toxins

MDPI

Review

Colubrid Venom Composition: An -Omics Perspective

Inácio L. M. Junqueira-de-Azevedo [1,*], Pollyanna F. Campos [1], Ana T. C. Ching [2] and Stephen P. Mackessy [3]

1 Laboratório Especial de Toxinologia Aplicada, Center of Toxins, Immune-Response and Cell Signaling (CeTICS), Instituto Butantan, São Paulo 05503-900, Brazil; pollyanna.fc@gmail.com
2 Laboratório de Imunoquímica, Instituto Butantan, São Paulo 05503-900, Brazil; anatung@gmail.com
3 School of Biological Sciences, University of Northern Colorado, Greeley, CO 80639-0017, USA; stephen.mackessy@unco.edu
* Correspondence: inacio.azevedo@butantan.gov.br; Tel.: +55-11-2627-9731

Academic Editor: Bryan Fry
Received: 7 June 2016; Accepted: 8 July 2016; Published: 23 July 2016

Abstract: Snake venoms have been subjected to increasingly sensitive analyses for well over 100 years, but most research has been restricted to front-fanged snakes, which actually represent a relatively small proportion of extant species of advanced snakes. Because rear-fanged snakes are a diverse and distinct radiation of the advanced snakes, understanding venom composition among "colubrids" is critical to understanding the evolution of venom among snakes. Here we review the state of knowledge concerning rear-fanged snake venom composition, emphasizing those toxins for which protein or transcript sequences are available. We have also added new transcriptome-based data on venoms of three species of rear-fanged snakes. Based on this compilation, it is apparent that several components, including cysteine-rich secretory proteins (CRiSPs), C-type lectins (CTLs), CTLs-like proteins and snake venom metalloproteinases (SVMPs), are broadly distributed among "colubrid" venoms, while others, notably three-finger toxins (3FTxs), appear nearly restricted to the Colubridae (sensu stricto). Some putative new toxins, such as snake venom matrix metalloproteinases, are in fact present in several colubrid venoms, while others are only transcribed, at lower levels. This work provides insights into the evolution of these toxin classes, but because only a small number of species have been explored, generalizations are still rather limited. It is likely that new venom protein families await discovery, particularly among those species with highly specialized diets.

Keywords: Colubridae; evolution; proteins; snake; toxins; transcriptomics

1. Introduction

More than one hundred years of biochemical and pharmacological studies have resulted in an exceptional depth of knowledge about snake venoms. The major toxins of the most medically important taxa of venomous snakes were determined by first generation approaches including protein chemistry, comparative pharmacology and cladistics methods borrowed from evolutionary biology. Advances in molecular biology, particularly protein and nucleic acid sequencing techniques, greatly expanded our understanding of compositional complexity, and more recent development in proteomics and early genomics greatly accelerated the pace of cataloguing venoms in exquisite detail. Recent next generation methods, including deep sequencing transcriptomics (RNAseq), genomic sequencing and high resolution mass spectrometry, including top-down proteomics, generally called venomics; (cf. [1–3]), have further accelerated the pace of sequence acquisition and compositional analysis and constituted the basis of large-scale biotechnological explorative initiatives (e.g., [4]). These studies collectively created very complete inventories of the toxin families and superfamilies present in

species representing significant risk to human health, further refined by a growing knowledge of the relative abundances, post-translational modifications and also structural conservation of proteins across numerous genera [5–18]. As a side product of the accumulation of this knowledge, the observed differences in venom composition among related taxa are becoming appreciated as a productive model for making evolutionary inferences about diversifying selection. In turn, the association of quantified toxins with empirically demonstrated activities have allowed predictions of functional and ecological roles for the species that produce these toxins (e.g., [19–22]).

However, because most of these efforts were driven by anthropocentric interests in understanding the mechanisms of actions of toxins causing severely debilitating effects, the major focus has been on the species of medical relevance, which are confined to only three families of modern snakes (Viperidae, Elapidae and Atractaspididae). As a consequence, a large part of the biodiversity of venom-producing snakes was not systematically evaluated in the same way, leaving a gap in our knowledge of the repertoire of toxins from other groups of venomous snakes, specifically those that do not typically result in serious human envenomations [23].

The advanced snakes (Caenophidea, superfamily Colubroidea) include a diverse assemblage of species with an evolutionary history of over 100 million years, most of which possess a venom production system [24–27]. The family "Colubridae" formerly referred to any caenophidian snake not included in the three medically important families of venomous snakes, and this assemblage, though acknowledged to be a paraphyletic group, resisted systematic consensus for many years [28–31]. More recently, several groups have reclassified the "Colubridae" into several families and subfamilies [32–35], but a true consensus classification is still lacking. In a more formal definition, a "colubrid" snake refers only to species belonging to the family Colubridae, which currently includes the subfamilies Natricinae, Pseudoxenodontinae, Dipsadinae, Scaphiodontophiinae, Calamariinae, Grayiinae and Colubrinae [34]. This family represents about 50% of the extant snake fauna (distributed in more than 1800 species), many with very distinct habits and diversification of species within each subfamily, and this classification scheme likely still masks considerable differentiation. Additional rear-fanged species, accounting for approximately 361 species, have been allocated to the families Homalopsidae and Lamprophiidae [34]; further, some authors consider Natricidae (approximately 226 species) and Dipsadidae (approximately 754 species) as distinct families [36,37]. Hereafter, we will use a broader definition of the term "colubrid" to refer to any of the families in the above paraphyletic group and not only to the family Colubridae sensu stricto, though the vast majority of data discussed here come from this family.

In spite of the uncertainties in phylogenies, snakes in these families often possess one or more enlarged rear teeth (opisthoglyph dentition) that are typically associated with a pair of Duvernoy's venom glands, homologs of the venom glands of families Viperidae, Elapidae and Atractaspididae [38,39]. Several species show no specialized or enlarged rear fangs (aglyph dentition), though in some cases they also contain other specialized oral glands that produce venom-like secretions [40].

Colubrids are rarely investigated using -omics approaches mainly because of their limited capacity to inject a debilitating dose of venom into humans, and so the biological activity of most species' venoms is wholly unknown. As are all snakes, they are predators, and venom is presumed to be of critical importance for capturing, killing and/or digesting the prey [41]. Thus, their venoms are expected to be highly efficient within the proper ecological context of each species, meaning that their venoms could be as rich and diverse in protein types as those from medically important species. Moreover, because the different colubrid venoms are utilized in very distinct ecological scenarios and evolved under different selective pressures, they may contain cryptic novel and unpredictable types of proteins.

Because so few studies have focused on colubrids venoms and a plethora of different methodological approaches were used by different labs, it is not clear which types of toxins are currently known in the various groups, what structural characteristics are known and what their evolutionary history has been. Some of the toxin sequences were obtained through direct

protein purification/sequencing, while others were deduced from transcriptomic and/or proteomic investigations. In addition to the different times when the investigations were performed, their specific goals sometimes hindered the perception of unusual new toxins. As a consequence, this has produced a distorted view of toxin repertoires that exist in colubrid venoms and hinders a more complete reconstruction of the evolutionary history of venom protein classes. The somewhat myopic view of venoms as occurring only in front-fanged snake species has interfered with a more holistic, fundamental perspective of the processes underpinning the evolution of venom, restricting the use of these exceptionally diversified animals as models for testing adaptive evolution by natural selection and negatively impacting the discovery of new bioactive molecules.

Here we survey previously discovered and several new venom proteins from venoms of colubrid species, focusing on those with known protein or cDNA-based sequences. Our intent is to provide an up-to-date catalog of proteins known to occur in colubrid snake venoms and present these in an evolutionary context, highlighting their (presently known) diversification. There are well over 2200 species of non-front-fanged snakes, many of which possess a Duvernoy's venom gland, so it should be immediately apparent that there is much work to do before a well-documented understanding of venom diversity among colubrids is possible. Nonetheless, by summarizing known information, we hope that this report will stimulate further investigation of the many genera of colubrid snakes for which we have no toxinological information.

2. Results and Discussion

2.1. Compiling the Venom Components of Colubrid Snakes

Our attempt to compile the toxins present in colubrids was based on three strategies: (1) generating transcriptomic sequences from the venom glands of three species of colubrids, *Erythrolamprus miliaris*, *Oxyrhopus guibei* and *Xenodon merremi* (Dipsadinae subfamily of Colubridae), to identify transcripts coding for known and putative types of snake toxins (Table S1); (2) prospecting public databases for toxin-related sequences in other colubrid species previously investigated; and (3) reviewing the literature on colubrid venoms that describes the isolation of toxins or provides clear evidence for the occurrence of specific proteins in colubrid venoms. For ease of presentation, the protein types compiled were organized into three categories: (a) "major snake venom components" (Table 1), referring to protein types generally encountered in high amounts in the venoms of many species of traditionally venomous snakes (Viperidae, Elapidae and Atractaspididae) and which certainly are important toxins; (b) "minor (or arguably) venom components" (Table 2), referring to protein types previously described in the venom of some species of venomous snakes, generally in low amounts, and which may represent toxins, ancillary venom proteins or housekeeping proteins; and (c) "putative new snake toxins in colubrid venoms" (Table 3), referring to protein types uncovered from colubrid venom analyses, occurring in high or low quantities, which may represent putative toxins, exclusive or not to the group. We should emphasize that the separation into major and minor components is unrelated to the level of expression (or protein quantity) of the components in colubrid venoms. Rather, it is related to a relative importance and frequency of the proteins in venoms of other venomous snakes. This organization is admittedly subjective and flexible, but it was adopted because it would be unrealistic to propose a division based on more tangible (but highly diverse) measures provided by the varied methodologies adopted in the studies reviewed. Because it reflects a particular point of view, it does not aim to establish a strict rule for toxin categorization or to define whether certain venom proteins do or do not have relevant functions in snake venoms. Additionally, because the strict definition of "toxin" would be dependent on the functional, ecological and behavioral contexts of the species, which are largely unavailable for colubrids, the protein types included here should be generally regarded as "venom components", which in some cases are very likely to be toxins and in other cases may or may not be toxins. The approximate phylogenetic relationships among the species for which venom components could be identified in our compilation are depicted in trees (Figure 1) based on the phylogenetic hypothesis of Colubroidea snakes as proposed by Pyron et al. [34].

Table 1. Major snake venom components and their occurrences in colubrid species.

Species	Enzymatic				Non-Enzymatic							Reference
	LAAO	PLA2 (IA)	SVMP	SVSP	3FTx	CNP	CRISP	CTL	DEFEN	KUN-1	KUN-2	
Boiga dendrophila					B							[42,43]
Boiga irregularis			TPB		TPB	T	TP	T			t	[12,44,45]
Borikenophis portoricensis			B				BP					[46,47]
Cerberus rynchops			TP				TP	TP				[48]
Coelognathus radiatus					B							[49]
Dispholidus typus			xP		x		x					[50–52]
Erythrolamprus miliaris			T		t		T	T				This work; [50]
Erythrolamprus poecilogyrus			x		x		x	x				[50,51]
Helicops angulatus							BP					[53]
Hypsiglena sp.			TP		T	T	TP	TP			t	[13]
Hypsiglena torquata							P					[54]
Leoiheterodon madagascarensis					x		x	x				[50]
Macropisthodon rudis			t									[55]
Opheodrys aestivus					x		t	t			t	[9]
Oxybelis fulgidus					B							[56]
Oxyrhopus guibei			T			t	T	T			t	This work
Phalotris mertensi	TP		T	tP	t	t	t	T	TP	TP		[57]
Pantherophis guttatus			t		x		t	t			t	[9]
Philodryas baroni							P					[54]
Philodryas chamissonis			x	x		x	x	x				[58]
Philodryas olfersii			xTP	xTP		T	xTP	TP		x		[51,59]
Philodryas patagoniensis							P					[54]
Pseudoferania polylepis			x		x		x	x				[50]
Rhabdophis tigrinus						x	x	t				[50]
Telescopus dhara			x		x		x			x		[50,51]
Thamnodynastes strigatus			TP	t *	t		TP	TP	T			[60]
Thrasops jacksonii			x		x			x				[51]
Trimorphodon biscutatus		B			B		B					[54,61]
Xenodon merremi			T		T	T *	T	T				This work

Protein categories are: LAAO, L-amino acid oxidase; PLA2 (IA), phospholipase A_2 (type IA); SVMP, snake venom metalloproteinase; SVSP, snake venom serine proteinase; 3FTx, three finger toxin; CNP, C-type natriuretic peptide; CRISP, cysteine rich secretory protein; CTL, C-type lectin, DEFEN, defensin (crotamine-like); KUN-1, Kunitz type protein (type 1); and KUN-2, Kunitz type protein (type 2). Types of evidence: T = Expressed in VG transcriptome at high level; t = Expressed in VG transcriptome at low (or uninformed) level; x = RT-PCR (non-quantitative); P = Detected in the proteome by MS/MS; and B = Protein purified and/or activity tested from the Duvernoy's venom. The green color graduation represents the strength of the combination of evidence for each product, from light (less) to dark (more). Note: * = only 3'UTR detected.

Table 2. Minor snake venom components and their occurrences in colubrid species.

Species	Enzymatic							Non-Enzymatic										Reference
	5NUCL	AChE	DPP	FactV	FactX	HYAL	PDE	AVIT	bPLA2i	CVF	CYST	gPLA2i	KU-WA	NGF *	OHA	VEGF-A **	WAP	
Boiga dendrophila																		[42,43]
Boiga irregularis		T				t	t		t	t	t	t				t	t	[12,44,45]
Borikenophis portoricensis		B					B											[46,47]
Cerberus rynchops																		[48]
Coelognathus radiatus																		[49]
Dispholidus typus																		[50–52]
Erythrolamprus miliaris		t	t			t				t		T				t	xt	This work; [50]
Erythrolamprus poecilogyrus																	x	[50,51]
Helicops angulatus																		[53]
Hypsiglena sp.			t										tP		t	t	t	[13]
Hypsiglena torquata																		[54]
Leoiheterodon madagascarensis																		[50]
Macropisthodon rudis							t											[55]
Opheodrys aestivus		t	t							t	x		t	x	t	t	t	[9]
Oxybelis fulgidus																		[56]
Oxyrhopus guibei		t	t							t		t				t	t	This work
Phalotris mertensi	tP	t							t		t	tP	t		tP	t	T	[57]
Pantherophis guttatus		t	t						***	t	t	***	t	x	t	t	t	[9]
Philodryas baroni																		[54]
Philodryas chamissonis																		[58]
Philodryas olfersii						t						T					x	[51,59]
Philodryas patagoniensis																		[54]
Pseudoferania polylepis																		[50]
Rhabdophis tigrinus																	x	[50]
Telescopus dhara																		[50,51]
Thamnodynastes strigatus		t				t												[60]
Thrasops jacksonii																	x	[51]
Trimorphodon biscutatus														x				[54,61]
Xenodon merremi			t *															This work

Protein categories are: 5NUCLEO, 5′nucleotidase; AChE, acetylcholinesterase; DPP, dipeptidyl peptidase; FactV, venom coagulation factor V; FactX, venom coagulation factor X; HYAL, hyaluronidase; PDE, phosphodiesterase; AVIT, AVIT protein; bPLA2i, beta type phospholipase A_2 inhibitor; CVF, cobra venom factor; CYST, cystatins; gPLA2i, gamma type phospholipase A_2 inhibitor; KU-WA-FU, ku-wap-fusin protein; NGF, nerve growth factor; OHA, ohanin (vesprin) protein; VEGF-A, vascular endothelial growth factor (type A); and WAP, waprin-like proteins. Types of evidence: T = Expressed in VG transcriptome at high level; t = Expressed in VG transcriptome at low (or uninformed) level; x = RT-PCR (non-quantitative); P = Detected in the proteome by MS/MS; and B = Protein purified and/or activity tested from the Duvernoy's venom. The color gradation represents the strength of the combination of evidence for each product, from light (less) to dark (more). Note: * partial sequences from other colubrids were PCR amplified as part of a phylogenetic study [62]; ** no VEGF-F (svVEGF) detected in colubrids; *** cDNA and protein isolated from liver and serum of *P. quadrivirgata* and *P. climacophora*.

Table 3. Putative new snake venom components identified from colubrid species.

Species	Enzymatic				Non-Enzymatic				Reference
	svLIPA	PLA_2 (IIE)	PLB	svMMP	EGFr	Lacta	LIPO	Vefico	
Boiga dendrophila									[42,43]
Boiga irregularis	t	t	t					t	[12,44,45]
Borikenophis portoricensis									[46,47]
Cerberus rynchops								TP	[48]
Coelognathus radiatus									[49]
Dispholidus typus		x					x		[50–52]
Erythrolamprus miliaris			t	T			T		This work; [50]
Erythrolamprus poecilogyrus				x					[50,51]
Helicops angulatus									[53]
Hypsiglena sp.								t	[13]
Hypsiglena torquata									[54]
Leoiheterodon madagascarensis		x							[50]
Macropisthodon rudis									[55]
Opheodrys aestivus	t	t						t	[9]
Oxybelis fulgidus									[56]
Oxyrhopus guibei	t	T	t	t		T	T	t	This work
Phalotris mertensi	TP		t	tP				t	[57]
Pantherophis guttatus	t	t	t						[9]
Philodryas baroni									[54]
Philodryas chamissonis									[58]
Philodryas olfersii	t		T	t				t	[51,59]
Philodryas patagoniensis									[54]
Pseudoferania polylepis								x	[50]
Rhabdophis tigrinus				TB			x		[50]
Telescopus dhara									[50,51]
Thamnodynastes strigatus				TP	T	TP			[60]
Thrasops jacksonii									[51]
Trimorphodon biscutatus							x		[54,61]
Xenodon merremi									This work

Protein categories are: LIPA, snake venom acid lipase; PLA2 (IIE), phospholipase A_2 (type IIE); PLB, phospholipase B; svMMP, snake venom matrix metalloproteinase; EGFr, EGF repeats protein; Lacta, lactadherin-like protein; LIPO, lipocalin; and Vefico, veficolin (ficolin-like). Types of evidence: T = Expressed in VG transcriptome at high level; t = Expressed in VG transcriptome at low (or uninformed) level; x = RT-PCR (non-quantitative); P = Detected in the proteome by MS/MS; and B = Protein purified and/or activity tested from the Duvernoy's venom. The color gradation represents the strength of the combination of evidence for each product, from light (less) to dark (more).

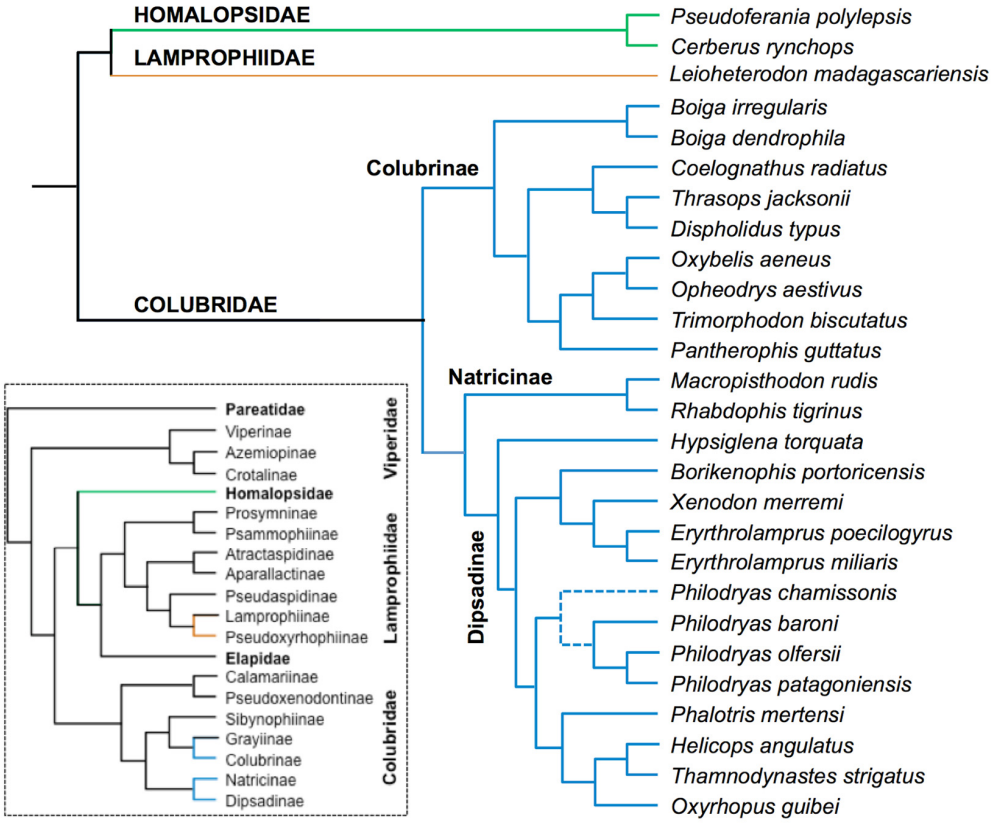

Figure 1. Schematic cladograms showing the phylogenetic relationships among families and species of snakes discussed in this work (colored branches). The cladogram was based on the phylogenetic tree proposed by Pyron et al. [34]. Dashed lines in *Philodryas* indicate the presumed placement of *P. chamissonis*.

It is interesting to note that most -omics characterizations of colubrid venoms have addressed members of the Dipsadinae subfamily of Colubridae, perhaps because a large number of genera in this subfamily are rear-fanged and possess Duvernoy's venom glands, and several have been involved in human envenomations, typically with mild effects [63–65]. The Dipsadinae species studied include *Philodryas olfersii* [59], *Thamnodynastes strigatus* [60] and *Hypsiglena* sp. [12], as well as *Erythrolamprus miliaris*, *Oxyrhopus guibei* and Xenodon merremi described here. For Colubrinae, transcriptomes of oral glands from *Pantherophis guttatus*, *Opheodrys aestivus* [9] and *Boiga irregularis* (Duvernoy's venom gland); [12] were generated, although only the last one was complemented by venom proteomic analysis. Nevertheless, many toxins from the other subfamilies have been investigated by more focused approaches, such as protein purification from the venom (e.g., *Borikenophis portoricensis* [47]) or specific cDNA cloning, including some genera with particularly toxic venom, such as the natricine *Rhabdophis* [66]. Very recently, full length mRNAs derived from secreted venoms of several colubrine and dipsadine colubrids were reverse transcribed and sequenced, demonstrating that it is possible to obtain transcript sequences from venom alone [67].

2.2. Major Snake Venom Enzymatic Components

For most colubrid species, especially in the subfamily Dipsadinae, snake venom metalloproteinases (SVMPs) are predominant components in the transcriptomes and in the proteomes. All sequences described in Colubridae to date belong to the P-III class of SVMPs, which include pre- and pro-domains, a metalloproteinase catalytic domain, a disintegrin-like domain and a cysteine-rich domain (Figure 2). The absence of P-II, P-I and short coding disintegrins in colubrid venoms is in accordance with the hypothesis that those proteins evolved within the family Viperidae from a P-III ancestor gene, after the split of this lineage [68,69]. A solely exception in Colubridae is the occurrence of a shortened P-III

SVMP in *Phalotris mertensi*. This protein was proteomically confirmed in the venom of the species and it has a partial disintegrin-like domain and no Cys-rich domain, as a result of a transcript with an early stop codon and a substituted 3′UTR sequence [57]. A phylogenetic tree of representative SVMPs indicates that, despite a high degree of diversity among the Colubridae SVMPs, they share a common ancestor with elapid and atractaspidid P-III SVMPs (Figure 2).

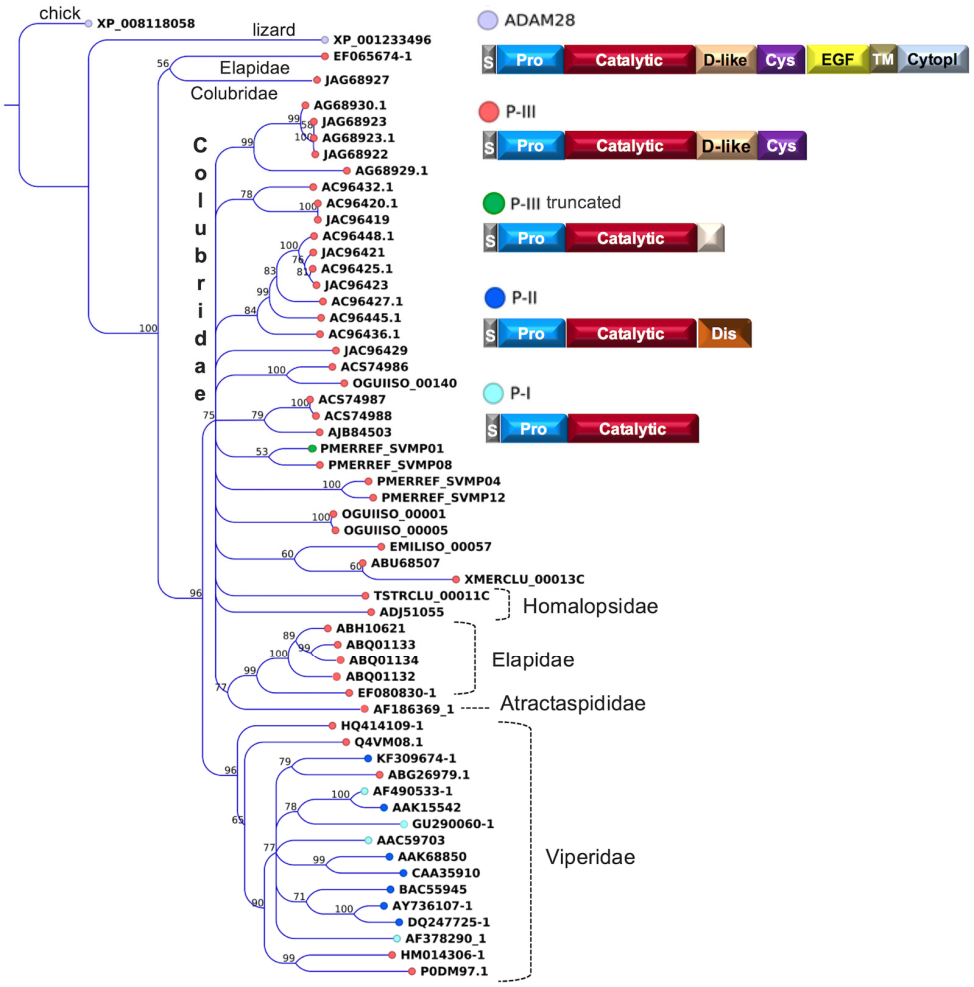

Figure 2. Maximum likelihood tree showing the relationship among representative SVMPs from different snake families. Bootstrap values are plotted close to the internal nodes. Colors in the terminal nodes indicate the types of the precursors, and their domain arrangements are depicted on the right. Abbreviated domains are: S, signal peptide; Pro, prodomain; Catalytic, metalloproteinase; D-like, disintegrin-like; Dis, disintegrin; Cys, cysteine rich; TM, transmembrane; EGF, epidermal growth factor; and Cytopl, cytoplasmic. The protein sequences are referred to by their accession numbers in GenBank, except those initiated by the codes EMILISO, OGUIISO, PMERREF, TSTRCLU and XMERCLU, which are mentioned in the "definition" field of sequence files deposited in the Transcriptome Shotgun Assembly (TSA) database.

Snake venom serine proteinases (SVSPs) are detected in some colubrid venoms and transcriptomes; however, they are not commonly present in these venoms, nor as abundantly expressed and diversified as observed in many viperid snakes. The few colubrid SVSPs sequenced are related to the kallikrein-like enzymes well characterized in viperid venoms, and they include a C-terminal extension that distinguishes them from the lizard venom kallikrein-like enzymes [70].

Phospholipases A_2 (PLA_2) are very common components in the venoms of the medically important snake families Elapidae and Viperidae, and they belong to type I and II PLA_2s, respectively.

In colubrids, they seem to not be among major components and have been detected in only a few species. In the colubrine *Trimorphodon biscutatus*, an enzyme was purified and its partial sequence indicated that it was a type IA PLA_2 [61]. However, in another colubrine (*Dispholidus typus*) [51], in the dipsadine *Oxyrhopus guibei* (this work) and in the pseudoxyrhophiine (family Lamprophiidae) *Leioheterodon madagascariensis* [51], among others, the reported type of PLA_2 is IIE.

The occurrence of transcripts coding for enzymes of IIE subtype in the venom glands indicates a possible independent recruitment of a PLA_2 to the venom, since they are distinct from the type IIA paralogs commonly expressed in the venom glands of viperid snakes [51]. Whether or not these type IIE PLA_2s represent truly new toxins or accessory proteins of the venom glands remains to be clarified, but Hargreaves et al. [9] found them to be exclusively expressed at low levels in the venom glands of the species tested.

Despite being very common in the venom of other groups of snakes, L-Amino Acid Oxidase (LAAO) was thought to be essentially non-existent in colubrid venoms (e.g., [71]). Very low levels of LAAO activity were detected in Brown Treesnake (*B. irregularis*) venom [72]; however, assays of venom from 13 different species of colubrine, dipsadine and natricine rear-fanged snakes detected no LAAO activity [73]. In a comparative transcriptomic analysis of tissues from *Pantherophis guttatus*, an LAAO was shown to be expressed in the scent gland but not in the salivary glands of this species [9], suggesting it is not a venom component. Recently, however, an LAAO was found moderately expressed in the venom glands of the colubrid *Phalotris mertensi*, and the MS/MS spectrometric analysis clearly showed its presence in the venom of this species [57].

2.3. Major Snake Venom Non-Enzymatic Components

Three-Finger Toxins (3FTx) are major constituents of Elapidae venoms and represent the lethal component of the majority of species of this family. These toxins seem to have differential importance in different subfamilies of colubrids. Alpha-colubritoxin from *Coelognathus radiatus* was the first colubrid toxin isolated and sequenced [49] and several other 3FTx, such as denmotoxin and irditoxin, functionally characterized in members of the subfamily Colubrinae, were demonstrated to be abundant toxins with taxon-specific activities [42,45]. The -omics characterization of *Boiga irregularis* venom showed that 3FTx dominate the transcriptome of this species (67.5% of toxin transcripts) [12]. The authors described 58 unique 3FTx sequences grouped into at least 10 sequence clusters that were proteomically confirmed in the venom. These clusters could be arranged in three groups based on the structural characteristics, but none of them were closely related to the above-mentioned irditoxin from the same species. Together with individual sequences isolated from other genera [9,50,56], 3FTx seem to be major components in many venoms of the subfamily Colubrinae. In contrast, in the Dipsadinae, 3FTx are not found [59,60] or are detected at minor abundance and diversity levels [57]. However, in the current work, we retrieved sequences from the transcriptome of the Duvernoy's venom gland of *Xenodon merremi* (a dipsadine colubrid) that were expressed at high level (Table 1). 3FTx-like sequences were also reported in venom glands of the family Lamprophiidae, as well as in species at the base of the Alethinophidia snake radiation (including species in the Cylindrophiidae and Pythonidae [50]). Nevertheless, it appears likely that 3FTx-like transcripts found in gland tissue of these latter two families may represent house-keeping genes, rather than toxins [9,74].

C-Type Lectins (CTL) are ubiquitous venom components in many snake groups, and they are also found abundantly in colubrid venoms. In Colubridae, amounts of venom CTL transcripts vary from 2% in *Philodryas olfersii* [59] to as much as 21% in *Phalotris mertensi* of total transcripts [57]. They were also reported in the snakes *Pseudoferania polylepis* and *Cerberus rynchops* (family Homalopsidae) and were highly expressed in *Cerberus* [48]. From a phylogenetic tree of all colubrid CTLs and related orthologs (Figure 3), it is possible to observe the existence of distinct types of CTLs in non-front fanged snakes, although the phylogenetic reconstruction failed to resolve the evolutionary relationships among them. Nevertheless, in addition to the presence of a CTL-like (snalec) clade, largely found in colubrids, colubrid sequences are observed to be nested within the clade of Elapidae and Viperidae "true" CTLs

sequences (i.e., those with a predicted galactose binding motif QPD substituting the plesiotypic motif EPN: [75,76]. One of them (PMERREF_CTL04) was confirmed in *P. mertensi* venom and in fact has the QPD motif, indicating that predicted galactose-binding lectins should also be present in other Colubridae venoms. Moreover, some orphan transcripts observed in the venom glands of snakes from different families clustered completely outside of the clades of typical venom CTLs (Figure 3). Some of them are suggested to code for venom proteins, such as two transcripts highly expressed and proteomically detected in the venom of *P. mertensi*, and similar transcripts are expressed at moderate levels in the venom glands of other colubrids (*Hypsiglena* sp. and *Boiga irregularis*; [12]), viperids (*Bothrops insularis*; [77]) and elapids (*Hoplocephalus bungaroides*; [78]). Interestingly, the encoded proteins from the transcripts of this group present not a single but various motifs (QPD, EPD, EPN, RPS, QVE, and EPK) for sugar binding at the second loop of the carbohydrate recognition domain. It indicates that these genes may have undergone a diversification process that parallels that experienced by other CTL types, i.e., the substitution of the binding motif of the original sugar ligand, mannose, by binding motifs to other types of carbohydrates.

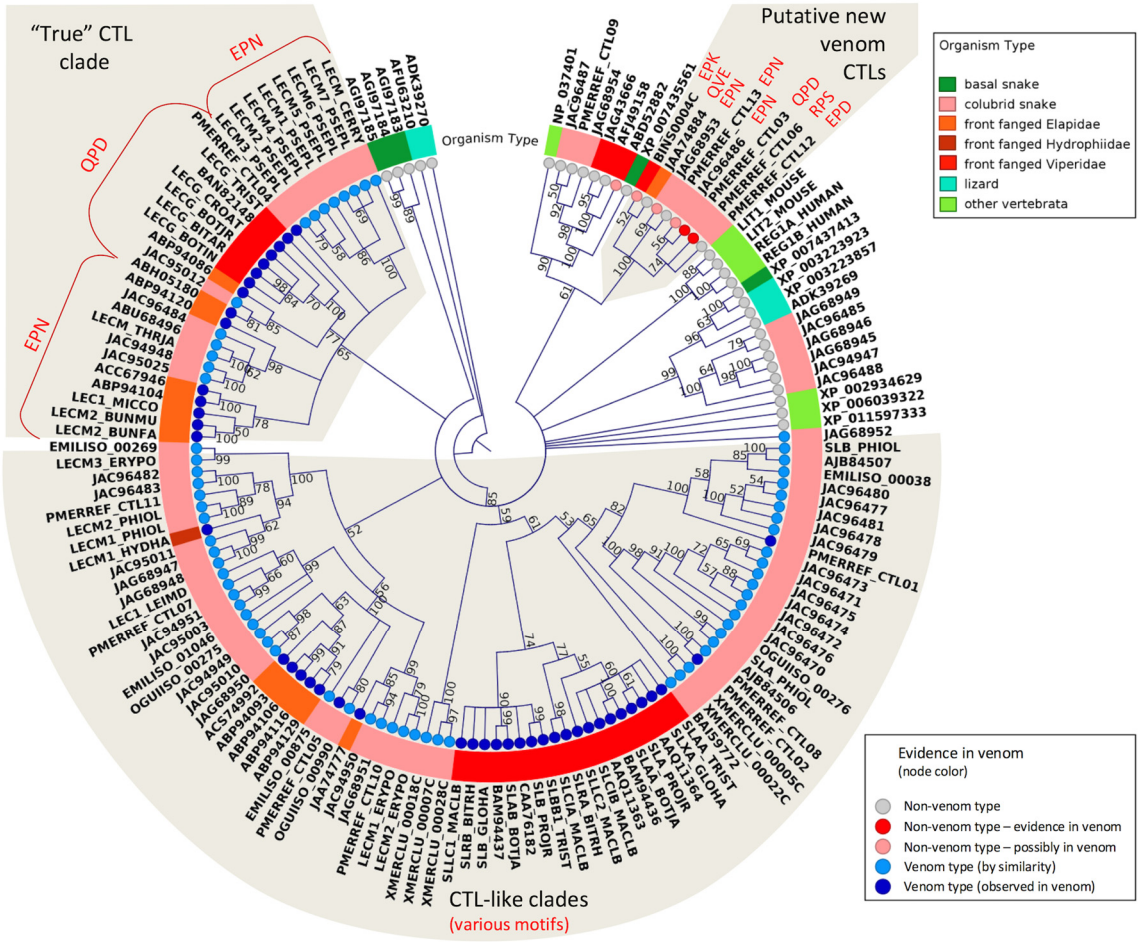

Figure 3. Maximum likelihood circular cladogram showing the relationship among representative CTLs from different snake families. Bootstrap values are plotted close to internal nodes. Colors at the terminal nodes (circles) indicate typical vs. atypical venom proteins and the evidence of occurrence in the venoms. Colors in the diagram surrounding the cladogram indicate the taxonomic groups. The carbohydrate binding motifs, as discussed in the text (EPN, QPD, etc.), are indicated by red type. The protein sequences are referred by their GenBank accession numbers, except those initiated by the codes EMILISO, OGUIISO, PMERREF and XMERCLU, which are mentioned in the "definition" field of sequence files deposited in TSA.

Although the role of cysteine rich secretory proteins (CRISPs) in venom is not yet clear, they are very ubiquitous venom components and are found in almost all snake species, including colubrids, and have been investigated via either classical protein techniques (e.g., [79]) or -omics profiling [80]. Contrary to the other highly expressed snake toxins, CRISPs seem to have not undergone multiple duplications during snake lineage evolution, and a single paralog is normally found abundantly expressed and translated to a venom protein in each colubrid species; in some species, such as *B. irregularis*, a minor isoform is also present in the venom (Mackessy, unpub. obs.). Nevertheless, positive Darwinian selection on CRISPs were observed to be higher in Colubridae and Viperidae proteins than on other reptiles, while negative selection occurs in mammalian CRISPs [80].

The first C-type natriuretic precursor (CNP) from a colubrid species was described from the *P. olfersii* transcriptome, where it was suggested to have a common ancestor with the natriuretic peptide precursor of elapid snakes and with the bradykinin-potentiating peptides precursor (BPP) of viperid snakes [59]. Currently, nine colubrid species in the three major subfamilies of Colubridae (Colubrinae, Dipsadinae and Natricinae) were shown to have this precursor generally highly expressed in the Duvernoy's venom glands. Most of them have the same general structure, i.e., the C-type peptide has no C-terminal extension and the CNP prodomain is not preceded by a BPP-containing region (Figure 4). Based on this organization, Jackson et al. [78] suggested that the acquisition of the *C*-terminal extension occurred within the Elapidae, while the acquisition of BBP repeats occurred along the viperid lineage diversification. We notice, however, a notable exception in the CNP precursor of the Dipsadinae *P. mertensi*: this precursor, transcribed at high levels in the venom glands, possesses a long sequence inserted at the middle of the CNP prodomain (linker domain), which is rich in Pro residues (including PP and PPP internal peptides) and resembles the BPP-containing region of the viperid precursor (Figure 4). At the C-terminal portion of this region, one particular motif, "QRFFPPPIPP", shows a high degree of similarity to the BPP signature. Besides the classical BPPs, which led to the development of successful anti-hypertensive drugs [81], the BPP precursors of Viperidae snakes were demonstrated to generate other bioactive peptides, including SVMP inhibitors [82–85]. It is thus reasonable to suppose that this region of the *P. mertensi* CNP precursor could also be processed to generate bioactive peptides and perhaps a BPP-like peptide.

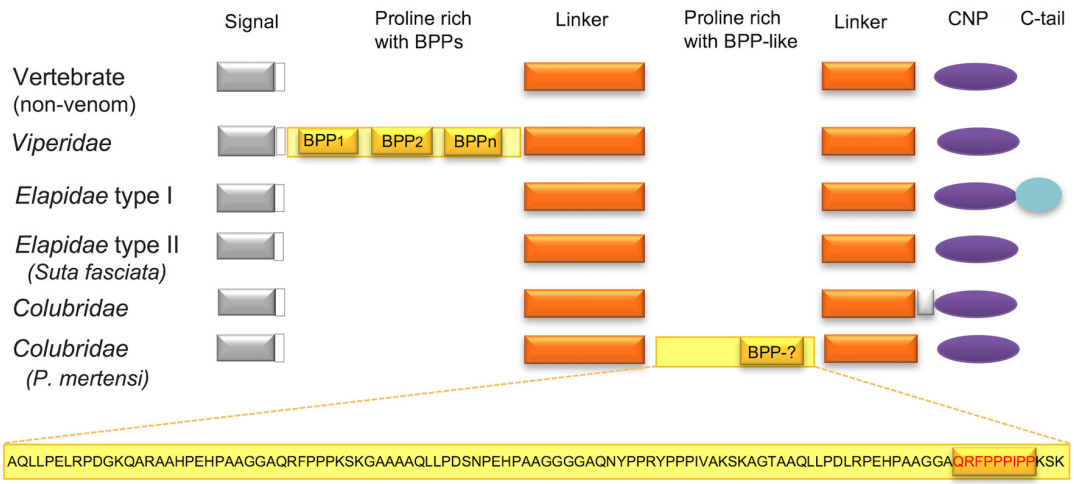

Figure 4. Schematic organization of CNP (and BPP) precursors in the different snake families and in other vertebrates. The precursor of *P. mertensi* [57] exhibits a Pro-rich insertion in the linker region (detached at the bottom), which includes a BPP-like segment that may generate a BPP after processing.

Crotamine is a beta-defensin-type polypeptide very well characterized from rattlesnake (Viperidae) venoms and thought to be restricted to the genus *Crotalus*. However, beta-defensin homologous genes were found in other viperid genera [86] and, more recently, venom gland transcripts were reported at relatively high expression levels in the transcriptome of the colubrids *Thamnodynastes strigatus* [60]

and *Phalotris mertensi* [57]. In the latter, the corresponding protein was detected by shotgun MS/MS analysis of the venom, suggesting it may be a valid colubrid venom component. The colubrid proteins have a highly conserved signal peptide, almost identical to that of crotamine (see Supporting Figure 3 from [60]); the mature polypeptides display the same cysteines involved in the disulfide arrangement of crotamine, but the other residues are highly variable, making it difficult to establish the evolutionary relationship between them.

Kunitz-type proteins appear in snake venoms in several forms, sometimes as single-product precursors (KUN-1), at other times with tandem repeated domains (KUN-2), and less frequently associated with WAP domains in a protein designated ku-wap-fusin (KU-WA-FU) [87]. Although in some species of colubrids these components have a transcriptional level not indicative of a relevant participant in venom, in at least two species, *Hypsiglena* sp. [12] and *Phalotris mertensi* [57], they have medium or elevated expression levels and were also detected in the venom. In *Phalotris mertensi*, three single-domain precursors are highly expressed and dominate the venom profile. The conservation of residues believed to be the protease inhibitory sites in their sequences [88,89] indicate they likely act as serine proteinase inhibitors, the plesiotypic function of this toxin, rather than as neurotoxins, as observed in some elapid Kunitz-like proteins.

2.4. Minor or Arguably Actual Venom Components

Other protein components previously reported in the venoms of the families Elapidae and/or Viperidae, generally in minor quantities, are also detected in low amounts in colubrid venoms and/or transcriptomes and are listed in Table 2. However, the actual contribution of these molecules to the venom is debatable, and some authors consider them non-toxins because of their occurrence in non-venom gland specific tissues [9].

Regarding minor occurring enzymes, venom-like acetylcholinesterase (Ache) sequences are found in many elapid species but were suggested as a colubrid venom component only in *Boiga* venom, where low activity was detected [44,73], as well as low expression levels in the transcriptome of *B. irregularis* [12]. A 5′-nucleotidase, on the other hand, was identified at low expression levels in the *Phalotris mertensi* transcriptome, and it was also detected in its venom proteome [57]. Factor Va- and Factor Xa-like proteins are venom serine proteinases distinct from the classical SVSPs [90], and they are believed to have been recruited into the venom proteome on the basis of their occurrences in venoms of the Australian elapid radiation [50]. Accordingly, no Factor Va-like sequence were retrieved in our searches, while only the endophysiological (non-venom) Factor Xa transcripts could be found in the venom glands of three colubrid species but were never identified in the secreted venom. These data indicate that the expression of endophysiological Factor Xa also may occur in the venom glands, in addition to the liver, although at low levels (this work and [9]). Transcripts for some other minor venom enzymes were only found at very low levels.

Growth factor sequences from colubrids and other snakes, such as vascular endothelial growth factor (VEGF) and nerve growth factor (NGF), are common in the databases. However, there is no clear evidence of them as venom components among colubrids. Whereas a venom-specific VEGF (VEGF-F) was extensively demonstrated to be a venom component in viperid snakes, possibly acting as a toxin dispersion agent [91], an endophysiological paralog (VEGF-A) was later shown to be co-expressed in venom glands and secreted only in low amounts in the venom of some species [92]. All the VEGF forms retrieved from colubrids are similar to VEGF-A, expressed at low levels, and thus they are more likely to be non-toxins, possibly corresponding to the endophysiological factor produced in the venom gland environment. The colubrid NGF sequences available in databases are mostly derived from phylogenetic studies based on this genetic marker. In contrast to other venomous snake families, where NGF was clearly demonstrated to be a venom component, in colubrids there is no support for this factor as a venom component, since it is not specifically transcribed in the venom glands of any species and the protein has to date not been isolated from the venom.

Cobra venom factor (CVF) was clearly demonstrated as a venom component only in elapid snakes [93]. Although very similar sequences could be found expressed at low levels in some colubrids, the absence of protein detection in their venom suggests the transcripts could also be the endophysiological complement factor C3 expressed by blood cells within the venom glands.

Enzymatic inhibitors that typically function to protect snakes from the bites of other snakes are mainly produced in the liver and secreted into the plasma of venomous and non-venomous snakes [94], but some of them seem to be produced in the venom glands. For example, a specific paralog of a gamma-PLA_2 inhibitor ($gPLA_2i$) was shown to be exclusively expressed in venom glands of *B. jararaca* (Viperidae) [10]. Accordingly, we could identify three colubrid species showing low to medium expression levels of $gPLA_2i$, and one of them was proteomically demonstrated in the venom. Protease inhibitors such as cystatins have been previous demonstrated in snake venoms [95], but their role in the venom is unclear. We retrieved transcripts coding for these proteins from some colubrids, but according to the analysis of Hargreaves et al. [9], they have undifferentiated levels of expression among tissues, and no further evidence of their presence in colubrid venoms have been noted yet, indicating that they are probably not colubrid venom components. Nevertheless, the common occurrence of many transcripts coding for all these toxin-like proteins in venom-producing tissues indicate that if they are not toxins, they may play important roles in the maintenance of this specialized secretory epithelium. We did not find transcripts related to sarafotoxins [96] in any colubrids, including *Leioheterodon madagascariensis* (Lamprophiidae), indicating that this component may be apotypic of Atractaspidinae.

2.5. Putative New Snake Toxins Suggested from Colubrid Venoms

Although SVMPs dominate many colubrid venom profiles, another type of metzincin, the snake venom matrix metalloproteinase (svMMP), was revealed to be a colubrid-specific venom component likely playing an important role in some species. svMMPs were abundantly found in the transcriptome and proteome of *Thamnodynastes strigatus* [60] and are highly expressed in the transcriptome of *Erythrolamprus miliaris* (this work), both Dipsadinae. Other Dipsadinae species investigated by similar -omics approaches showed lower abundance of svMMPs, though they were still detected in the venoms of some species. A svMMP was also purified, sequenced and functionally characterized from the venom of *Rhabdophis tigrinus* (Natricinae) [66]. For many of the species in which svMMPs were detected, SVMPs were also present in the venom, seemingly indicating that svMMPs are not substituting for the function of SVMPs but perhaps are adding a possible synergistic effect toward producing extracellular matrix lesions caused by the venoms. The colubrid svMMPs show important differences related to the presence of ancillary domains, as illustrated in Figure 5: whereas in some species, such as *Rhabdophis tigrinus*, the protein has a classical MMP9-like structure, in others, such as *Thamnodynastes strigatus*, they do not include the fibronectin repeats nor the hemopexin domains, thus resembling a MMP7-like arrangement. The *E. miliaris* svMMPs found in this work revealed a more complex situation, since some of the precursors have the fibronectin repeats inserted in the catalytic domain, whereas other precursors do not show these domains (Figure 5). Both forms are highly transcribed in the venom glands, representing the major toxin type found in the transcriptome of this species, but unfortunately, we did not have access to the venom of this species to evaluate its effective secretion. It is interesting to observe in a phylogenetic tree of svMMP precursors (Figure 5) that there is a strong clustering of svMMPs within the MMP-9 clade. This result indicates that all svMMPs seem to derive from a single MMP-9 ancestor gene, regardless of the presence or the absence of ancillary domains. Additionally, the clustering of species-specific proteins in monophyletic groups signifies intra-clade gene duplications, with independent losses of the fibronectin and hemopexin domains in some clades (*Thamnodynastes* and *Erythrolamprus*). Moreover, it clearly points out that the simplified MMP7-like arrangement observed in some svMMPs is a derived trait from the modification of a MMP9-type svMMP, rather than originating from an MMP7 gene.

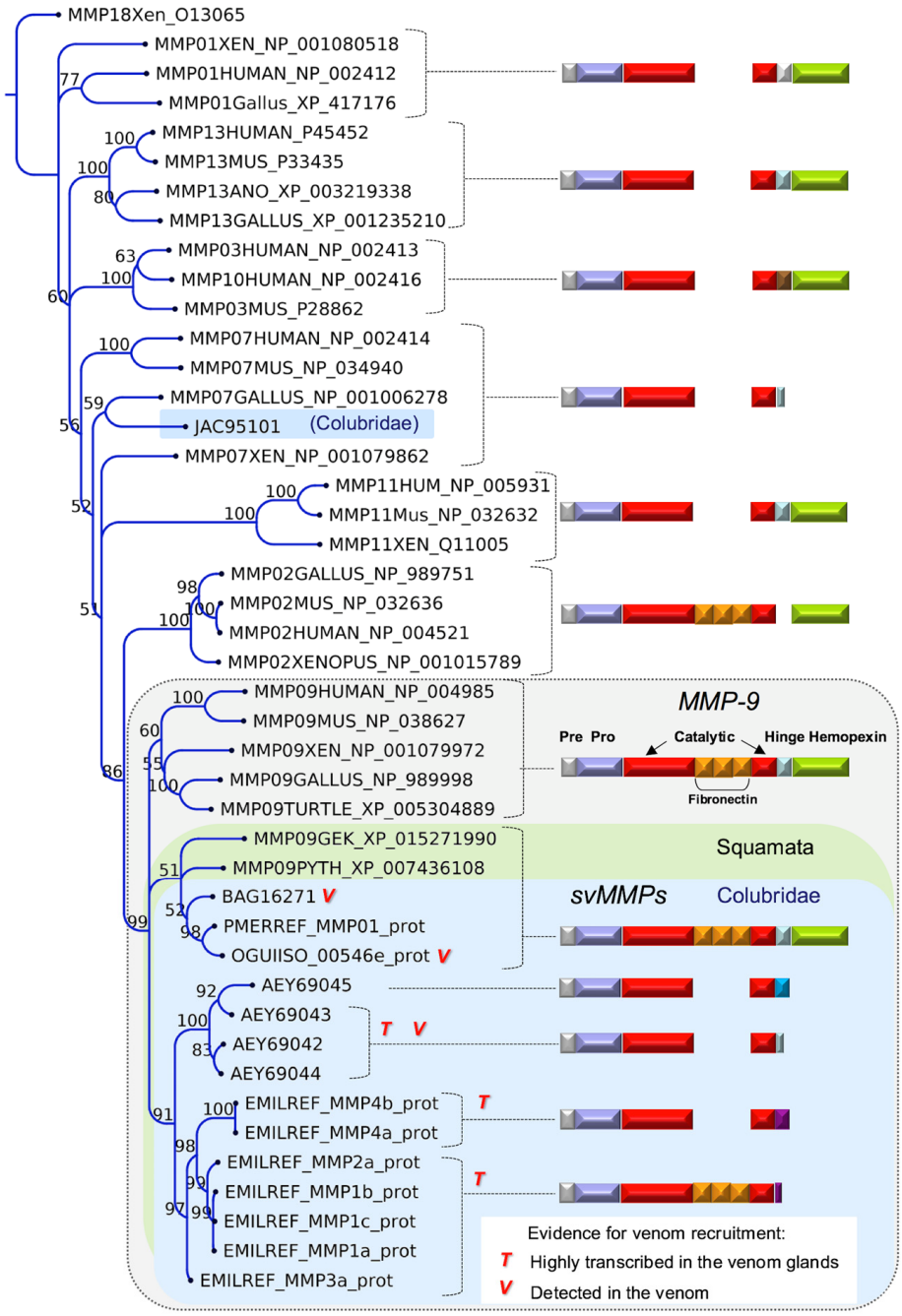

Figure 5. Maximum likelihood tree showing the relationship among svMMPs from different snake families and MMPs from other vertebrate groups. Bootstrap values are plotted close to internal nodes. The domain arrangement of each precursor type is depicted on the right. The types of evidence for the occurrence in venoms are indicated by "T" (transcribed) and "V" (detected in venom). The protein sequences are labeled by their accession numbers in GenBank, except those initiated by the codes EMILREF, OGUIISO, and PMERREF, which are mentioned in the "definition" field of sequence files deposited in TSA.

Another enzyme representing an example of a putative toxin from colubrid venom is an acid lipase (svLIPA), similar to mammalian lysosomal acid lipases. In *P. mertensi*, this protein was proteomically and immunochemically detected in the venom and its mRNA was highly expressed in the venom glands [57]. Interestingly, this *P. mertensi* sequence is closely related to acid lipases previously suggested as possible venom components in species of other snake families but not clearly demonstrated in their venoms [97,98], as well as a prominent protein component of saliva from several species of *Varanus* (BLAST search). By comparing acid lipase sequences from different reptiles, we could demonstrate that all transcripts showing evidence for venom proteins in different snakes (i.e., high expression in the venom glands, proteomics detection, or immunoreactivity in venom) form a monophyletic group, and thus LIPA may represent a novel type of venom component, and perhaps a toxin [57].

Novel non-enzymatic components were also proposed from the venoms of non-front fanged snakes. Venom ficolin (veficolin) is a class of putative toxins initially characterized from the homalopsid *Cerberus rynchops* venom and transcriptome [48]. Other related transcripts could be retrieved from several colubrid species but they are generally expressed at low levels, and the encoded proteins were not detected in any other venom. A lactadherin-like protein, a secreted carrier protein containing a FA58C (coagulation factor V and VIII C-terminal) domain, was first identified from a partial clone in the transcriptome of *T. strigatus*. Since it was found proteomically in the venom of that species, it was suggested as a possible venom component [60]. In the present work, we identified a complete transcript coding for this protein in the *Oxyrhopus guibei* transcriptome, but we did not evaluate the venom of this species. A search for similar transcripts in other snakes revealed a complete sequence only in the transcriptome of the viperid *Crotalus horridus* (JAA96713, [15]). An EGF repeat-containing cDNA was found in relatively high levels in the transcriptome of *T. strigatus* but was not confirmed in this venom nor was it retrieved from other species [60].

An interesting case of a potentially new venom component identified from Colubridae -omics analysis is a type of lipocalin. Transcripts coding for lipocalin-structured proteins were retrieved from several snake venom glands by transcriptomic analysis or by RT-PCR amplification and they were shown to be homologous [51]. In the transcriptomic analysis of the Atractaspidinae *Atractaspis aterrima*, some lipocalin sequences were identified as among the most expressed transcripts in the venom glands [18]. Since lipocalins are common components from some invertebrate venoms and from the saliva of hematophagous animals [99], they were suggested as possible venom components [47]. These proteins also show weak sequence similarities to a putative olfactory protein specifically expressed in high amounts in the Bowman's glands of the olfactory tissue from a frog [100]. Interestingly, among the original data generated in the present work, we found an extremely highly expressed transcript coding for a lipocalin in *Oxyrhopus guibei*. Alone, this mRNA accounts for 29% of the sequencing reads in the transcriptomic analysis. A phylogenetic tree of all available lipocalin sequences, from snakes and from several other sources, showed that the transcripts highly expressed in snake venom glands, including those from Colubridae and Atractaspidinae, are likely to be orthologs, whereas other transcripts expressed at low levels correspond to a paralogous snake gene (Figure 6). Although it is not possible to confirm, without a proteomic analysis, if lipocalin is indeed a venom component, the high expression of the same gene in the venom glands of distinct snake species suggests that its product should have an important role for this animal, perhaps as a new toxin or perhaps involved in olfactory-mediated behavior.

Finally, a putative new toxin proposed from a highly expressed transcript from *Atractaspsis aterrima* (Atractaspidine) [18] displayed some similarity with an unknown protein predicted from a high expressed contig from *Erythrolamprus miliaris*. However, the areas of conservation were restricted to the signal peptide and to the *C*-terminal and thus it is not likely that the two putative proteins correspond to a common toxin (data not shown).

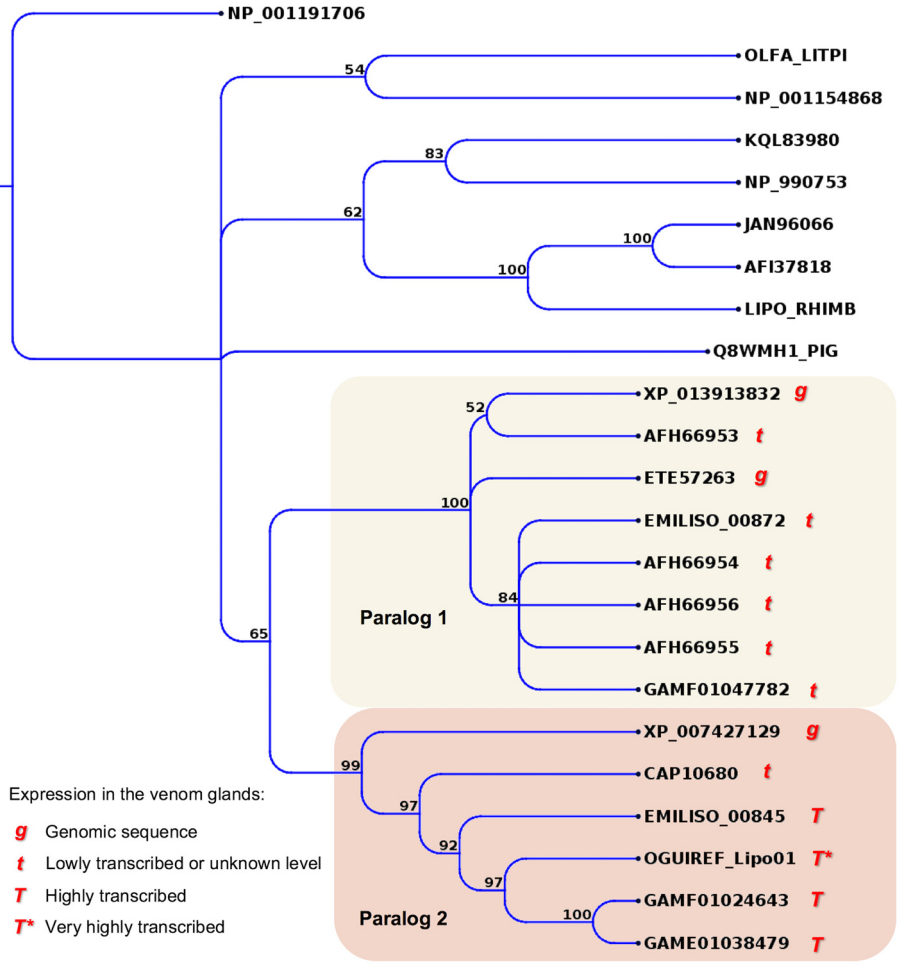

Figure 6. Maximum likelihood tree showing the relationship among lipocalin proteins from different snake families and from other vertebrate groups. Note that transcripts highly expressed in venom glands are all in the same clade. Bootstrap values are plotted close to internal nodes. The protein sequences are labeled by their accession numbers in GenBank, except those initiated by the code EMIL, which is mentioned in the "definition" field of sequence files deposited in TSA and the sequence OGUIREF_Lipo1 (Accession Number KX450875). Sequences labeled GAMF from *A aterrima* were translated from the original nucleotide contigs retrieved from TSA.

3. Conclusions

It is now abundantly clear that the venoms produced among the colubrid rear-fanged snakes are homologous with the much better characterized venoms of the front-fanged snakes. As trophic adaptations that facilitate feeding, venoms vary in composition with several important factors, including phylogeny, and so it is to be expected that among the diverse colubrid lineages, novel compounds, and new functional variants of better-known venom proteins, will be encountered. Much progress toward understanding rear-fanged snake venom composition has been made in the last decade, but, as indicated above, we have barely begun to explore the diversity of advanced snakes that comprise the colubrids. Transcriptomic and genomic approaches will greatly facilitate this work, but it must be remembered that functional assays should also accompany analysis of any venom, because the common recurring motif in venom biochemistry is to make the most of a stable molecular scaffold, perhaps best exemplified by the varied pharmacologies of the three-finger toxin superfamily. These small, structurally conservative peptides have very similar crystal structures but affect systems as diverse as neurotransmission, the blood clot cascade, ion channel function, and salamander limb regeneration and courtship. As Dr. Jay Fox once said, in venoms "we find only what we are looking

for", and, to find truly novel toxins that will likely be present in some colubrid venoms, we will have to look beyond the "normal" families of venom proteins.

4. Materials and Methods

4.1. Original Transcriptomic Data

4.1.1. Animals

Three specimens of *Erythrolamprus miliaris* (one male and two females) five specimens of *Oxyrhopus guibei* (two males and three females) and two specimens of *Xenodon merremi* (both female) were provided by the Laboratory of Herpetology at the Instituto Butantan. These animals were collected from the wild by the local population, delivered at Instituto Butantan and kept in captivity for a short time (up to one month); all snakes were provided water *ad lib* but not fed. Manual extraction of the venom was performed 4 days prior to euthanizing the animals and dissecting out both Duvernoy's venom glands, which were frozen in liquid nitrogen. All animal procedures were authorized by the Ethical Committee for Animal Research of Butantan Institute (protocols 164/2004 and 935/12, approved on 11 May 2004 and 1 June 2012, respectively), according to principles adopted by the Brazilian College of Animal Experimentation.

4.1.2. RNA-Seq

Erythrolamprus miliaris and *Oxyrhopus guibei* transcriptomes were investigated using RNA-Seq, in a 454 pyrosequencing platform. Pairs of glands from each specimen were ground into a powder in liquid nitrogen and homogenized using a Polytron Tissue Homogenizer (Kinematica, Luzern, Switzerland). Total RNA was extracted with TRIZOL Reagent (Life Technologies, Thermo Fisher Scientific, Carlsbad, CA, USA) and mRNA was prepared using the Dynabeads mRNA DIRECT kit (Life Technologies, Thermo Fisher Scientific, Carlsbad, CA, USA). mRNA was quantified by the Quant-iT^{TM} RiboGreen RNA reagent and kit (Life Technologies, Thermo Fisher Scientific, Carlsbad, CA, USA). To obtain 500 ng of mRNA needed to prepare cDNA libraries for pyrosequencing with cDNA Synthesis System kit (Roche Diagnostics, Basel, Switzerland), we pooled mRNAs from individual specimens of each species. Emulsion PCR amplification and library sequencing were performed individually for each species, using a GS Junior 454 Sequencing System (Roche Diagnostics, Basel, Switzerland) according to the manufacturer's protocols. The raw sequences were deposited in GenBank SRA with the accession numbers SRR3141951-SRR3141952 (*Erythrolamprus miliaris*) and SRR3141953 (*Oxyrhopus guibei*).

The raw reads from each species were assembled with Newbler 2.7 (Roche Diagnostics, Basel, Switzerland), which first removes adaptors and contaminating ribosomal RNA sequences. The assembly parameters were set to: (i) a minimum overlap length of 50% of the read; and (ii) a minimum overlap identity of 98%, with all other parameters set as default. The resulting unigenes were deposited in the GenBank TSA repository with the accession numbers GEFK00000000.1 linked to Bioproject PRJNA310611 (*Erythrolamprus miliaris*) and GEFL00000000.1 linked to Bioproject PRJNA310661 (*Oxyrhopus guibei*). Unigene sequences were automatically annotated using Blast2Go [101] by performing a Blast search against the UniProt database with the algorithm BlastX to identify similar sequences. Toxin categories were manually assigned by comparing the unigenes to a compiled list of known snake toxins. Final manual curation of relevant unigene sequences was undertaken to improve the quality and the extension of the automatically assembled unigenes. The levels of expression of individual unigenes were calculated using the RNA-Seq function of CLC Genomics Workbench v8 (Qiagen, Hilden, Germany, 2015)) by mapping cleaned reads (without known contaminants and rRNAs) back to the unigenes and normalizing the count number by the length of the unigene using RPKM (reads per kilobase per million of mapped reads) formula [102].

4.1.3. Expressed Sequence Tags (ESTs) Generation

Xenodon merremi transcriptome was investigated by means of EST generation, prior to the common use of NGS (next generation sequencing). The pairs of glands from *Xenodon merremi* specimens were ground into a powder in liquid nitrogen and homogenized using a Polytron Tissue Homogenizer (Kinematica, Luzern, Switzerland). Total RNA was extracted with TRIZOL reagent (Life Technologies, Thermo Fisher Scientific, Carlsbad, CA, USA) and mRNA was prepared using a column of oligo-dT cellulose (GE). To obtain 5 µg of mRNA needed to prepare cDNA libraries using the Superscript Plasmid System for cDNA Synthesis and Cloning (Life Technologies, Thermo Fisher Scientific, Carlsbad, CA, USA), we pooled mRNAs from the two specimens. The cDNA was ligated with the adaptors included in the kit, size selected into two ranges (250–600 bp and over 600 bp), directionally cloned into pSPORT-1 plasmids and transformed in *E. coli* DH5α electrocompetent cells. Plasmid DNA was isolated using alkaline lysis and sequenced on an ABI 3100 sequencer using the BigDye 3.1 kit (Applied Biosystems, Foster City, CA, USA) with a standard 5′ primer (M13R). The electropherogram files were analyzed in a semi-automatic way and then assembled in clusters of contiguous sequences using the CAP3 program [103] set for 98% or more of base identity in a high-quality region. The resulting unigenes were deposited in GenBank TSA repository with the accession number GETV00000000 linked to Bioproject PRJNA310192. The relative representation of each cluster was given by the number of ESTs used in its assembly, as described elsewhere [104].

4.2. Public Database Sequence Retrieval

Prototypical sequences of the different kinds of proteins known in the venoms of Viperidae, Elapidae, Atractaspididae and Colubridae snakes were compiled from GenBank and from our archives and used as in silico probes to a more extensive search for related protein sequences from Colubroidea snakes. Searches were performed using a stand-alone Blast tool of CLC Genomics Workbench v8 against a downloaded version of the GenBank nr database (December 2015), which includes non-redundant protein sequences from GenBank database and protein sequences from TSA (transcriptome shotgun assembly) repository. An initial e-value cutoff of lower than 10^{-5} was considered but the alignment of each sequence identified was individually evaluated to decide for the retrieving of a given protein. Whenever possible, the expression level of the transcript coding for each protein was examined from the bibliographical reference associated with the sequence to assign the "T" or "t" symbols on the summarizing tables, corresponding, respectively, to "high" (meaning highly expressed or higher expressed than in other tissues) or "low" (meaning lowly expressed or lower expressed than in other tissues) transcriptional level in the venom glands. Evidence for proteomic identification of proteins was also obtained from the literature.

4.3. Sequence Comparisons and Evolutionary Analyses

Protein sequences were incorporated into gene family sequence alignments containing toxin and non-toxin protein homologs and paralogs and representative outgroup sequences. Alignments were performed using MUSCLE or CLUSTALW tools of CLC Genomics Workbench v8 (Qiagen) and checked manually. Phylogenetic trees were generated by the Maximum Likelihood method, with WAG substitution model and bootstrapping 1000 replicates, using CLC Genomics Workbench.

Supplementary Materials: The following are available online at www.mdpi.com/2072-6651/8/8/230/s1, Table S1: Assembled contigs from the species investigated and their expression values.

Acknowledgments: This work was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, grant numbers 12/00177-5, 2013/07467-1) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) to Inácio L.M. Junqueira-de-Azevedo and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (fellowship 387422) to Pollyanna Fernandes Campos.

Author Contributions: I.L.M.J.-d.-A. conceived the research, analyzed sequences and wrote the paper; P.F.C. and A.T.C.C. performed the transcriptomic experiments and analyzed sequences; and S.P.M. discussed the results and wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

3FTx	three finger toxin
5NUCLEO	5′nucleotidase
AChE	acetylcholinesterase
AVIT	AVIT protein
bPLA2i	beta type phospholipase A_2 inhibitor
CDS	coding DNA sequence
CNP	C-type natriuretic peptide
CRISP	cysteine rich secretory protein
CTL	C-type lectin
DEFEN	defensin (crotamine-like)
CVF	cobra venom factor
CYST	cystatins
DEF	defensin
DPP	dipeptidyl peptidase
EGFr	EGF repeats protein
ESTs	expressed sequence tags
FactV	venom coagulation factor V
FactX	venom coagulation factor X
$gPLA_2i$	gamma type phospholipase A_2 inhibitor
HYAL	hyaluronidase
KUN-1	kunitz type protein (type 1)
KUN-2	kunitz type protein (type 2)
KU-WA-FU	ku-wap-fusin protein
LAAO	L-amino acid oxidase
Lactha	lactadherin-like protein
LIPA	snake venom acid lipase
LIPO	lipocalin
NGF	nerve growth factor
OHA	ohanin (vesprin) protein
PDE	phosphodiesterase
PLA_2 (IA)	phospholipase A_2 (type IA)
PLA_2 (IIE)	phospholipase A_2 (type IIE)
PLB	phospholipase B
RPKM	reads per kilobase per million of mapped reads
svMMP	snake venom matrix metalloproteinase
SVMP	snake venom metalloproteinase
SVSP	snake venom serine proteinase
TSA	transcriptome shotgun assembly
Vefico	veficolin (ficolin-like)
VEGF-A	vascular endothelial growth factor (type A)
WAP	waprin-like proteins

References

1. Calvete, J.J. Snake venomics: From the inventory of toxins to biology. *Toxicon* **2013**, *75*, 44–62. [CrossRef] [PubMed]
2. Ducancel, F.; Durban, J.; Verdenaud, M. Transcriptomics and venomics: Implications for medicinal chemistry. *Future Med. Chem.* **2014**, *15*, 1629–1643. [CrossRef] [PubMed]
3. Brahma, R.K.; McCleary, R.J.; Kini, R.M.; Doley, R. Venom gland transcriptomics for identifying, cataloging, and characterizing venom proteins in snakes. *Toxicon* **2015**, *93*, 1–10. [CrossRef] [PubMed]
4. Gilles, N.; Servent, D. The European FP7 Venomics Project. *Future Med. Chem.* **2014**, *15*, 1611–1612. [CrossRef] [PubMed]
5. Aird, S.D.; Watanabe, Y.; Villar-Briones, A.; Roy, M.C.; Terada, K.; Mikheyev, A.S. Quantitative high-throughput profiling of snake venom gland transcriptomes and proteomes (*Ovophis okinavensis* and *Protobothrops flavoviridis*). *BMC Genom.* **2013**, *14*. [CrossRef] [PubMed]
6. Aird, S.D.; Aggarwal, S.; Villar-Briones, A.; Tin, M.M.; Terada, K.; Mikheyev, A.S. Snake venoms are integrated systems, but abundant venom proteins evolve more rapidly. *BMC Genom.* **2015**, *16*. [CrossRef] [PubMed]

7. Chapeaurouge, A.; Reza, M.A.; Mackessy, S.P.; Carvalho, P.C.; Valente, R.H.; Teixeira-Ferreira, A.; Perales, J.; Lin, Q.; Kini, R.M. Interrogating the venom of the viperid snake Sistrurus catenatus edwardsii by a combined approach of electrospray and MALDI mass spectrometry. *PLoS ONE* **2015**, *10*. [CrossRef] [PubMed]
8. Durban, J.; Juárez, P.; Angulo, Y.; Lomonte, B.; Flores-Diaz, M.; Alape-Girón, A.; Sasa, M.; Sanz, L.; Gutiérrez, J.M.; Dopazo, J.; et al. Profiling the venom gland transcriptomes of Costa Rican snakes by 454 pyrosequencing. *BMC Genom.* **2011**, *12*. [CrossRef] [PubMed]
9. Hargreaves, A.D.; Swain, M.T.; Logan, D.W.; Mulley, J.F. Testing the Toxicofera: Comparative transcriptomics casts doubt on the single, early evolution of the reptile venom system. *Toxicon* **2014**, *92*, 140–156. [CrossRef] [PubMed]
10. Junqueira-de-Azevedo, I.L.; Bastos, C.M.; Ho, P.L.; Luna, M.S.; Yamanouye, N.; Casewell, N.R. Venom-related transcripts from *Bothrops jararaca* tissues provide novel molecular insights into the production and evolution of snake venom. *Mol. Biol. Evol.* **2015**, *32*, 754–766. [CrossRef] [PubMed]
11. Margres, M.J.; McGivern, J.J.; Wray, K.P.; Seavy, M.; Calvin, K.; Rokyta, D.R. Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *J. Proteom.* **2014**, *96*, 145–158. [CrossRef] [PubMed]
12. McGivern, J.J.; Wray, K.P.; Margres, M.J.; Couch, M.E.; Mackessy, S.P.; Rokyta, D.R. RNA-seq and high-definition mass spectrometry reveal the complex and divergent venoms of two rear-fanged colubrid snakes. *BMC Genom.* **2014**, *15*. [CrossRef] [PubMed]
13. Petras, D.; Heiss, P.; Süssmuth, R.D.; Calvete, J.J. Venom Proteomics of Indonesian King cobra, *Ophiophagus. hannah*: Integrating top-down and bottom-up approaches. *J. Proteome Res.* **2015**, *14*, 2539–2556. [CrossRef] [PubMed]
14. Reeks, T.; Lavergne, V.; Sunagar, K.; Jones, A.; Undheim, E.; Dunstan, N.; Fry, B.; Alewood, P.F. Deep venomics of the *Pseudonaja* genus reveals inter- and intra-specific variation. *J. Proteom.* **2016**, *133*, 20–32. [CrossRef] [PubMed]
15. Rokyta, D.R.; Wray, K.P.; Margres, M.J. The genesis of an exceptionally lethal venom in the timber rattlesnake (*Crotalus horridus*) revealed through comparative venom-gland transcriptomics. *BMC Genom.* **2013**, *14*. [CrossRef] [PubMed]
16. Tan, C.H.; Fung, S.Y.; Yap, M.K.; Leong, P.K.; Liew, J.L.; Tan, N.H. Unveiling the elusive and exotic: Venomics of the Malayan blue coral snake (*Calliophis bivirgata flaviceps*). *J. Proteom.* **2016**, *132*, 1–12. [CrossRef] [PubMed]
17. Viala, V.L.; Hildebrand, D.; Trusch, M.; Fucase, T.M.; Sciani, J.M.; Pimenta, D.C.; Arni, R.K.; Schlüter, H.; Betzel, C.; Mirtschin, P.; et al. Venomics of the Australian eastern brown snake (*Pseudonaja textilis*): Detection of new venom proteins and splicing variants. *Toxicon* **2015**, *107*, 252–265. [CrossRef] [PubMed]
18. Terrat, Y.; Sunagar, K.; Fry, B.G.; Jackson, T.N.; Scheib, H.; Fourmy, R.; Verdenaud, M.; Blanchet, G.; Antunes, A.; Ducancel, F. *Atractaspis aterrima* toxins: The first insight into the molecular evolution of venom in side-stabbers. *Toxins (Basel)* **2013**, *5*, 1948–1964. [CrossRef] [PubMed]
19. Casewell, N.R.; Wagstaff, S.C.; Wüster, W.; Cook, D.A.; Bolton, F.M.; King, S.I.; Pla, D.; Sanz, L.; Calvete, J.J.; Harrison, R.A. Medically important differences in snake venom composition are dictated by distinct postgenomic mechanisms. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9205–9210. [CrossRef] [PubMed]
20. Gibbs, H.L.; Mackessy, S.P. Functional basis of a molecular adaptation: Prey-specific toxic effects of venom from *Sistrurus rattlesnakes*. *Toxicon* **2009**, *53*, 672–679. [CrossRef] [PubMed]
21. Rokyta, D.R.; Wray, K.P.; McGivern, J.J.; Margres, M.J. The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the Timber Rattlesnake (*Crotalus horridus*). *Toxicon* **2015**, *98*, 34–48. [CrossRef] [PubMed]
22. Margres, M.J.; McGivern, J.J.; Seavy, M.; Wray, K.P.; Facente, J.; Rokyta, D.R. Contrasting modes and tempos of venom expression evolution in two snake species. *Genetics* **2015**, *199*, 165–176. [CrossRef] [PubMed]
23. Modahl, C.; Saviola, A.J.; Mackessy, S.P. Proteomic and genomic approaches to the study of rear-fanged ("colubrid") snake venoms. In *Handbooks of Toxinology. Venom Genomics and Proteomics*; Gopalakrishnakone, P., Ed.; Springer Science: Dordrecht, The Netherlands, 2015; p. 23.
24. Kardong, K.V. Evolutionary patterns in advanced snakes. *Am. Zool.* **1980**, *20*, 269–282. [CrossRef]
25. Savitzky, A.H. The role of venom delivery strategies in snake evolution. *Evolution* **1980**, *34*, 1194–1204. [CrossRef]
26. Vidal, N. Colubroid systematics: Evidence for an early appearance of the venom apparatus followed by extensive evolutionary tinkering. *J. Toxicol.-Toxin Rev.* **2002**, *21*, 27–47. [CrossRef]

27. Mackessy, S.P.; Saviola, A.J. Venoms from "non-venomous" snakes: Rear-fanged snake venoms as sources of novel compounds. In *Snake Venoms and Envenomation: Modern Trends and Future Prospects*; Utkin, Y., Krivoshein, A.V., Eds.; Nova Science Publishers, Inc.: New York, NY, USA, 2016; p. 29.
28. Dowling, H.G.; Hass, C.A.; Hedges, S.B.; Highton, R. Snake relationships revealed by slow-evolving proteins: A preliminary survey. *J. Zool. Lond.* **1996**, *240*, 1–28. [CrossRef]
29. Heise, P.J.; Maxson, L.R.; Dowling, H.G.; Hedges, S.B. Higherlevel snake phylogeny inferred from mitochondrial DNA sequences of 12S rRNA and 16S rRNA genes. *Mol. Biol. Evol.* **1995**, *12*, 259–265. [PubMed]
30. Kelly, C.M.R.; Barker, N.P.; Villet, M.H. Phylogenetics of advanced snakes (Caenophidia) based on four mitochondrial genes. *Syst. Biol.* **2003**, *52*, 439–459. [CrossRef] [PubMed]
31. Kraus, F.; Brown, W.M. Phylogenetic relationships of colubroid snakes based on mitochondrial DNA sequences. *Zool. J. Linn. Soc.* **1998**, *122*, 455–487. [CrossRef]
32. Lee, M.S.Y.; Hugall, A.F.; Lawson, R.; Scanlon, J.D. Snake phylogeny based on multiple morphological and molecular data sets. *Syst. Biodivers.* **2007**, *5*, 371–389. [CrossRef]
33. Pyron, R.A.; Burbrink, F.T.; Colli, G.R.; de Oca, A.N.; Vitt, L.J.; Kuczynski, C.A.; Wiens, J.J. The phylogeny of advanced snakes (Colubroidea), with discovery of a new subfamily and comparison of support methods for likelihood trees. *Mol. Phylogenet. Evol.* **2011**, *58*, 329–342. [CrossRef] [PubMed]
34. Pyron, R.A.; Burbrink, F.T.; Wiens, J.J. A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evol. Biol.* **2013**, *13*. [CrossRef] [PubMed]
35. Vidal, N.; Delmas, A.S.; David, P.; Cruaud, C.; Couloux, A.; Hedges, S.B. The phylogeny and classification of caenophidian snakes inferred from seven nuclear protein-coding genes. *C. R. Biol.* **2007**, *330*, 182–187. [CrossRef] [PubMed]
36. Uetz, P. The original descriptions of reptiles. *Zootaxa* **2010**, *2334*, 59–68.
37. Uetz, P. The Reptile Database. Available online: http://www.reptile-database.org (accessed on 3 April 2016).
38. Weinstein, S.A.; Smith, T.L.; Kardong, K.V. Reptile venom glands. Form, function, and future. In *Handbook of Venoms and Toxins of Reptiles*; Mackessy, S.P., Ed.; CRC Press/Taylor & Francis Group: Boca Raton, FL, USA, 2009; pp. 65–91.
39. Saviola, A.J.; Peichoto, M.E.; Mackessy, S.P. Rear-fanged snake venoms: An untapped source of novel compounds and potential drug leads. *Toxin Rev.* **2014**, *33*, 185–201. [CrossRef]
40. Zaher, H.; de Oliveira, L.; Grazziotin, F.G.; Campagner, M.; Jared, C.; Antoniazzi, M.M.; Prudente, A.L. Consuming viscous prey: A novel protein-secreting delivery system in neotropical snail-eating snakes. *BMC Evol. Biol.* **2014**, *14*. [CrossRef] [PubMed]
41. Mackessy, S.P. Biochemistry and pharmacology of colubrid snake venoms. *J. Toxicol.-Toxin Rev.* **2002**, *21*, 43–83. [CrossRef]
42. Pawlak, J.; Mackessy, S.P.; Fry, B.G.; Bhatia, M.; Mourier, G.; Fruchart-Gaillard, C.; Servent, D.; Ménez, R.; Stura, E.; Ménez, A.; et al. Denmotoxin, a three-finger toxin from the colubrid snake *Boiga dendrophila* (Mangrove Catsnake) with bird-specific activity. *J. Biol. Chem.* **2006**, *281*, 29030–29041. [CrossRef] [PubMed]
43. Lumsden, N.G.; Fry, B.G.; Ventura, S.; Kini, R.M.; Hodgson, W.C. Pharmacological characterisation of a neurotoxin from the venom of *Boiga dendrophila* (mangrove catsnake). *Toxicon* **2005**, *45*, 329–334. [CrossRef] [PubMed]
44. Mackessy, S.P.; Sixberry, N.M.; Heyborne, W.H.; Fritts, T. Venom of the Brown Treesnake, *Boiga irregularis*: Ontogenetic shifts and taxa-specific toxicity. *Toxicon* **2006**, *47*, 537–548. [CrossRef] [PubMed]
45. Pawlak, J.; Mackessy, S.P.; Sixberry, N.M.; Stura, E.A.; Le Du, M.H.; Ménez, R.; Foo, C.S.; Ménez, A.; Nirthanan, S.; Kini, R.M. Irditoxin, a novel covalently linked heterodimeric three-finger toxin with high taxon-specific neurotoxicity. *FASEB J.* **2009**, *23*, 534–545. [CrossRef] [PubMed]
46. Weldon, C.L.; Mackessy, S.P. Biological and proteomic analysis of venom from the Puerto Rican Racer (*Alsophis portoricensis*: Dipsadidae). *Toxicon* **2010**, *55*, 558–569. [CrossRef] [PubMed]
47. Weldon, C.L.; Mackessy, S.P. Alsophinase, a new P-III metalloproteinase with alpha-fibrinogenolytic and hemorrhagic activity from the venom of the Puerto Rican Racer *Alsophis portoricensis* (Serpentes: Dipsadidae). *Biochimie* **2012**, *94*, 1189–1198. [CrossRef] [PubMed]
48. OmPraba, G.; Chapeaurouge, A.; Doley, R.; Devi, K.R.; Padmanaban, P.; Venkatraman, C.; Velmurugan, D.; Lin, Q.; Kini, R.M. Identification of a novel family of snake venom proteins Veficolins from *Cerberus rynchops* using a venom gland transcriptomics and proteomics approach. *J. Proteome Res.* **2010**, *9*, 1882–1893. [CrossRef] [PubMed]

49. Fry, B.G.; Lumsden, N.G.; Wüster, W.; Wickramaratna, J.C.; Hodgson, W.C.; Kini, R.M. Isolation of a neurotoxin (alpha-colubritoxin) from a nonvenomous colubrid: Evidence for early origin of venom in snakes. *J. Mol. Evol.* **2003**, *57*, 446–452. [CrossRef] [PubMed]
50. Fry, B.G.; Scheib, H.; van der Weerd, L.; Young, B.; McNaughtan, J.; Ramjan, S.F.; Vidal, N.; Poelmann, R.E.; Norman, J.A. Evolution of an arsenal: Structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol. Cell. Proteom.* **2008**, *7*, 215–246. [CrossRef] [PubMed]
51. Fry, B.G.; Scheib, H.; Junqueira-de-Azevedo, I.L.M.; Silva, D.A.; Casewell, N.R. Novel transcripts in the maxillary venom glands of advanced snakes. *Toxicon* **2012**, *59*, 696–708. [CrossRef] [PubMed]
52. Kamiguti, A.S.; Theakston, R.D.; Sherman, N.; Fox, J.W. Mass spectrophotometric evidence for P-III/P-IV metalloproteinases in the venom of the Boomslang (*Dispholidus typus*). *Toxicon* **2000**, *38*, 1613–1620. [CrossRef]
53. Estrella, A.; Sánchez, E.E.; Galán, J.A.; Tao, W.A.; Guerrero, B.; Navarrete, L.F.; Rodríguez-Acosta, A. Characterization of toxins from the broad-banded water snake *Helicops angulatus* (Linnaeus, 1758): Isolation of a cysteine-rich secretory protein, Helicopsin. *Arch. Toxicol.* **2011**, *85*, 305–313. [CrossRef] [PubMed]
54. Peichoto, M.E.; Tavares, F.L.; Santoro, M.L.; Mackessy, S.P. Venom proteomes of South and North American opisthoglyphous (Colubridae and Dipsadidae) snake species: A preliminary approach to understanding their biological roles. *Comp. Biochem. Physiol. Part D* **2012**, *7*, 361–369.
55. Zhang, Z.; Zhang, X.; Hu, T.; Zhou, W.; Cui, Q.; Tian, J.; Zheng, Y.; Fan, Q. Discovery of toxin-encoding genes from the false viper *Macropisthodon rudis*, a rear-fanged snake, by transcriptome analysis of venom gland. *Toxicon* **2015**, *106*, 72–78. [CrossRef] [PubMed]
56. Heyborne, W.H.; Mackessy, S.P. Identification and characterization of a taxon-specific three-finger toxin from the venom of the Green Vinesnake (*Oxybelis fulgidus*; family Colubridae). *Biochimie* **2013**, *95*, 1923–1932. [CrossRef] [PubMed]
57. Campos, P.F.; Silva, D.A.; Zelanis, A.; Paes Leme, A.F.; Rocha, M.M.T.; Menezes, M.C.; Serrano, S.M.; Junqueira-de-Azevedo, I.L. Trends in the evolution of snake toxins underscored by an integrative omics approach to profile the venom of the colubrid *Phalotris mertensi*. *Genome Biol. Evol.* **2016**. [CrossRef] [PubMed]
58. Urra, F.A.; Pulgar, R.; Gutiérrez, R.; Hodar, C.; Cambiazo, V.; Labra, A. Identification and molecular characterization of five putative toxins from the venom gland of the snake *Philodryas chamissonis* (Serpentes: Dipsadidae). *Toxicon* **2015**, *108*, 19–31. [CrossRef] [PubMed]
59. Ching, A.T.; Rocha, M.M.; Paes Leme, A.F.; Pimenta, D.C.; de Furtado, M.F.; Serrano, S.M.; Ho, P.L.; Junqueira-de-Azevedo, I.L. Some aspects of the venom proteome of the Colubridae snake *Philodryas olfersii* revealed from a Duvernoy's (venom) gland transcriptome. *FEBS Lett.* **2006**, *580*, 4417–4422. [CrossRef] [PubMed]
60. Ching, A.T.; Paes Leme, A.F.; Zelanis, A.; Rocha, M.M.; de Furtado, M.F.; Silva, D.A.; Trugilho, M.R.; da Rocha, S.L.; Perales, J.; Ho, P.L.; et al. Venomics profiling of *Thamnodynastes strigatus* unveils matrix metalloproteinases and other novel proteins recruited to the toxin arsenal of rear-fanged snakes. *J. Proteome Res.* **2012**, *11*, 1152–1162. [CrossRef] [PubMed]
61. Huang, P.; Mackessy, S.P. Biochemical characterization of phospholipase A_2 (trimorphin) from the venom of the Sonoran Lyre Snake *Trimorphodon biscutatus* lambda (family Colubridae). *Toxicon* **2004**, *44*, 27–36. [CrossRef] [PubMed]
62. Wiens, J.J.; Kuczynski, C.A.; Smith, S.A.; Mulcahy, D.G.; Sites, J.W., Jr.; Townsend, T.M.; Reeder, T.W. Branch lengths, support, and congruence: Testing the phylogenomic approach with 20 nuclear loci in snakes. *Syst. Biol.* **2008**, *57*, 420–431. [CrossRef] [PubMed]
63. Taub, A.M. Ophidian cephalic glands. *J. Morphol.* **1966**, *118*, 529–542. [CrossRef] [PubMed]
64. Salomão, M.G.; Albolea, A.B.P.; Santos, S.M.A. Colubrid snakebite: A public health problem in Brazil. *Herpetol. Rev.* **2003**, *34*, 307–312.
65. Weinstein, S.A.; Griffin, R.; Ismail, A.K. Non-front-fanged colubroid ("colubrid") snakebites: Three cases of local envenoming by the mangrove or ringed cat-eyed snake (*Boiga dendrophila*; Colubridae, Colubrinae), the Western beaked snake (*Rhamphiophis oxyrhynchus*; Lamprophiidae, Psammophinae) and the rain forest cat-eyed snake (*Leptodeira frenata*; Dipsadidae). *Clin. Toxicol. (Phila.)* **2014**, *52*, 277–282. [PubMed]
66. Komori, K.; Konishi, M.; Maruta, Y.; Toriba, M.; Sakai, A.; Matsuda, A.; Hori, T.; Nakatani, M.; Minamino, N.; Akizawa, T. Characterization of a novel metalloproteinase in Duvernoy's gland of *Rhabdophis tigrinus tigrinus*. *J. Toxicol. Sci.* **2006**, *31*, 157–168. [CrossRef] [PubMed]

67. Modahl, C.M.; Mackessy, S.P. Full-length venom protein cDNA sequences from venom-derived mRNA: Exploring compositional variation and adaptive multigene evolution. *PLoS Negl. Trop. Dis.* **2016**, *10*, e0004587. [CrossRef] [PubMed]
68. Moura-da-Silva, A.M.; Theakston, R.D.; Crampton, J.M. Evolution of disintegrin cysteine-rich and mammalian matrix-degrading metalloproteinases: Gene duplication and divergence of a common ancestor rather than convergent evolution. *J. Mol. Evol.* **1996**, *43*, 263–269. [CrossRef] [PubMed]
69. Casewell, N.R.; Wagstaff, S.C.; Harrison, R.A.; Renjifo, C.; Wüster, W. Domain loss facilitates accelerated evolution and neofunctionalization of duplicate snake venom metalloproteinase toxin genes. *Mol. Biol. Evol.* **2011**, *28*, 2637–2649.
70. Vaiyapuri, S.; Sunagar, K.; Gibbins, J.M.; Jackson, T.N.W.; Reeks, T.; Fry, B.G. Kallikrein enzymes. In *Venomous Reptiles and Their Toxins: Evolution, Pathophysiology and Biodiscovery*, 1st ed.; Fry, B.G., Ed.; Oxford University Press: Oxford, UK, 2015; p. 576.
71. Weinstein, S.A.; Kardong, K.V. Properties of Duvernoy's secretions from opisthoglyphous and aglyphous colubrid snakes. *Toxicon* **1994**, *32*, 1161–1185. [CrossRef]
72. Vest, D.K.; Mackessy, S.P.; Kardong, K.V. The unique Duvernoy's secretion of the brown tree snake (*Boiga irregularis*). *Toxicon* **1991**, *29*, 532–535. [CrossRef]
73. Hill, R.E.; Mackessy, S.P. Characterization of venom (Duvernoy's secretion) from twelve species of colubrid snakes and partial sequence of four venom proteins. *Toxicon* **2000**, *38*, 1663–1687. [CrossRef]
74. Reyes-Velasco, J.; Card, D.C.; Andrew, A.L.; Shaney, K.J.; Adams, R.H.; Schield, D.R.; Casewell, N.R.; Mackessy, S.P.; Castoe, T.A. Expression of venom gene homologs in diverse python tissues suggests a new model for the evolution of snake venom. *Mol. Biol. Evol.* **2015**, *32*, 173–183. [CrossRef] [PubMed]
75. Drickamer, K. C-type lectin-like domains. *Curr. Opin. Struct. Biol.* **1999**, *9*, 585–590. [CrossRef]
76. Arlinghaus, F.T.; Fry, B.G.; Sunagar, K.; Jackson, T.N.W.; Eble, J.A.; Reeks, T.; Clemetson, K.J. Lectin proteins. In *Venomous Reptiles and Their Toxins: Evolution, Pathophysiology and Biodiscovery*, 1st ed.; Fry, B.G., Ed.; Oxford University Press: Oxford, UK, 2015; p. 576.
77. Junqueira-de-Azevedo, I.L.; Ho, P.L. A survey of gene expression and diversity in the venom glands of the pitviper snake *Bothrops insularis* through the generation of expressed sequence tags (ESTs). *Gene* **2002**, *299*, 279–291. [CrossRef]
78. Jackson, T.N.; Sunagar, K.; Undheim, E.A.; Koludarov, I.; Chan, A.H.; Sanders, K.; Ali, S.A.; Hendrikx, I.; Dunstan, N.; Fry, B.G. Venom down under: Dynamic evolution of Australian elapid snake toxins. *Toxins* **2013**, *5*, 2621–2655. [CrossRef] [PubMed]
79. Peichoto, M.E.; Mackessy, S.P.; Teibler, P.; Tavares, F.L.; Burckhardt, P.L.; Breno, M.C.; Acosta, O.; Santoro, M.L. Purification and characterization of a cysteine-rich secretory protein from *Philodryas patagoniensis* (Dipsadidae) snake venom. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* **2009**, *150*, 79–84. [CrossRef] [PubMed]
80. Sunagar, K.; Johnson, W.E.; O'Brien, S.J.; Vasconcelos, V.; Antunes, A. Evolution of CRISPs associated with toxicoferan-reptilian venom and mammalian reproduction. *Mol. Biol. Evol.* **2012**, *29*, 1807–1822. [CrossRef] [PubMed]
81. Ondetti, M.A.; Rubin, B.; Cushman, D.W. Design of specific inhibitors of angiotensin-converting enzyme: New class of orally active anti-hypertensive agents. *Science* **1977**, *196*, 441–444. [CrossRef] [PubMed]
82. Munekiyo, S.M.; Mackessy, S.P. Presence of peptide inhibitors in rattlesnake venoms and their effects on endogenous metalloproteases. *Toxicon* **2005**, *45*, 255–263. [CrossRef] [PubMed]
83. Higuchi, S.; Murayama, N.; Saguchi, K.; Ohi, H.; Fujita, Y.; da Silva, N.J., Jr.; de Siqueira, R.J.; Lahlou, S.; Aird, S.D. A novel peptide from the ACEI/BPP-CNP precursor in the venom of *Crotalus durissus collilineatus*. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* **2006**, *144*, 107–121. [CrossRef] [PubMed]
84. Wagstaff, S.C.; Favreau, P.; Cheneval, O.; Laing, G.D.; Wilkinson, M.C.; Miller, R.L.; Stöcklin, R.; Harrison, R.A. Molecular characterisation of endogenous snake venom metalloproteinase inhibitors. *Biochem. Biophys. Res. Commun.* **2008**, *365*, 650–656. [CrossRef] [PubMed]
85. Graham, R.L.; Graham, C.; McClean, S.; Chen, T.; O'Rourke, M.; Hirst, D.; Theakston, D.; Shaw, C. Identification and functional analysis of a novel bradykinin inhibitory peptide in the venoms of New World Crotalinae pit vipers. *Biochem. Biophys. Res. Commun.* **2005**, *338*, 1587–1592. [CrossRef] [PubMed]
86. Correa, P.G.; Oguiura, N. Phylogenetic analysis of β-defensin-like genes of *Bothrops*, *Crotalus* and *Lachesis* snakes. *Toxicon* **2013**, *69*, 65–74. [CrossRef] [PubMed]

87. Doley, R.; Pahari, S.; Reza, M.A.; Mackessy, S.P.; Kini, K.M. The gene structure and evolution of ku-wap-fusin (Kunitz Waprin Fusion Protein), a novel evolutionary intermediate of the Kunitz serine protease inhibitors and waprins from *Sistrurus catenatus* (Massasauga Rattlesnake) venom glands. *Open Evol. J.* **2010**, *4*, 31–41.
88. Laskowski, M., Jr.; Kato, I. Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **1980**, *49*, 593–626. [CrossRef] [PubMed]
89. Millers, E.K.; Trabi, M.; Masci, P.P.; Lavin, M.F.; de Jersey, J.; Guddat, L.W. Crystal structure of textilinin-1, a Kunitz-type serine protease inhibitor from the venom of the Australian common brown snake (*Pseudonaja textilis*). *FEBS J.* **2009**, *276*, 3163–3175. [CrossRef] [PubMed]
90. Rao, V.S.; Joseph, J.S.; Kini, R.M. Group D prothrombin activators from snake venom are structural homologues of mammalian blood coagulation factor Xa. *Biochem. J.* **2003**, *369*, 635–642. [CrossRef] [PubMed]
91. Junqueira-de-Azevedo, I.L.; Farsky, S.H.; Oliveira, M.L.; Ho, P.L. Molecular cloning and expression of a functional snake venom vascular endothelium growth factor (VEGF) from the *Bothrops insularis* pit viper. A new member of the VEGF family of proteins. *J. Biol. Chem.* **2001**, *276*, 39836–39842. [CrossRef] [PubMed]
92. Takahashi, H.; Hattori, S.; Iwamatsu, A.; Takizawa, H.; Shibuya, M. A novel snake venom vascular endothelial growth factor (VEGF) predominantly induces vascular permeability through preferential signaling via VEGF receptor-1. *J. Biol. Chem.* **2004**, *279*, 46304–46314. [CrossRef] [PubMed]
93. Vogel, C.W.; Bredehorst, R.; Fritzinger, D.C.; Grunwald, T.; Ziegelmüller, P.; Kock, M.A. Structure and function of cobra venom factor, the complement-activating protein in cobra venom. *Adv. Exp. Med. Biol.* **1996**, *391*, 97–114. [PubMed]
94. Okumura, K.; Masui, K.; Inoue, S.; Ikeda, K.; Hayashi, K. Purification, characterization and cDNA cloning of a phospholipase A_2 inhibitor from the serum of the non-venomous snake *Elaphe quadrivirgata*. *Biochem. J.* **1999**, *341*, 165–171. [CrossRef] [PubMed]
95. Brillard-Bourdet, M.; Nguyên, V.; Ferrer-di Martino, M.; Gauthier, F.; Moreau, T. Purification and characterization of a new cystatin inhibitor from Taiwan cobra (*Naja naja atra*) venom. *Biochem. J.* **1998**, *331*, 239–244. [CrossRef] [PubMed]
96. Ducancel, F.; Matre, V.; Dupont, C.; Lajeunesse, E.; Wollberg, Z.; Bdolah, A.; Kochva, E.; Boulain, J.C.; Ménez, A. Cloning and sequence analysis of cDNAs encoding precursors of sarafotoxins. Evidence for an unusual "rosary-type" organization. *J. Biol. Chem.* **1993**, *268*, 3052–3055. [PubMed]
97. Casewell, N.R.; Harrison, R.A.; Wüster, W.; Wagstaff, S.C. Comparative venom gland transcriptome surveys of the saw-scaled vipers (Viperidae: *Echis*) reveal substantial intra-family gene diversity and novel venom transcripts. *BMC Genom.* **2009**, *10*. [CrossRef] [PubMed]
98. Corrêa-Netto, C.; Junqueira-de-Azevedo, I.L.; Silva, D.A.; Ho, P.L.; Leitão-de-Araújo, M.; Alves, M.L.; Sanz, L.; Foguel, D.; Zingali, R.B.; Calvete, J.J. Snake venomics and venom gland transcriptomic analysis of Brazilian coral snakes, *Micrurus altirostris* and *M. corallinus*. *J. Proteom.* **2011**, *74*, 1795–1809. [CrossRef] [PubMed]
99. Fry, B.G.; Roelants, K.; Norman, J.; King, G.; Tyndal, J.; Lewis, R.; Norton, R.; Renjifo, C.; Rodriguez de la Vega, R.C. Toxicogenomic multiverse: Convergent recruitment of proteins into animal venoms. *Annu. Rev. Genom. Hum. Genet.* **2009**, *10*, 483–511. [CrossRef] [PubMed]
100. Lee, K.H.; Wells, R.G.; Reed, R.R. Isolation of an olfactory cDNA: Similarity to retinol-binding protein suggests a role in olfaction. *Science* **1987**, *235*, 1053–1056. [CrossRef] [PubMed]
101. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **2005**, *21*, 3674–3676. [CrossRef] [PubMed]
102. Mortazavi, A.; Williams, B.A.; McCue, K.; Schaeffer, L.; Wold, B. Mapping and quantifying mammalian transcriptomes by rna-seq. *Nat. Methods* **2008**, *5*, 621–628. [CrossRef] [PubMed]
103. Huang, X.; Madan, A. CAP3: A DNA sequence assembly program. *Genome Res.* **1999**, *9*, 868–877. [CrossRef] [PubMed]
104. Junqueira-de-Azevedo, I.L.; Ching, A.T.; Carvalho, E.; Faria, F.; Nishiyama, M.Y., Jr.; Ho, P.L.; Diniz, M.R. *Lachesis muta* (Viperidae) cDNAs reveal diverging pit viper molecules and scaffolds typical of cobra (Elapidae) venoms: implications for snake toxin repertoire evolution. *Genetics* **2006**, *173*, 877–889. [CrossRef] [PubMed]

© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).