *Journal of molecular biology* 350, 735-743

62 Peichoto,M.E. *et al.* (2009) Purification and characterization of a cysteine-rich secretory protein from Philodryas patagoniensis snake venom. *Comp Biochem. Physiol C. Toxicol. Pharmacol.* 150, 79-84

63 Clemetson,K.J. *et al.* (1998) Snake venom C-type lectins as tools in platelet research. *Platelets.* 9, 165-169

64 Clemetson,K.J. *et al.* (2001) Multifunctional snake C-type lectins affecting platelets. *Haemostasis* 31, 148-154

65 Hirabayashi,J. *et al.* (1991) Complete primary structure of a galactose-specific lectin from the venom of the rattlesnake Crotalus atrox. Homologies with Ca2(+)-dependent-type lectins. *J. Biol. Chem.* 266, 2320-2326

66 Lin,L.P. *et al.* (2007) Cloning, expression and characterization of two C-type lectins from the venom gland of Bungarus multicinctus. *Toxicon* 50, 411-419

67 Drickamer,K. (1993) Recognition of complex carbohydrates by Ca(2+)-dependent animal lectins. *Biochem. Soc Trans.* 21, 456-459

68 Morita,T. (2004) C-type lectin-related proteins from snake venoms. *Curr. Drug Targets. Cardiovasc. Haematol. Disord.* 4, 357-373

69 Gould,R.J. *et al.* (1990) Disintegrins: a family of integrin inhibitory proteins from viper venoms. *Proc. Soc. Exp. Biol. Med.* 195, 168-171

70 Lin,E. *et al.* (2010) The disintegrin contortrostatin in combination with docetaxel is a potent inhibitor of prostate cancer in vitro and in vivo. *Prostate* 70, 1359-1370

71 Markland,F.S., Jr. (1998) Snake venom fibrinogenolytic and fibrinolytic enzymes: an updated inventory. Registry of Exogenous Hemostatic Factors of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. *Thrombosis and haemostasis* 79, 668-674

72 Markland,F.S. (1998) Snake venoms and the hemostatic system. *Toxicon* 36, 1749-1800

73 Hartman,G.D. *et al.* (1992) Non-peptide fibrinogen receptor antagonists. 1. Discovery and design of exosite inhibitors. *J. Med. Chem.* 35, 4640-4642

74 Funk,C. *et al.* (1971) Reptilase-R--a new reagent in blood coagulation. *Br. J. Haematol.* 21, 43-52

75 Madaras,F. *et al.* (1981) Automated estimation of factor Xa using the chromogenic substrate S-2222. *Haemostasis* 10, 271-275

76 Cushman,D.W. *et al.* (1979) Development of specific inhibitors of angiotensin I converting enzyme (kininase II). *Fed. Proc* 38, 2778-2782

77 Pu,X.C. *et al.* (1995) A novel analgesic toxin (hannalgesin) from the venom of king cobra (Ophiophagus hannah). *Toxicon* 33, 1425-1431

78 Trikha,M. *et al.* (1994) Contortrostatin, a snake venom disintegrin, inhibits beta 1 integrin-mediated human metastatic melanoma cell adhesion and blocks experimental metastasis. *Cancer Res.* 54, 4993-4998

79 DeWys,W.D. *et al.* (1976) Effect of defibrination on tumor growth and response to chemotherapy. *Cancer Res.* 36, 3584-3587

80 Wood S Jr and Hilgard,P.H. (1973) Arvin-induced hypofibrinogenemia and metastasis formation from blood-borne cancer cells. *Johns. Hopkins. Med. J.* 133, 207-213

81 Eggertsen,G. *et al.* (1981) Molecular characterization of the complement activating protein in the venom of the Indian cobra (Naja N. siamensis). *Mol. Immunol.* 18, 125-133

82 Vogt,W. (1990) Snake venom constituents affecting the complement system. In *Medical Use of Snake Venom Proteins* (Stocker,K.F., ed), pp. 79-96, CRC Press

83 Vogel,C.-W. (1987) *Immunoconjugates. Antibody Conjugates in Radioimaging and Therapy of Cancer*, Oxford University Press

84 Stiles,B.G. *et al.* (1991) Antibacterial effects of different snake venoms: purification and characterization of antibacterial proteins from Pseudechis australis (Australian king brown or mulga snake) venom. *Toxicon* 29, 1129-1141

85 Costa Torres,A.F. *et al.* (2010) Antibacterial and antiparasitic effects of Bothrops marajoensis venom and its fractions: Phospholipase A2 and L-amino acid oxidase. *Toxicon* 55, 795-804

86 Petricevich,V.L. and Mendonca,R.Z. (2003) Inhibitory potential of Crotalus durissus terrificus venom on measles virus growth. *Toxicon* 42, 143-153

87 Fenard,D. *et al.* (1999) Secreted phospholipases A(2), a new class of HIV inhibitors that block virus entry into host cells. *J. Clin. Invest* 104, 611-618

88 Chu,N.S. (2005) Contribution of a snake venom toxin to myasthenia gravis: the discovery of alpha-bungarotoxin in Taiwan. *Journal of the history of the neurosciences* 14, 138-148

89 Vincent,A. (2002) Unravelling the pathogenesis of myasthenia gravis. *Nat. Rev. Immunol.* 2, 797-804

90 Zelanis,A. *et al.* (2010) Preliminary biochemical characterization of the venoms of five Colubridae species from Brazil. *Toxicon* 55, 666-669

91 Acosta,O. *et al.* (2003) Hemorrhagic activity of the Duvernoy's gland secretion of the xenodontine colubrid Philodryas patagoniensis from the north-east region of Argentina. *Toxicon* 41, 1007-1012

92 Acosta de,P.O. *et al.* (2003) Edematogenic and myotoxic activities of the Duvernoy's gland secretion of Philodryas olfersii from the north-east region of Argentina. *Biocell* 27, 363-370

93 Ompraba,G. *et al.* (2010) Identification of a Novel Family of Snake Venom Proteins Veficolins from Cerberus rynchops Using a Venom Gland Transcriptomics and Proteomics Approach. *J Proteome. Res.*

94 Junqueira-de-Azevedo,I.L. *et al.* (2006) Lachesis muta (Viperidae) cDNAs reveal diverging pit viper molecules and scaffolds typical of cobra (Elapidae) venoms: implications for snake toxin repertoire evolution. *Genetics* 173, 877-889

95 Mirtschin,P.J. *et al.* (2006) Venom yields from Australian and some other species of snakes. *Ecotoxicology.* 15, 531-538

96 Wuster,W. and Broadley,D.G. (2007) Get an eyeful of this: a new species of giant spitting cobra from eastern and north-eastern Africa (Squamata: Serpentes: Elapidae: Naja). *Zootaxa* 1532, 51-68

97 Mirtschin,P.J. *et al.* (2002) Influences on venom yield in Australian tigersnakes (Notechis scutatus) and brownsnakes (Pseudonaja textilis: Elapidae, Serpentes). *Toxicon* 40, 1581-1592

98 Kostiza,T. and Meier,J. (1996) Nerve growth factors from snake venoms: chemical properties, mode of action and biological significance. *Toxicon* 34, 787-806

99 Cohen,S. and Levi-Montalcini,R. (1956) A Nerve Growth-Stimulating Factor Isolated from Snake Venom. *Proceedings of the National Academy of Sciences of the United States of America* 42, 571-574

100 Callahan,B.L. *et al.* (2008) Modulation of mechanical and thermal nociceptive sensitivity in the laboratory mouse by behavioral state. *J. Pain* 9, 174-184

101 Brittijn,S.A. *et al.* (2009) Zebrafish development and regeneration: new tools for biomedical research. *Int. J Dev. Biol.* 53, 835-850

102 Vogt,A. *et al.* (2010) Development of automated imaging and analysis for zebrafish chemical screens. *J. Vis. Exp.*

103 Vogt,A. *et al.* (2009) Automated image-based phenotypic analysis in zebrafish embryos. *Dev. Dyn.* 238, 656-663

104 Reading,C.J. *et al.* (2010) Are snake populations in widespread decline? *Biology Letters*

105 Shine,R. *et al.* (2002) Reproductive isolating mechanisms between two sympatric sibling species of sea snakes. *Evolution Evolution* 56, 1655-1662

106 Kochva,E. (1987) The origin of snakes and evolution of the venom apparatus. *Toxicon* 25, 65-106

107 Vidal,N. (2002) Colubroid systematics: Evidence for an early appearance of the venom apparatus followed by extensive evolutionary tinkering. *Journal of Toxicology-Toxin Reviews* 21, 21-41

108 Kochva,E. (1978) Oral Glands of the Reptilia. In *Biology of the Reptilia* (Gans,C. and Gans,K.A., eds), Academic press

109 Jackson,K. (2003) The Evolution of Venom-delivery Systems in Snakes. *Zoological Journal of the Linnean Society* 137, 337-354

110 Kardong,K.V. (1979) Protovipers and the Evolution of Snake Fangs. *Evolution Evolution* 33, 433-443

111 Jackson,K. (2007) The evolution of venom-conducting fangs: insights from developmental biology. *Toxicon* 49, 975-981

112 Jackson,K. (1995) Evicence from tooth surface morphology for a posterior maxillary origin of the proteroglyph fang. *Amphibia - Reptilia* 16, 273-288

113 Kardong,K. (1980) Evolutionary Patterns in Advanced Snakes. *American Zoologist* 20, 269-282

114 Kardong,K. (1982) The Evolution of the Venom Apparatus in Snakes from Colubrids to Viperids & Elapids. *Memorias do Instituto Butantan* 46, 105-118

115 Cobourne,M.T. and Sharpe,P.T. (2005) Sonic hedgehog signaling and the developing tooth. *Current topics in developmental biology* 65, 255-287

116 Fraser,G.J. *et al.* (2006) Gene deployment for tooth replacement in the rainbow trout (Oncorhynchus mykiss): a developmental model for evolution of the osteichthyan dentition. *Evol. Dev.* 8, 446-457

117 Fraser,G.J. *et al.* (2006) Developmental and evolutionary origins of the vertebrate dentition: molecular controls for spatio-temporal organisation of tooth sites in osteichthyans. *J. Exp. Zool. B Mol. Dev. Evol.* 306, 183-203

118 Yamanaka,A. *et al.* (2007) Development of heterodont dentition in house shrew (Suncus murinus) 1. *Eur J Oral Sci* 115, 433-440

119 Kuch,U. *et al.* (2006) Snake fangs from the Lower Miocene of Germany: evolutionary stability of perfect weapons. *Die Naturwissenschaften* 93, 84-87

120 Young,B. (1996) Dentitional surface features in snakes (Reptilia: Serpentes). *Amphibia - Reptilia* 17, 261-276

121 Frazzetta,T.H. (1966) Studies on the morphology and function of the skull in the Boidae (Serpentes). II. Morphology and function of the jaw apparatus in Phython sebae and Python molurus. *J. Morphol.* 118, 217-295

122 Bogert,C.M. (1943) Dentitional phenomena in cobras and other elapids with notes on adaptive modifications of fangs. *Bulletin American Museum of Natural History* 285-357

123 Cope,E.D. (1900). In *US National Museum Annual Report for 1898*, US Govt Print. Off.

124 Boulenger,G.A. (1896) Remarks on the Dentition of Snakes and on the Evolution of the Poison-fangs. *Proceedings of the Zoological Society of London*, 614-616

125 Anthony,J. (1955) Essai sur l'evolution anatomique de l'appareil venimeux des Ophidiens. *Ann. Sci. Nat. Zool. Biol. Anim.* II, 7-53

126 Shayer-Wollberg,M. and Kochva,E. (1967) Embryonic development of the venom apparatus in *Causus rhombeatus* (Viperidae, Ophidia). *herpetologica* 23, 249-259

127 Kochva,E. (1963) Development of the venom gland and trigeminal muscles in Vipera palaestinae. *Acta Anat.* 52, 49-89

128 Ovadia,M. (1984) Embryonic Development of Duvernoy's Gland in the Snake Natrix tesselata (Colubridae). *Copeia* 2, 516-521

129 Gygax,P. (1971) Entwicklung, bau und funktion der giftdrusse (Duvernoy's gland) von *Natrix tesselata*. *Acta Trop.* 28, 226-274

130 Ovadia,M. (1985) Embryonic Development of Duvernoy's Gland in the Snake Spalerosophis cliffordi (Colubridae). *Copeia* 1, 101-106

131 Buchtova,M. *et al.* (2007) Embryonic development of Python sebae - II: Craniofacial microscopic anatomy, cell proliferation and apoptosis. *Zoology. (Jena)* 110, 231-251

132 Fry,B.G. *et al.* (2008) Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol. Cell Proteomics.* 7, 215-246

133 Reshef,R. (1995) Ontogenetic mechanisms of teeth and oral gland development in snake - a possible contribution to the understanding of evolutionary processes in reptiles, Tel Aviv University

134 Reshef,R. *et al.* (1995) Snake yolk sac as a site for in vivo organ incubation: a new method in the research of snake embryo development. *J. Exp. Zool.* 270, 538-546

135 Pearson,J.C. *et al.* (2005) Modulating Hox gene functions during animal body patterning. *Nat. Rev. Genet.* 6, 893-904

136 Kessel,M. and Gruss,P. (1991) Homeotic transformations of murine vertebrae and concomitant alteration of Hox codes induced by retinoic acid. *Cell* 67, 89-104

137 Duboule,D. and Morata,G. (1994) Colinearity and functional hierarchy among genes of the homeotic complexes. *Trends Genet.* 10, 358-364

138 Burke,A.C. *et al.* (1995) Hox genes and the evolution of vertebrate axial morphology. *Development (Cambridge, England)* 121, 333-346

139 Gaunt,S.J. (1994) Conservation in the Hox code during morphological evolution. *Int. J. Dev. Biol.* 38, 549-552

140 Caroll,R.L. *et al.* (2001) *From DNA to Diversity*, Blackwell Science, London

141 Caldwell,M.W. (2003) "Without a leg to stand on": on the evolution and development of axial elongation and limblessness in tetrapods. (40 edn) pp. 573-588

142 Caroll,R.L. (2003) *Vertebrate Paleontology and Evolution.*

143 Gomez,C. *et al.* (2008) Control of segment number in vertebrate embryos. *Nature* 454, 335-339

144 Vonk,F.J. and Richardson,M.K. (2008) Developmental biology: Serpent clocks tick faster. *Nature* 454, 282-283

145 Cohn,M.J. and Tickle,C. (1999) Developmental basis of limblessness and axial patterning in snakes. *Nature* 399, 474-479

146 McIntyre,D.C. *et al.* (2007) Hox patterning of the vertebrate rib cage. *Development (Cambridge, England)* 134, 2981-2989

147 Chen,F. *et al.* (1998) Analysis of Hoxa7/Hoxb7 mutants suggests periodicity in the generation of the different sets of vertebrae. *Mech. Dev.* 77, 49-57

148 van den Akker,E. *et al.* (2001) Axial skeletal patterning in mice lacking all paralogous group 8 Hox genes. *Development (Cambridge, England)* 128, 1911-1921

149 Wellik,D.M. and Capecchi,M.R. (2003) Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. *Science (New York, N. Y* 301, 363-367

150 Carapuco,M. *et al.* (2005) Hox genes specify vertebral types in the presomitic mesoderm. *Genes Dev.* 19, 2116-2121

151 Hostikka,S.L. and Capecchi,M.R. (1998) The mouse Hoxc11 gene: genomic structure and expression pattern. *Mech. Dev.* 70, 133-145

152 Becker,D. *et al.* (1996) Conserved regulatory element involved in the early onset of Hoxb6 gene expression. *Dev. Dyn.* 205, 73-81

153 Toth,L.E. *et al.* (1987) Region-specific expression of mouse homeobox genes in the embryonic mesoderm and central nervous system. *Proc. Natl. Acad. Sci. U. S. A* 84, 6790-6794

154 Puschel,A.W. *et al.* (1990) Position-specific activity of the Hox1.1 promoter in transgenic mice. *Development (Cambridge, England)* 108, 435-442

155 Godwin,A.R. and Capecchi,M.R. (1998) Hoxc13 mutant mice lack external hair. *Genes Dev.* 12, 11-20

156 Chen,F. and Capecchi,M.R. (1997) Targeted mutations in hoxa-9 and hoxb-9 reveal synergistic interactions. *Dev. Biol.* 181, 186-196

157 Cohn,M.J. *et al.* (1997) Hox9 genes and vertebrate limb specification. *Nature* 387, 97-101

158 Rallis,C. *et al.* (2003) Tbx5 is required for forelimb bud formation and continued outgrowth. *Development (Cambridge, England)* 130, 2741-2751

159 Gibson-Brown,J.J. *et al.* (1998) Involvement of T-box genes Tbx2-Tbx5 in vertebrate limb specification and development. *Development (Cambridge, England)* 125, 2499-2509

160 Greene,H.W. (1997) *Snakes : the evolution of mystery in nature*, University of California Press

161 Mogil,J.S. *et al.* (2006) Screening for pain phenotypes: analysis of three congenic mouse strains on a battery of nine nociceptive assays. *Pain* 126, 24-34

162 Hoekstra,H.E. and Coyne,J.A. (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution Evolution* 61, 995-1016

163 Wagner,G.P. and Lynch,V.J. (2008) The gene regulatory logic of transcription factor evolution. *Trends Ecol. Evol.* 23, 377-385

164 Galant,R. and Carroll,S.B. (2002) Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* 415, 910-913

165 Ronshaugen,M. *et al.* (2002) Hox protein mutation and macroevolution of the insect body plan. *Nature* 415, 914-917

166 Vonk,F.J. *et al.* (2011) Snake venom: From fieldwork to the clinic: Recent insights into snake biology, together with new technology allowing high-throughput screening of venom, bring new hope for drug discovery. *Bioessays* 33, 269-279

167 Zupunski,V. *et al.* (2003) Adaptive evolution in the snake venom Kunitz/BPTI protein family. *FEBS letters* 547, 131-136

168 Kordis,D. and Gubensek,F. (2000) Adaptive evolution of animal toxin multigene families. *Gene* 261, 43-52

169 Reza,M.A. *et al.* (2007) Structure of two genes encoding parallel prothrombin activators in Tropidechis carinatus snake: gene duplication and recruitment of factor X gene to the venom gland. *J Thromb Haemost* 5, 117-126

170 Fry,B.G. (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 research* 15, 403-420

171 Kwong,S. *et al.* (2009) The recruitment of blood coagulation factor X into snake venom gland as a toxin: the role of promoter cis-elements in its expression. *Thromb. Haemost.* 102, 469-478

172 Minh Le,T.N. *et al.* (2005) Gene duplication of coagulation factor V and origin of venom prothrombin activator in Pseudonaja textilis snake. *Thrombosis and haemostasis* 93, 420-429

173 Cousin,X. *et al.* (1998) Identification of a novel type of alternatively spliced exon from the acetylcholinesterase gene of Bungarus fasciatus. Molecular forms of acetylcholinesterase in the snake liver and muscle. *J. Biol. Chem.* 273, 9812-9820

174 Vidal,N. *et al.* (2007) The phylogeny and classification of caenophidian snakes inferred from seven nuclear protein-coding genes. *C. R. Biol* 330, 182-187

175 Fry,B.G. *et al.* (2006) Early evolution of the venom system in lizards and snakes. *Nature* 439, 584-588

176 Vonk,F.J. *et al.* (2011) Snake venom: From fieldwork to the clinic. *Bioessays* 33, 269-279

177 Reza,M.A. *et al.* (2005) Gene structures of trocarin D and coagulation factor X, two functionally diverse prothrombin activators from Australian rough scaled snake. *Pathophysiology of haemostasis and thrombosis* 34, 205-208

178 Stanke,M. *et al.* (2006) Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC. Bioinformatics.* 7, 62

179 Warren,W.C. *et al.* (2008) Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 453, 175-183

180 Li,R. *et al.* (2010) The sequence and de novo assembly of the giant panda genome. *Nature* 463, 311-317

181 Wallis,J.W. *et al.* (2004) A physical map of the chicken genome. *Nature* 432, 761-764

182 Kini,R.M. and Doley,R. (2010) Structure, function and evolution of three-finger toxins - Mini proteins with multiple targets. *Toxicon*

183 R.Manjunatha Kini (1997) *Venom Phospholipase A2 Enzymes*, John Wiley & Sons, Inc.

184 Casewell,N.R. *et al.* (2009) Comparative venom gland transcriptome surveys of the saw-scaled vipers (Viperidae: Echis) reveal substantial intra-family gene diversity and novel venom transcripts. *BMC. Genomics* 10, 564

185 Gibbs,G.M. and O'Bryan,M.K. (2007) Cysteine rich secretory proteins in reproduction and venom. *Soc. Reprod. Fertil. Suppl* 65, 261-267

186 Moura-da-Silva,A.M. *et al.* (2007) Importance of snake venom metalloproteinases in cell biology: effects on platelets, inflammatory and endothelial cells. *Curr. Pharm. Des* 13, 2893-2905

187 Du,X.Y. and Clemetson,K.J. (2002) Snake venom L-amino acid oxidases. *Toxicon* 40, 659-665

188 Kemparaju,K. and Girish,K.S. (2006) Snake venom hyaluronidase: a therapeutic target. *Cell Biochem. Funct.* 24, 7-12

189 Reitinger,S. *et al.* (2008) High-yield recombinant expression of the extremophile enzyme, bee hyaluronidase in Pichia pastoris. *Protein Expr. Purif.* 57, 226-233

190 Chatrath,S.T. *et al.* (2011) Identification of novel proteins from the venom of a cryptic snake Drysdalia coronoides by a combined transcriptomics and proteomics approach. *J. Proteome. Res.* 10, 739-750

191 Li,J.W. and Vederas,J.C. (2009) Drug discovery and natural products: end of an era or an endless frontier? *Science (New York, N. Y* 325, 161-165

192 Deshimaru,M. *et al.* (1996) Accelerated evolution of crotalinae snake venom gland serine proteases. *FEBS letters* 397, 83-88

193 Calvete,J.J. *et al.* (2007) Snake Venomics of Bitis Species Reveals Large Intragenus Venom Toxin Composition Variation: Application to Taxonomy of Congeneric Taxa. *Journal of proteome research*

194 Doley,R. *et al.* (2008) Accelerated exchange of exon segments in Viperid three-finger toxin genes (Sistrurus catenatus edwardsii; Desert Massasauga). *BMC. Evol. Biol.* 8, 196

195 Doley,R. *et al.* (2008) Unusual accelerated rate of deletions and insertions in toxin genes in the venom glands of the pygmy copperhead (Austrelaps labialis) from Kangaroo island. *BMC. Evol. Biol.* 8, 70

196 Harvey,A.L. (2008) Natural products in drug discovery. *Drug Discov. Today* 13, 894-901

197 Kapoor,V.K. (2010) Natural toxins and their therapeutic potential. *Indian J. Exp. Biol.* 48, 228-237

198 Hamidi,M. *et al.* (2007) Applications of carrier erythrocytes in delivery of biopharmaceuticals. *J. Control Release* 118, 145-160

199 Bailon,P. and Won,C.Y. (2009) PEG-modified biopharmaceuticals. *Expert. Opin. Drug Deliv.* 6, 1-16

200 Takahashi,H. *et al.* (2007) Delivery of large biopharmaceuticals from cardiovascular stents: a review. *Biomacromolecules.* 8, 3281-3293

201 Dinarvand,R. *et al.* (2004) Polymeric delivery systems for biopharmaceuticals. *Biotechnol. Genet. Eng Rev.* 21, 147-181

202 Decker,E.L. and Reski,R. (2008) Current achievements in the production of complex biopharmaceuticals with moss bioreactors. *Bioprocess. Biosyst. Eng* 31, 3-9

203 Wurm,F.M. (2007) Manufacturing of biopharmaceuticals and implications for biosimilars. *Kidney Blood Press Res.* 30 Suppl 1, 6-8

204 Komarova,T.V. *et al.* (2010) Transient expression systems for plant-derived biopharmaceuticals. *Expert. Rev. Vaccines.* 9, 859-876

205 Daniell,H. *et al.* (2009) Plant-made vaccine antigens and biopharmaceuticals. *Trends Plant Sci.* 14, 669-679

206 Kessler,M. *et al.* (2006) Immunogenicity of biopharmaceuticals. *Nephrol. Dial. Transplant.* 21 Suppl 5, v9-12

207 Bhogal,N. (2010) Immunotoxicity and immunogenicity of biopharmaceuticals: design concepts and safety assessment. *Curr. Drug Saf* 5, 293-307

208 Bussiere,J.L. *et al.* (2009) Alternative strategies for toxicity testing of species-specific biopharmaceuticals. *Int. J. Toxicol.* 28, 230-253

209 Giezen,T.J. *et al.* (2009) Pharmacovigilance of biopharmaceuticals: challenges remain. *Drug Saf* 32, 811-817

210 Reichert,J.M. (2000) New biopharmaceuticals in the USA: trends in development and marketing approvals 1995-1999. *Trends Biotechnol.* 18, 364-369

211 Reichert,J.M. and Healy,E.M. (2001) Biopharmaceuticals approved in the EU 1995-1999: a European Union-United States comparison. *Eur. J. Pharm. Biopharm.* 51, 1-7

212 Ponzio,T.A. *et al.* (2011) License Compliance Issues For Biopharmaceuticals: Special Challenges For Negotiations Between Companies And Non-Profit Research Institutions. *LES Nouv.* 46, 216-225

213 Ratner,M. (2010) Pfizer stakes a claim in plant cell-made biopharmaceuticals. *Nat. Biotechnol.* 28, 107-108

214 Wallace,M.S. (2006) Ziconotide: a new nonopioid intrathecal analgesic for the treatment of chronic pain. *Expert. Rev. Neurother.* 6, 1423-1428

215 Gallwitz,B. (2006) Exenatide in type 2 diabetes: treatment effects in clinical studies and animal study data. *Int. J. Clin. Pract.* 60, 1654-1661

216 van Os,G.M. *et al.* (2011) Detection of lupus anticoagulant in the presence of rivaroxaban using Taipan snake venom time. *J. Thromb. Haemost.* 9, 1657-1659

217 Magazanik,L.G. and Vyskocyl,F. (1972) [Mechanism of action of snake venom neurotoxins]. *Zhurnal evoliutsionnoi biokhimii i fiziologii* 8, 555-557

218 Koehn,F.E. and Carter,G.T. (2005) The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.* 4, 206-220

219 Brack,W. (2003) Effect-directed analysis: a promising tool for the identification of organic toxicants in complex mixtures? *Anal. Bioanal. Chem.* 377, 397-407

220 Brack,W. *et al.* (2007) Effect-directed analysis of key toxicants in European river basins a review. *Environ. Sci. Pollut. Res. Int.* 14, 30-38

221 Houtman,C.J. *et al.* (2007) Biomonitoring of estrogenic exposure and identification of responsible compounds in bream from Dutch surface waters. *Environ. Toxicol. Chem.* 26, 898-907

222 Yu,S. *et al.* (2010) Screening and isolation of a natural dopamine D1 receptor antagonist using cell-based assays. *J. Biotechnol.* 145, 304-309

223 Nyila,M.A. *et al.* (2009) Bioactivities of Plectranthus ecklonii constituents. *Nat. Prod. Commun.* 4, 1177-1180

224 Ebada,S.S. *et al.* (2008) Methods for isolation, purification and structural elucidation of bioactive secondary metabolites from marine invertebrates. *Nat. Protoc.* 3, 1820-1831

225 Fernandez,J. *et al.* (2011) Venomic and antivenomic analyses of the Central American coral snake, Micrurus nigrocinctus (Elapidae). *J. Proteome. Res.* 10, 1816-1827

226 Hill,R.E. and Mackessy,S.P. (1997) Venom yields from several species of colubrid snakes and differential effects of ketamine. *Toxicon* 35, 671-678

227 Fry,B.G. *et al.* (2003) Molecular evolution and phylogeny of elapid snake venom three-finger toxins. *Journal of molecular evolution* 57, 110-129

228 Schebb,N.H. *et al.* (2009) Analysis of glutathione adducts of patulin by means of liquid chromatography (HPLC) with biochemical detection (BCD) and electrospray ionization tandem mass spectrometry (ESI-MS/MS). *Anal. Bioanal. Chem.* 394, 1361-1373

229 Kool,J. *et al.* (2007) Online biochemical detection of glutathione-S-transferase P1-specific inhibitors in complex mixtures. *J. Biomol. Screen.* 12, 396-405

230 Kool,J. *et al.* (2007) Cytochrome P450 bio-affinity detection coupled to gradient HPLC: on-line screening of affinities to cytochrome P4501A2 and 2D6. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 858, 49-58

231 Kool,J. *et al.* (2007) An on-line post-column detection system for the detection of reactive-oxygen-species-producing compounds and antioxidants in mixtures. *Anal. Bioanal. Chem.* 388, 871-879

232 Kool,J. *et al.* (2011) Advances in mass spectrometry-based post-column bioaffinity profiling of mixtures. *Anal. Bioanal. Chem.* 399, 2655-2668

233 de Jong,C.F. *et al.* (2006) High-performance liquid chromatography-mass spectrometry-based acetylcholinesterase assay for the screening of inhibitors in natural extracts. *J. Chromatogr. A* 1112, 303-310

234 van Elswijk,D.A. *et al.* (2004) Rapid dereplication of estrogenic compounds in pomegranate (Punica granatum) using on-line biochemical detection coupled to mass spectrometry. *Phytochemistry* 65, 233-241

235 Schobel,U. *et al.* (2001) High resolution screening of plant natural product extracts for estrogen receptor alpha and beta binding activity using an online HPLC-MS biochemical detection system. *J. Biomol. Screen.* 6, 291-303

236 Ingkaninan,K. *et al.* (2000) The application of HPLC with on-line coupled UV/MS-biochemical detection for isolation of an acetylcholinesterase inhibitor from narcissus 'Sir Winston Churchill'. *J. Nat. Prod.* 63, 803-806

237 Heus,F. *et al.* (2010) Development of a microfluidic confocal fluorescence detection system for the hyphenation of nano-LC to on-line biochemical assays. *Anal. Bioanal. Chem.* 398, 3023-3032

238 Dutertre,S. *et al.* (2007) AChBP-targeted alpha-conotoxin correlates distinct binding orientations with nAChR subtype selectivity. *EMBO J.* 26, 3858-3867

239 Smit,A.B. *et al.* (2006) Acetylcholine-binding proteins: functional and structural homologs of nicotinic acetylcholine receptors. *J. Mol. Neurosci.* 30, 9-10

240 Smit,A.B. *et al.* (2001) A glia-derived acetylcholine-binding protein that modulates synaptic transmission. *Nature* 411, 261-268

241 Edink,E. *et al.* (2011) Fragment growing induces conformational changes in acetylcholine-binding protein: a structural and thermodynamic analysis. *J. Am. Chem. Soc.* 133, 5363-5371

242 de Kloe,G.E. *et al.* (2010) Surface plasmon resonance biosensor based fragment screening using acetylcholine binding protein identifies ligand efficiency hot spots (LE hot spots) by deconstruction of nicotinic acetylcholine receptor alpha7 ligands. *J. Med. Chem.* 53, 7192-7201

243 Kool,J. *et al.* (2010) Online fluorescence enhancement assay for the acetylcholine binding protein with parallel mass spectrometric identification. *J. Med. Chem.* 53, 4720-4730

244 Celie,P.H. *et al.* (2004) Nicotine and carbamylcholine binding to nicotinic acetylcholine receptors as studied in AChBP crystal structures. *Neuron* 41, 907-914

245 Chang,L. *et al.* (1999) Genetic organization of alpha-bungarotoxins from Bungarus multicinctus (Taiwan banded krait): evidence showing that the production of alpha-bungarotoxin isotoxins is not derived from edited mRNAs. *Nucleic Acids Res.* 27, 3970-3975

246 Fiordalisi,J.J. and Grant,G.A. (1993) Evidence for a fast-exchange conformational process in alpha-bungarotoxin. *Toxicon* 31, 767-775

247 McDowell,R.S. *et al.* (1992) Mambin, a potent glycoprotein IIb-IIIa antagonist and platelet aggregation inhibitor structurally related to the short neurotoxins. *Biochemistry* 31, 4766-4772

248 Marchot,P. *et al.* (1997) Expression and activity of mutants of fasciculin, a peptidic acetylcholinesterase inhibitor from mamba venom. *J. Biol. Chem.* 272, 3502-3510

249 Marchot,P. *et al.* (1998) Inhibition of mouse acetylcholinesterase by fasciculin: crystal structure of the complex and mutagenesis of fasciculin. *Toxicon* 36, 1613-1622

250 jas-Bailador,F. *et al.* (1998) Effects of alpha-erabutoxin, alpha-bungarotoxin, alpha-cobratoxin and fasciculin on the nicotine-evoked release of dopamine in the rat striatum in vivo. *Neurochem. Int.* 33, 307-312

251 Carsi,J.M. *et al.* (1999) m2-toxin: A selective ligand for M2 muscarinic receptors. *Mol. Pharmacol.* 56, 933-937

252 Carsi,J.M. and Potter,L.T. (2000) m1-toxin isotoxins from the green mamba (Dendroaspis angusticeps) that selectively block m1 muscarinic receptors. *Toxicon* 38, 187-198

253 Gong,N. *et al.* (1999) Postsynaptic short-chain neurotoxins from Pseudonaja textilis. cDNA cloning, expression and protein characterization. *Eur. J. Biochem.* 265, 982-989

254 Gong,N. *et al.* (2000) Molecular cloning, characterization and evolution of the gene encoding a new group of short-chain alpha-neurotoxins in an Australian elapid, Pseudonaja textilis. *FEBS Lett.* 473, 303-310

255 Chaki,S. *et al.* (1992) Purification and partial characterization of K+ channel blockers from the venom of Dendroaspis angusticeps. *Neurochem. Int.* 20, 553-558

256 Albrand,J.P. *et al.* (1995) NMR and restrained molecular dynamics study of the three-dimensional solution structure of toxin FS2, a specific blocker of the L-type calcium channel, isolated from black mamba venom. *Biochemistry* 34, 5923-5937

257 Kini,R.M. *et al.* (1998) Flanking proline residues identify the L-type Ca2+ channel binding site of calciseptine and FS2. *Biochemistry* 37, 9058-9063

258 Jang,J.Y. *et al.* (1997) Comparison of the hemolytic activity and solution structures of two snake venom cardiotoxin analogues which only differ in their N-terminal amino acid. *Biochemistry* 36, 14635-14641

259 Fryklund,L. and Eaker,D. (1973) Complete amino acid sequence of a nonneurotoxic hemolytic protein from the venom of Haemachatus haemachates (African ringhals cobra). *Biochemistry* 12, 661-667

260 Choudhury,S.R. *et al.* (2006) Purification, crystallization and preliminary X-ray structural studies of a 7.2 kDa cytotoxin isolated from the venom of Daboia russelli russelli of the Viperidae family. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.* 62, 292-294

261 Dubovskii,P.V. *et al.* (2005) Interaction of three-finger toxins with phospholipid membranes: comparison of S- and P-type cytotoxins. *Biochem. J.* 387, 807-815

262 Ploug,M. and Ellis,V. (1994) Structure-function relationships in the receptor for urokinase-type plasminogen activator. Comparison to other members of the Ly-6 family and snake venom alpha-neurotoxins. *FEBS letters* 349, 163-168

263 Kornisiuk,E. *et al.* (1995) Binding of muscarinic toxins MTx1 and MTx2 from the venom of the green mamba Dendroaspis angusticeps to cloned human muscarinic cholinoceptors. *Toxicon* 33, 11-18

264 Jerusalinsky,D. *et al.* (1997) Workshop: the use of muscarinic toxins in the study of muscarinic receptors. *Life Sci.* 60, 1161-1162

265 Kanda,T. *et al.* (1983) Neuromuscular junctions of the posterior cricoarytenoid muscle in the cat. *Acta Otolaryngol. Suppl* 393, 25-32

266 Unwin,N. (1993) Nicotinic acetylcholine receptor at 9 A resolution. *J. Mol. Biol.* 229, 1101-1124

267 Pawlak,J. *et al.* (2006) Denmotoxin, a three-finger toxin from the colubrid snake Boiga dendrophila (Mangrove Catsnake) with bird-specific activity. *The Journal of biological chemistry* 281, 29030-29041

268 Kumar,T.K. *et al.* (1997) Snake venom cardiotoxins-structure, dynamics, function and folding. *J. Biomol. Struct. Dyn.* 15, 431-463

269 Konshina,A.G. *et al.* (2011) Snake cytotoxins bind to membranes via interactions with phosphatidylserine head groups of lipids. *PLoS. One.* 6, e19064

270 Dubovskii,P.V. *et al.* (2003) Interaction of the P-type cardiotoxin with phospholipid membranes. *Eur. J. Biochem.* 270, 2038-2046

271 Vulfius,C.A. *et al.* (2011) An unusual phospholipase A(2) from puff adder Bitis arietans venom--a novel blocker of nicotinic acetylcholine receptors. *Toxicon* 57, 787-793

272 Petras,D. *et al.* (2011) Snake venomics of African spitting cobras: toxin composition and assessment of congeneric cross-reactivity of the pan-African EchiTAb-Plus-ICP antivenom by antivenomics and neutralization approaches. *J. Proteome. Res.* 10, 1266-1280

273 Ackermann,E.J. and Taylor,P. (1997) Nonidentity of the alpha-neurotoxin binding sites on the nicotinic acetylcholine receptor revealed by modification in alpha-neurotoxin and receptor structures. *Biochemistry* 36, 12836-12844

274 Malany,S. *et al.* (2000) Orientation of alpha-neurotoxin at the subunit interfaces of the nicotinic acetylcholine receptor. *Biochemistry* 39, 15388-15398

275 Strydom,A.J. (1973) Snake venom toxins. The amino acid sequences of two toxins from Dendroaspis jamesoni kaimosae (Jameson's mamba) venom. *Biochimica et biophysica acta* 328, 491-509

276 Joubert,F.J. (1976) Snake venom toxins. The amino-acid sequences of three toxins (CM-8, CM-11 and CM-13a) from Naja haje annulifera (Egyptian cobra) venom. *European journal of biochemistry / FEBS* 64, 219-232

277 Joubert,F.J. (1975) The purification and amino acid sequence of toxin CM-13b from Naja haje annulifera (Egyptian cobra) venom. *Hoppe Seylers. Z. Physiol Chem.* 356, 1901-1908

278 Jeyaseelan,K. *et al.* (2003) Structurally conserved alpha-neurotoxin genes encode functionally diverse proteins in the venom of Naja sputatrix. *FEBS Lett.* 553, 333-341

279 Joubert,F. and Taljaard,N. (1978) Purification, some properties and the primary structures of three reduced and S-carboxymethylated toxins (CM-5, CM-6 and CM-10a) from Naje haje haje (Egyptian cobra) venom. *Biochim. Biophys. Acta* 537, 1-8

280 Pook,C.E. and McEwing,R. (2005) Mitochondrial DNA sequences from dried snake venom: a DNA barcoding approach to the identification of venom samples. *Toxicon* 46, 711-715

281 Lewis,R.J. and Garcia,M.L. (2003) Therapeutic potential of venom peptides. *Nat. Rev. Drug Discov.* 2, 790-802

282 Mouhat,S. *et al.* (2008) Animal toxins acting on voltage-gated potassium channels. *Curr. Pharm. Des* 14, 2503-2518

283 Dutertre,S. *et al.* (2006) Isolation and characterisation of conomap-Vt, a D-amino acid containing excitatory peptide from the venom of a vermivorous cone snail. *FEBS Lett.* 580, 3860-3866

284 Hansen,S.B. *et al.* (2002) Tryptophan fluorescence reveals conformational changes in the acetylcholine binding protein. *J. Biol. Chem.* 277, 41299-41302

285 Wielhouwer,E.M. *et al.* (2011) Zebrafish embryo development in a microfluidic flow-through system. *Lab Chip.* 11, 1815-1824

# CV Freek Vonk

Born: 24th of February 1983 (Dordrecht, The Netherlands).

### Voorbereidend Wetenschappelijk Onderwijs

Titus Brandsma College, Dordrecht, 1995-2001.

### Doctoraal biologie 2001-2008

University of Leiden, Faculty of Science; May 24, 2008; Grade average: 8,8 (out of 10 points).

### Work experience

*2012-*   Research Fellow, Naturalis Biodiversity Center

*2008 - 2012* PhD project (Leiden University), promotor Prof.  Dr. MK Richardson

*June 2010 -* Guest-researcher (Netherlands Centre for Biodiversity Naturalis)

*2008*   Doctoraal biologie (Leiden University)

*2006*   Internship (the Hubrecht laboratory, Utrecht)

*2006*   Internship (Australian Venom Research Unit, Melbourne University & Population and Evolutionary Genetics Unit, Museum Victoria)

*2002-2004* Guest-researcher (National Museum of Natural History Naturalis)

### Fieldwork

Led several research expeditions to collect rare and unique venom and genetic samples for the Leiden University compound libraries. These include two long (four month) expeditions to Australia (2009 and 2010; to Indonesia and Malaysia (2008); India (2011); Israel (2008); Zambia (2010); Tanzania (2010); and South-Africa (2011), and many other countries.

Freek has performed labwork in several foreign labs including The University of South-Australia (2005); the Australian Venom Research Unit at the University of Melbourne (2006); the Population and Evolutionary Genetics Unit, Museum Victoria (2006); and the University of Paris (2006).

**Academic activities**

Freek gave over 50 (mainly invited) conference talks and seminars all over the world. He reviewed scientific papers for the following journals: *Nature Communications; Applied Herpetology; Contributions to Zoology; Proc. R. Soc. B; European Journal of Biology; Journal of Experimental Zoology Part B: Molecular and Developmental Evolution;* and the *Journal of Molecular Evolution.* He is an editor of the *Journal of Venom Research* and has been asked to review scientific (research) grants for the Research Foundation Flanders (FWO), Belgium and the Sultan Qaboos University, Oman. Freek is a member of the International Society for Toxinology.

**Scholarships and prizes**

Apart from the Toptalent grant that supported this study, Freek received several grants from various sources, including the Naturalis Biodiversity center, the Leiden University Fund, the Gratama Stichting, the Curatoren fund, STW, the LUSTRA fund, the Schure-Beijerinck-Popping Fonds (KNAW), and a Rubicon grant from NWO, for a total of €410.000,-

**Peer-reviewed publications**

*Published*

- Vidal N, Marin J, Sassi J, Battistuzzi FU, Donnellan S, Fitch AJ, Fry BG, Vonk FJ, Rodriguez de la Vega RC, Couloux R, & SB Hedges. Molecular Evidence for an Asian Origin of Monitor Lizards Followed by Tertiary Dispersals to Africa and Australasia. *Biol Lett.* 2012 Jul 18
- Vonk FJ, Jackson K, Doley R, Madaras F, Mirtschin PJ & N Vidal. Snake Venom: from Fieldwork to the Clinic. *BioEssays* 33, 269-79 (2011).
- Siang AS, Doley R, Vonk FJ, & RM Kini. Transcriptomic Analysis of the Venom Gland of the Red-headed Krait (Bungarus flaviceps) Using Expressed Sequence Tags. *BMC Molecular Biology* 11, 24 - 34 (2010).
- Brittijn SA, Duivesteijn SJ, Belmamoune M, Bertens LFM, Bitter W, de Bruijn JD, Champagne D, Cuppen E, Flik G, Vandenbroucke-Grauls CM, Janssen RAJ, de Jong IM, de Kloet ER, Kros A, Meijer AH, Metz JR, van der Sar A, Schaaf MJM, Schulte-Merker S, Spaink HP, Tak PP, Verbeek FJ, Vervoordeldonk MJ, Vonk FJ, Witte F, Yuan H & MK Richardson. Zebrafish Development and Regeneration: New Tools for Biomedical Research. *Int. J. Dev. Biol.* 53, 835-850 (2009).
- Woltering JM, Vonk FJ, Muller H, Bardine N, Tuduce IL, de Bakker MAG, Knochel W, Sirbu IO, Durston AJ & MK Richardson. Axial Patterning in Snakes and Caecilians: Evidence for an Alternative Interpretation of the Hox Code. *Dev Biol.* 332, 82-9 (2009).
- Vonk FJ, Admiraal JF, Jackson K, Reshef R, de Bakker MAG, Vanderschoot K, van den Berge I, van Atten M, Burgerhout E, Beck A, Mirtschin PJ, Kochva E, Witte F, Fry BG, Woods A & MK Richardson. Evolutionary Origin and Development of Snake Fangs. *Nature* 454, 630-633 (2008).
- Vonk FJ & Richardson MK. Serpent Clocks Tick Faster. *Nature* 454, 282-283 (2008).
- Bertola LD, Ott EB, Griepsma S, Vonk FJ & CP Bagowski. Developmental Expression of the Alpha-Skeletal Actin Gene. *BMC Evol Biol.* 8, 166 (2008).
- Vonk FJ & Wüster W. Roles of CITES in Protecting New Species. *Science* 313, 915-916. (2006).

- Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan SF, Kuruppu S, Fung K, Hedges SB, Richardson MK, Hodgson WC, Ignjatovic V, Summerhayes V & E Kochva. Early Evolution of the Venom-System in Lizards and Snakes. *Nature* 439, 584-588 (2006).

*In preparation*

- Casewell N, Wüster W, Vonk FJ, Harrison R, & BG Fry. Complex Cocktails: the Evolutionary Novelty of Venoms. Invited review by *Tre nds in Ecology & Evolution*, under review.
- Heus F, Vonk FJ, Bruyneel B, Smit AB, Lingeman H, Richardson MK, Niessen W and J Kool. An Efficient Analytical Platform for On-line Microfluidic Profiling of Neurotoxic Snake Venoms Towards Nicotinic Receptor Like Affinity. Under review at *Toxicon*.
- Vonk FJ, Henkel CV, Casewell NV, Kini RM, Kerkkamp HM, Wuster W, Castoe TA, Ribeiro JMC, Spaink HP, Jansen HJ, Hyder SA, Arntzen JW, Pollock DD, van den Thillart GEEJ, Boetzer M, Pirovano W, Dirks RP and MK Richardson. Massive Evolutionary Expansion of Venom Genes in the King Cobra Genome. To be submitted late 2012.
- Castoe TA, de Koning AJP, Hall K, Yokoyama KY, Gu W, Smith EN, Feschotte C, Uetz P, Ray DA, Bogden R, Mackessy SP, Bronikowski AM, Vonk FJ, Henkel CV, Yandel M, Richardson MK, Warren WC, Secor S, and DD Pollock. The Burmese Python genome and the extreme adaptations of the trancriptome and proteome of snakes. To be submitted late 2012.

# Acknowledgements