SUPPORTING INFORMATION

TITLE:

Medically important differences in snake venom composition are dictated by distinct postgenomic mechanisms

AUTHORS:

Nicholas R. Casewell, Simon C. Wagstaff, Wolfgang Wüster, Darren A.N. Cook, Fiona M.S. Bolton, Sarah I. King, Davinia Pla, Libia Sanz, Juan J. Calvete, Robert A. Harrison

ABSTRACT:

Variation in venom composition is a ubiquitous phenomenon in snakes and occurs both interspecifically and intraspecifically. Venom variation can have severe outcomes for snakebite victims by rendering the specific antibodies found in antivenoms ineffective against heterologous toxins found in different venoms. The rapid evolutionary expansion of different toxin-encoding gene families in different snake lineages is widely perceived as the main cause of venom variation. However, this view is simplistic and disregards the understudied influence that processes acting on gene transcription and translation may have on the production of the venom proteome. Here, we assess the venom composition of six related viperid snakes and compare interspecific changes in the number of toxin genes, their transcription in the venom gland and their translation into proteins secreted in venom. Our results reveal that multiple levels of regulation are responsible for generating variation in venom composition between related snake species. We demonstrate that differential levels of toxin transcription, translation and their posttranslational modification have a substantial impact upon the resulting venom protein mixture. Notably, these processes act to varying extents on different toxin paralogs found in different snakes, and are therefore likely to be as important as ancestral gene duplication events for generating compositionally distinct venom proteomes. Our results suggest that these processes may also contribute to altering the toxicity of snake venoms, and we demonstrate how this variability can undermine the treatment of a neglected tropical disease, snakebite.

1. SI MATERIALS AND METHODS

1.1. Venom gland transcriptomes

The construction of venom gland transcriptomes for four saw-scaled viper species [*E. ocellatus* (Nigeria), *E. p. leakeyi* (Kenya), *E. coloratus* (Egypt) and *E. c. sochureki* (United Arab Emirates)] have previously been described (1, 2). Here, we constructed complementary venom gland transcriptomes for *B. arietans* (Nigeria) and *C. cerastes* (Egypt) in an identical manner for the purposes of direct comparison. Briefly, the libraries were constructed by using mRNA extracted from venom glands pooled from 10 individuals of each species by using the CloneMiner method, with ~1000 clones sequenced by using Sanger sequencing for each library. Ten individuals were used for each species to provide a representative transcriptome that accounts for any intra-specific variation in venom toxin expression. All venom gland material was dissected 3 d after venom extraction, and each transcriptome therefore represents a snap-shot of toxin transcription at this time point, where transcription during venom replenishment is at its peak (3). Generated ESTs were subsequently assembled into contigs (putative gene products) and annotated as described (1, 2). The number of ESTs contributing to each contig was used as a measure of transcription. Assembled contigs and individual ESTs were then subjected to six frame translations and used as reference databases to facilitate proteomic identification.

1.2. Venom proteomes

Proteins from each crude, lyophilized venom (2 mg extracted from the same individuals used for venom gland transcriptomics) were separated by reverse-phase HPLC as previously described (4). Fractions were collected manually, dried in a vacuum centrifuge (Savant), redissolved in water, and submitted to SDS-PAGE analysis (under non-reducing and reducing conditions) with Coomassie brilliant blue R-250 staining. Isolated RP-HPLC fractions were subjected to N-terminal sequencing and molecular mass determination (4). For MS/MS analyses, protein bands were excised and subjected to in-gel reduction and alkylation, and trypsin digestion (4). Tryptic peptide mass fingerprints were recorded with an Applied Biosystems' 5800 MALDI-TOF-TOF™ instrument. Up to 50 precursor ions were submitted to collision-induced dissociation (CID) MS/MS analysis. For ESI-CID-MS/MS, each tryptic peptide digest was either loaded in a nanospray capillary column and subjected to nanoelectrospray ionization (ESI) mass spectrometric analysis using a QTrap™ 2000 mass spectrometer (Applied Biosystems) or submitted to nano-Acquity UltraPerformance uPLC separation in-line with a Waters SYNAPT G2 HD mass spectrometry system. Production spectra of doubly- and triply-charged ions were interpreted manually or searched against the National Center for Biotechnology Information and UniProt/SwissProt databases, and against each species-specific venom gland transcriptome dataset by using ProteinPilot v.4 and the Paragon® algorithm (ABSciex) at ≥95% confidence (MALDI-TOF-TOF spectra); the on-line form of MASCOT (nESI-MS/MS spectra); or processed in ProteinLynx Global SERVER 2013 version 2.5.2. MS/MS

mass tolerance was set to ± 0.6 Da. Carbamidomethyl cysteine and oxidation of methionine were selected as fixed and variable modifications, respectively. The relative abundances (expressed as percentage of the total venom proteins) of the different protein families were calculated from the relation of the sum of the areas of the reverse-phase chromatographic peaks (containing proteins from the same family), to the total area of venom protein peaks in the reverse-phase chromatogram. The relative contributions of different proteins eluting in the same chromatographic fraction was estimated by densitometry of Coomassie brilliant blue-stained SDS-PAGE gels, as previously outlined (4).

1.3. Toxin evolution

Toxin gene sequences annotated as SVMP, PLA_2, CTL, SP, LAAO or CRISP in the transcriptomes were extracted and analysed to reconstruct their evolutionary history. These gene families were selected based on the results of the proteomic analyses, which identified the presence of these toxin types in the venom of the majority of the sampled species. The remaining toxin family identified and analysed, the short-coding disintegrins (DIS), were discarded from phylogenetic analysis due to their apparent convergent evolution from SVMPs (5, 6). For each toxin family, non-redundant nucleotide sequences from each of the six transcriptomes were aligned with published gene homologs isolated from the venom systems of other viperid snakes, using the MUSCLE algorithm (7). Where available, full-length primer-walked toxins sequences generated by our previous work on these species were incorporated into the alignments (5, 8). Phylogenetic analyses for each toxin family were performed by incorporating optimised models of sequence evolution (Table S8), selected by the Akaike information criterion in MrModelTest (http://www.abc.se/~nylander/mrmodeltest2/mrmodeltest2.html), into MrBayes v3.2 (9). Nucleotide gene trees were generated in duplicate using four chains for $1x10^7$ generations, sampling every 500^{th} cycle from the chain and using default settings in regards to priors. Tracer v1.4 (10) was used to estimate effective sample sizes for all parameters and to verify the point of convergence (burnin), with trees generated prior to this point discarded. The resulting sequence alignments have been submitted to the Dryad data repository (doi:10.5061/dryad.1j292).

1.4. Gene and protein comparisons

Proteomic matches to transcriptome gene products were overlaid on to the generated gene trees, alongside calculations of protein abundance. We applied the following rules to the data: (*i*) We conservatively retained the longest nucleotide sequence for each transcriptome contig and discarded all other sequence variants. These sequences are unlikely to represent distinct genes, but rather are the result of allelic variation (as ten specimens were used for library construction). We removed these sequences to prevent artificial inflation of gene numbers generated by the reconciliation analyses described below. (*ii*) Where different contigs from the same species exhibited monophyly and did not result in distinct proteome matches, we merged the contigs and

discarded non-matching sequences due to sequence redundancy. (*iii*) Sequences from contigs that were non-monophyletic with other contigs from the same species and had no proteomic match were retained and annotated with 'no proteomic hit' - indicative of non-translation of the transcript. (*iv*) Post-translational modifications (i.e. proteolytic cleavage) were assigned where multiple proteomic matches to the same transcriptome cluster were found in distinct non-overlapping proteomic fractions.

1.5. Gene/species tree reconciliation

We reconciled the resulting gene trees with the species tree presented in Fig. 1 using NOTUNG.v2.6 (11). Each *Echis*/*Bitis*/*Cerastes* toxin clade was analysed separately using the reconciliation option, which produces a reconciled tree displaying the timing of gene duplication and loss events, and the predicted number of gene loci. As NOTUNG permits the reconciliation of genes trees to non-binary species trees (11), we utilised this option when clade support in the gene tree exhibited a Bayesian posterior probability of <0.95. This approach allowed nodes that were not robustly supported in the gene tree to be reconciled with a polytomous species tree. We used this conservative approach to mitigate the generation of spurious gene duplication events that would otherwise be generated as artefacts of uncertainty in the gene tree.

1.6. Functional studies

The same venom samples used for proteomic analyses were used for functional assessments. All animal experimentation was conducted using standard protocols approved by the Liverpool School of Tropical Medicine Animal Welfare and Ethical Review Board and the UK Home Office (licence # 40/3216, 40/3151, 40/3718). Murine *in vivo* lethality studies were conducted to calculate the Lethal Dose 50 (LD_{50} – the amount of venom that kills 50% of injected mice) for *B. arietans* and *C. cerastes* venom. Experiments were undertaken as previously described (12), with groups of five male CD-1 mice (18-20 g) receiving varying doses of venom (in 100 µL PBS) intravenously. The number of mice surviving after seven hours was recorded and the venom LD_{50} and 95% confidence limits calculated by probit analysis (13). Comparisons with our previous results for the four *Echis* species (12) were undertaken using the 95% confidence limits. LD_{50} values for venom lethality to scorpions and locusts were reproduced from our earlier work (14, 15). We next tested the efficacy of the *E. ocellatus* monospecific antivenom, EchiTAbG, at neutralising five times the venom LD_{50} for *B. arietans* and *C. cerastes* in the same murine model. These effective dose 50 (ED_{50}) assays were undertaken as previously described for the four *Echis* species (12), with groups of mice receiving the maximum amount of antivenom permitted by the assay (150 µL). The antivenom was mixed with five venom LD_{50} doses in 50 µL PBS, pre-incubated at 37°C for 30 mins, injected intravenously into five mice and survival recorded seven hours later. The ED_{50} and 95% confidence limits used for comparisons among species were calculated by probit analysis (13). Modified minimum haemorrhagic dose (MHD) experiments (16) were undertaken to compare

the haemorrhagic activity of each venom. Ten microgram doses of venom were injected intradermally into the shaved dorsal skin of groups of six male CD-1 mice (18-20 g) under light anaesthesia (three mice for *C. cerastes*). After 24 hours the dorsal skin was removed and the size of the lesion on the inner surface of the skin measured in two directions at right angles using callipers and background illumination. The mean diameter of each lesion was calculated and comparative statistical analysis undertaken using paired two-tailed t-tests with a *P*-value threshold of 0.05. The procoagulant activity of each venom was determined using the minimum coagulant dose (MCD-P) assay (16). Varying doses of venom were incubated with 200 µL of human or mouse plasma (Sigma) at 37°C and the clotting time recorded in triplicate. The MCD was calculated by plotting clotting time against venom dose and calculating the 60-s clotting time from the equation of the line. Statistical comparisons were undertaken using regression analysis of the plotted lines with a *P*-value threshold of 0.05.

2. SI FIGURES

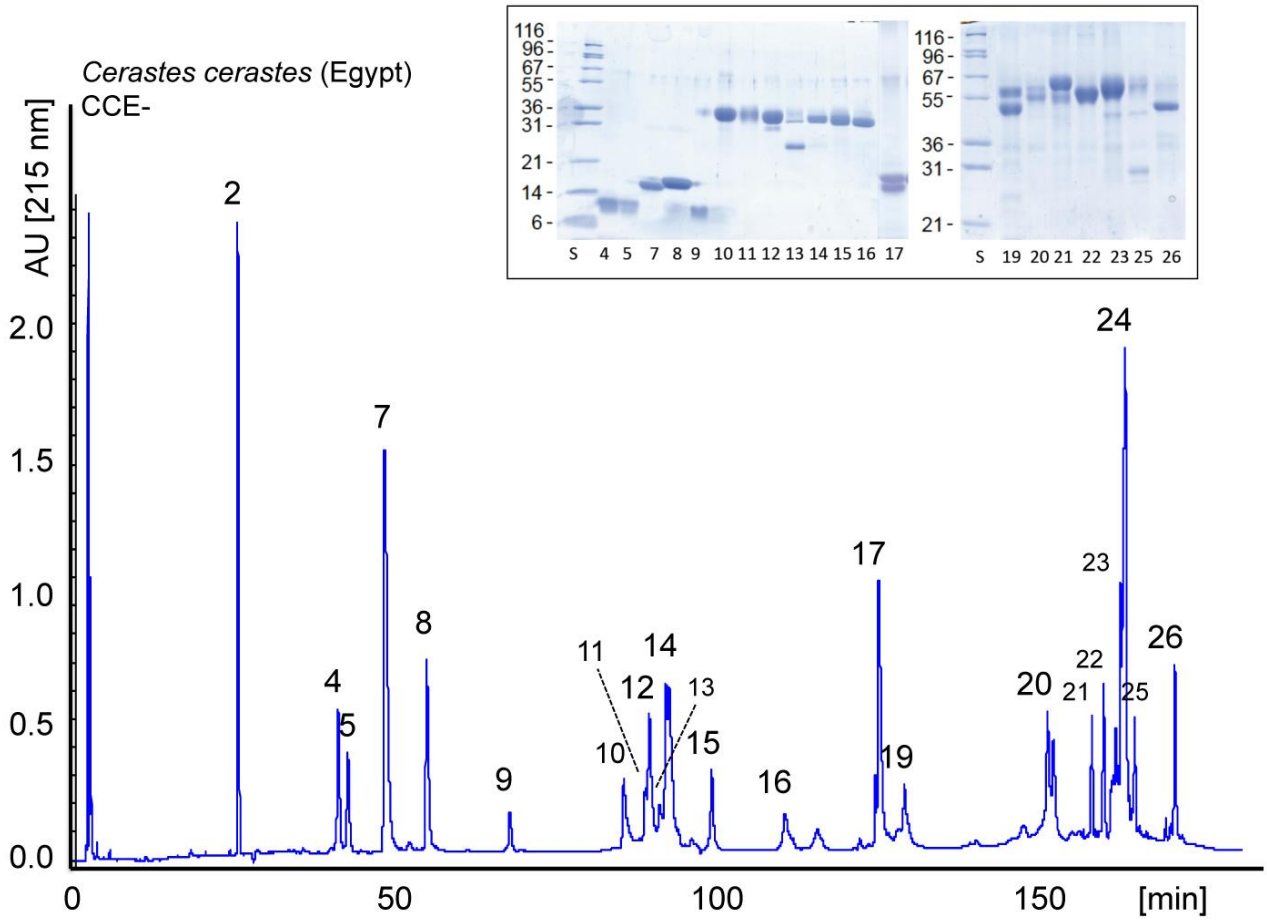

Fig. S1. Reverse-phase HPLC separation of *Cerastes cerastes* (Egypt) venom proteins. Insert, SDS-PAGE analysis of the isolated, numbered, protein fractions.

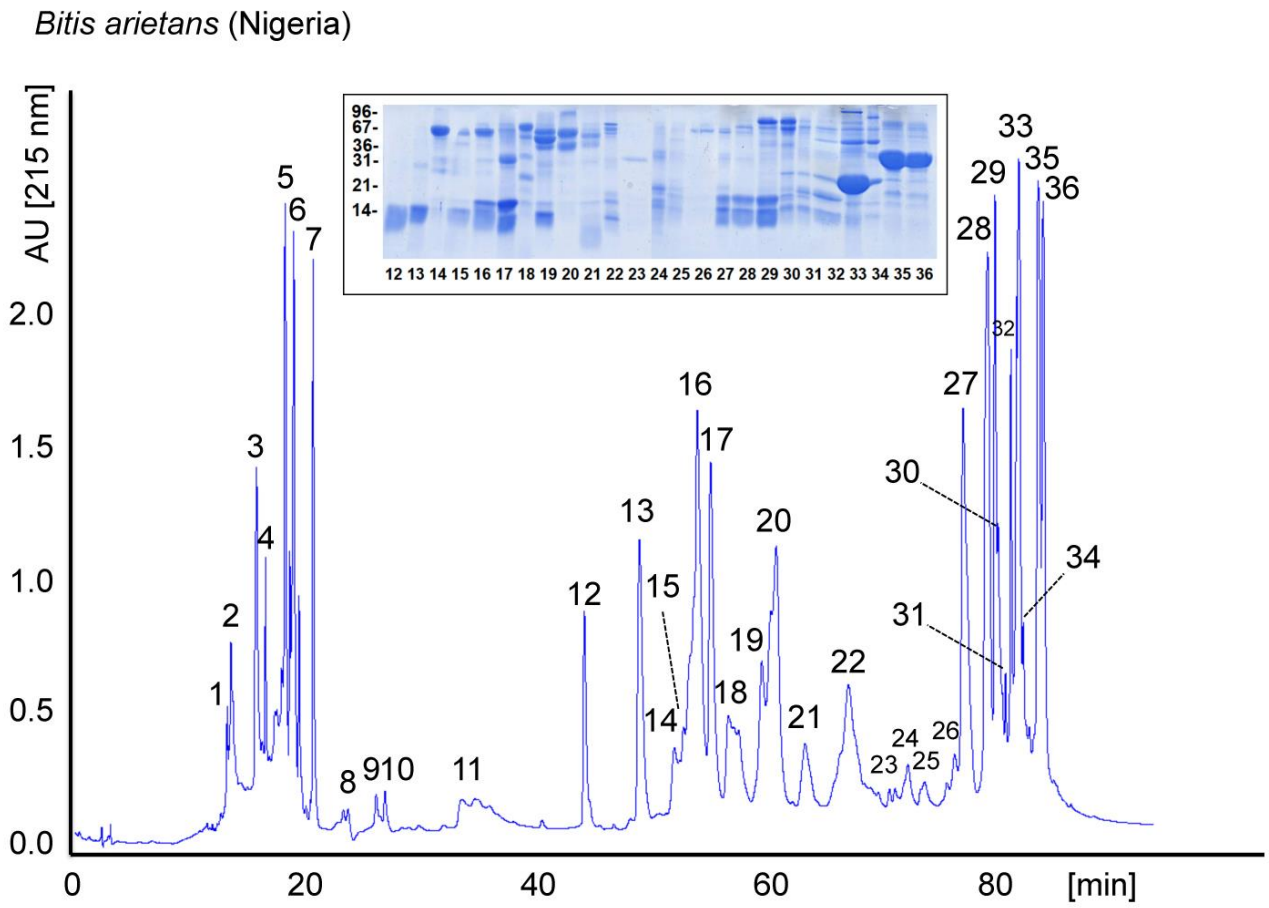

Fig. S2. Reverse-phase HPLC separation of *Bitis arietans* (Nigeria) venom proteins. Insert, SDS-PAGE analysis of the isolated, numbered, protein fractions.

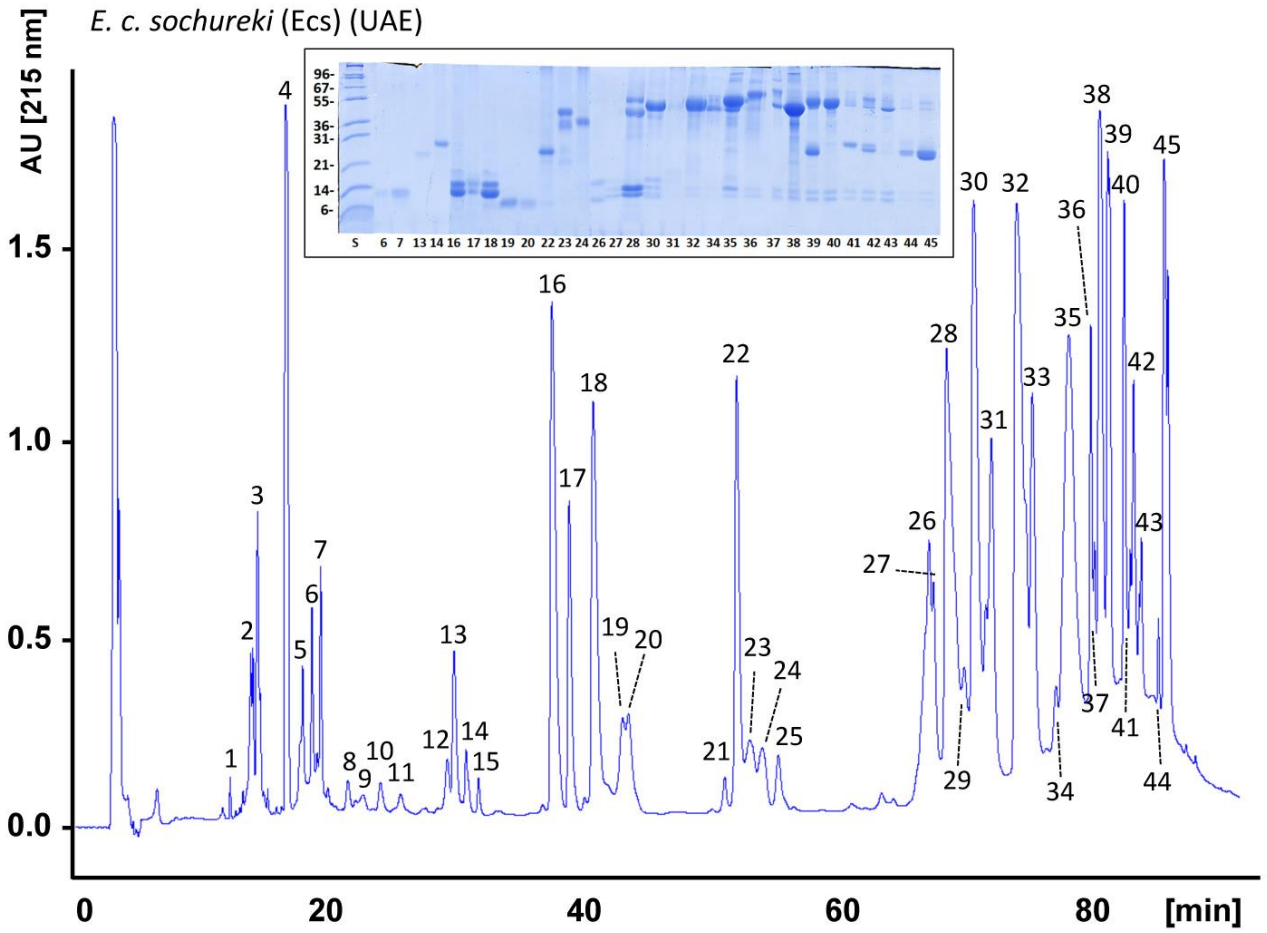

Fig. S3. Reverse-phase HPLC separation of *Echis carinatus sochureki* (UAE) venom proteins. Insert, SDS-PAGE analysis of the isolated, numbered, protein fractions.

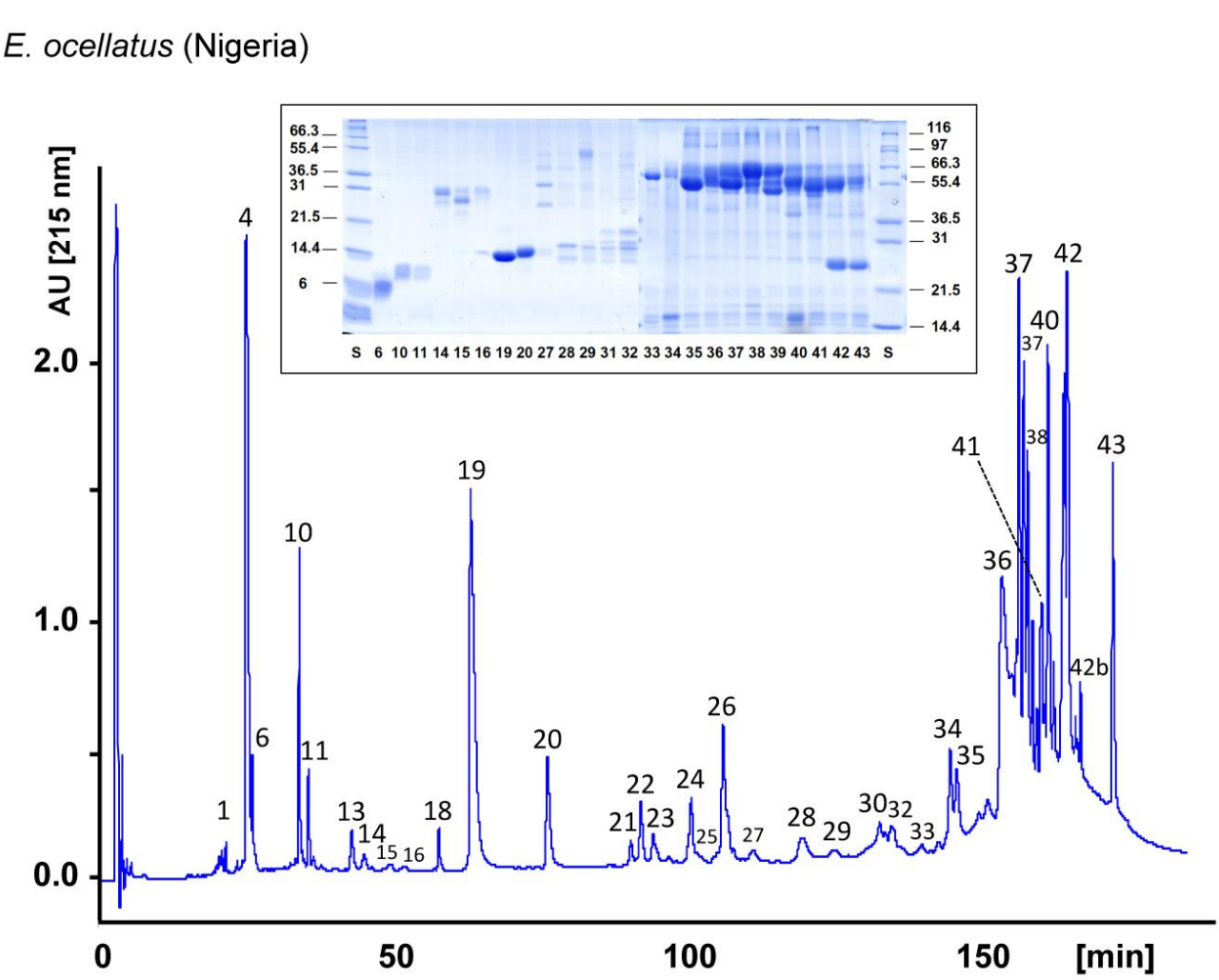

Fig. S4. Reverse-phase HPLC separation of *Echis ocellatus* (Nigeria) venom proteins. Insert, SDS-PAGE analysis of the isolated, numbered, protein fractions.

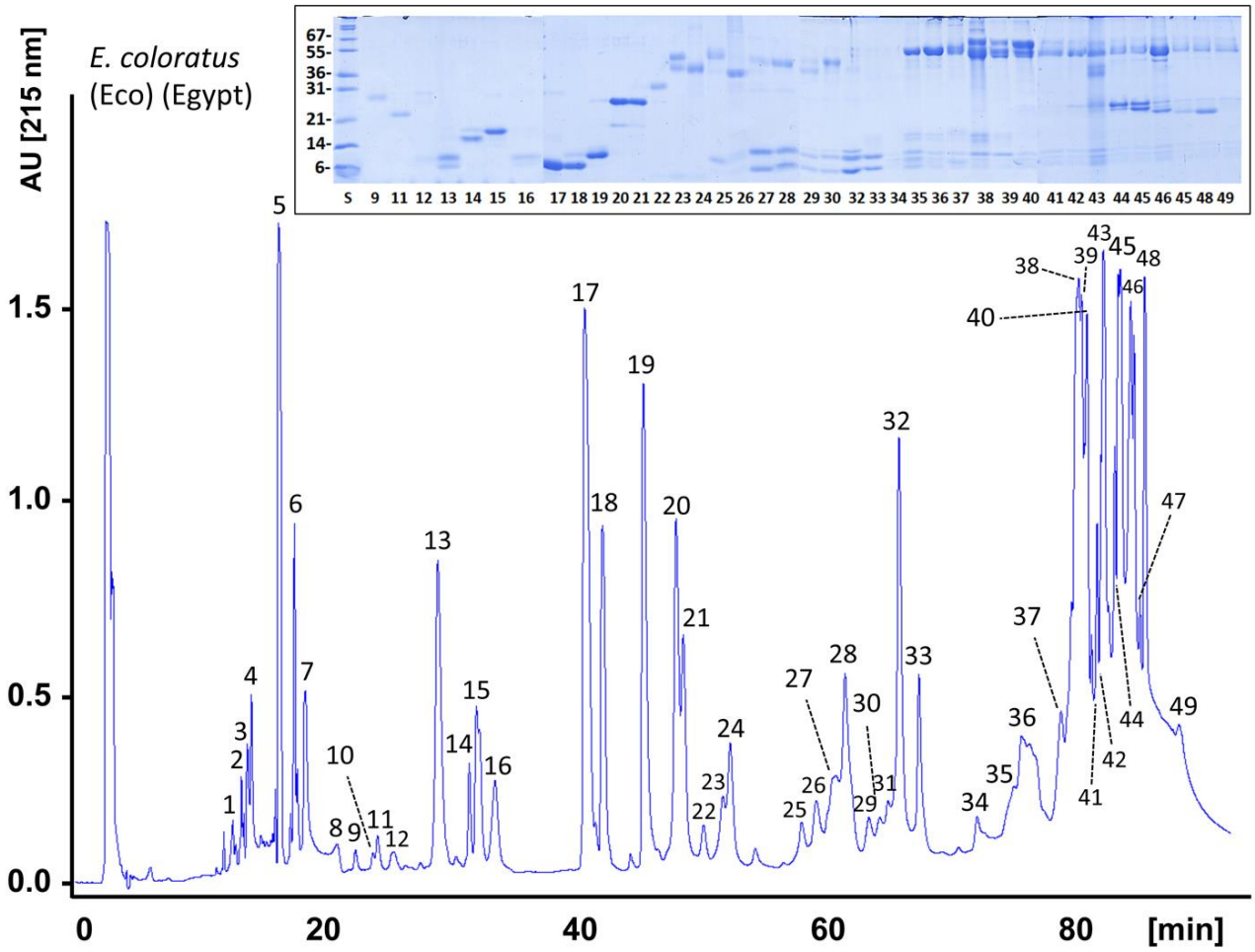

Fig. S5. Reverse-phase HPLC separation of *Echis coloratus* (Egypt) venom proteins. Insert, SDS-PAGE analysis of the isolated, numbered, protein fractions.

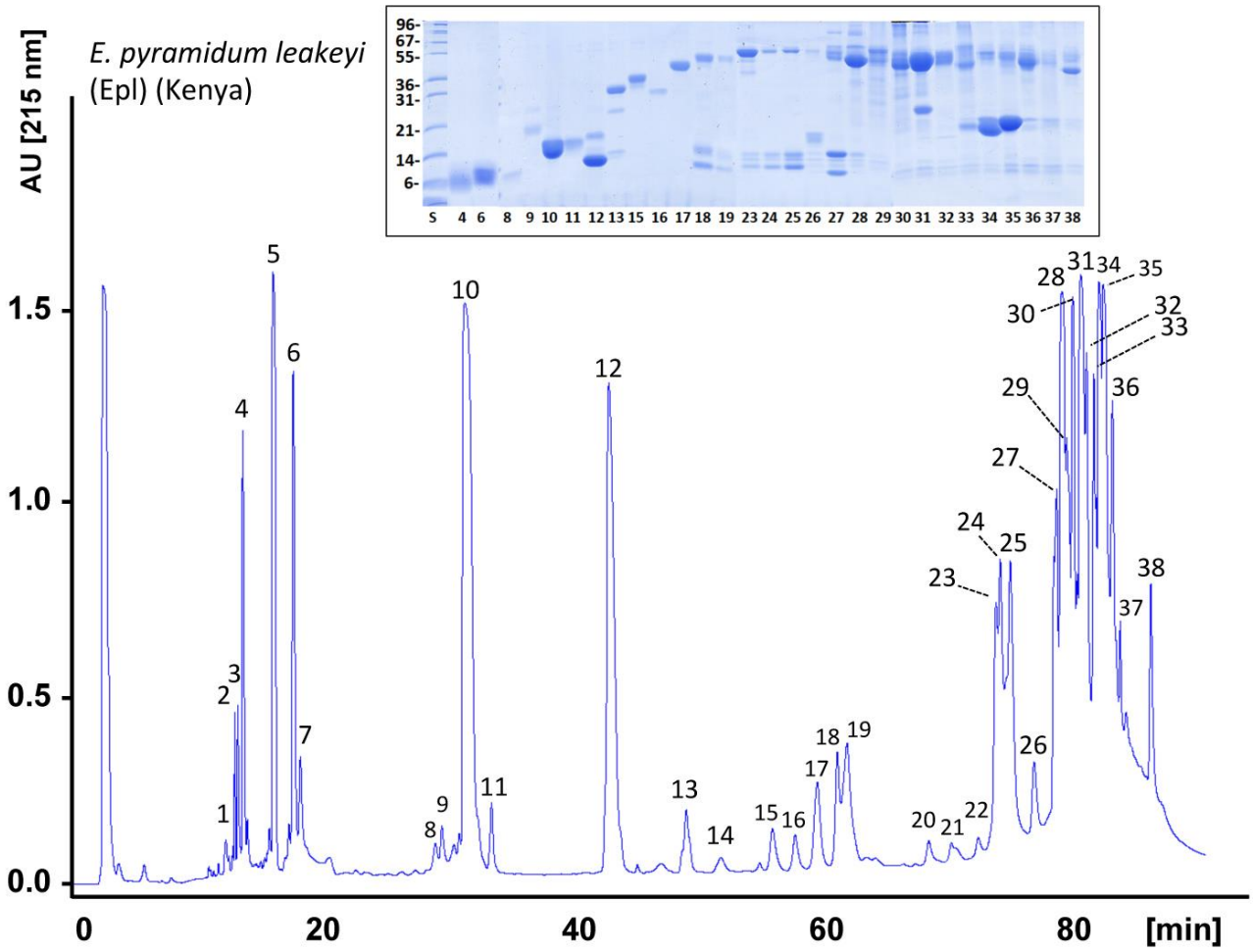

Fig. S6. Reverse-phase HPLC separation of *Echis pyramidum leakeyi* (Kenya) venom proteins. Insert, SDS-PAGE analysis of the isolated, numbered, protein fractions.

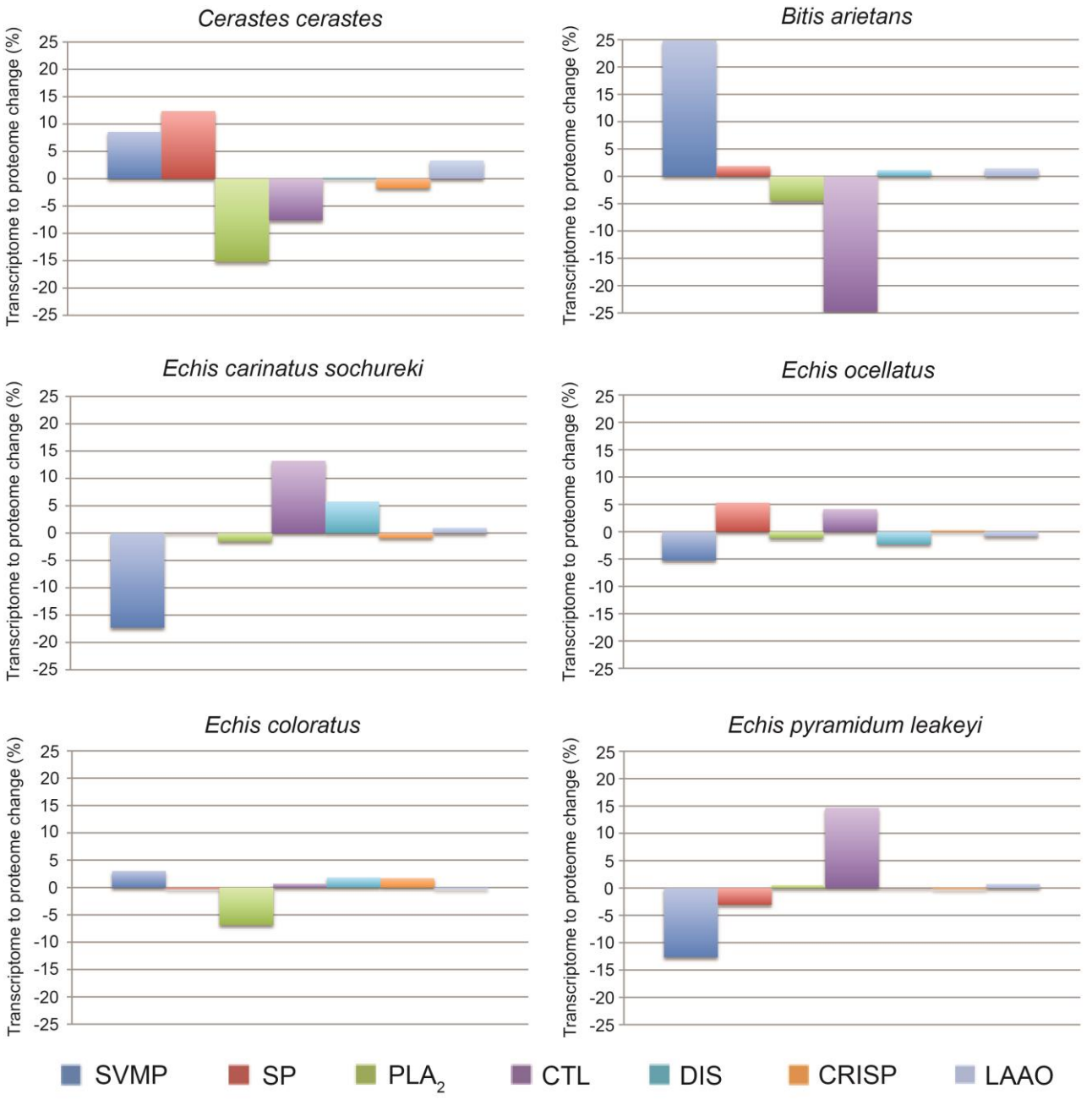

Fig. S7. The percentage change from gene expression to protein abundance for each toxin family identified in each species. Gene expression is calculated from venom gland transcriptome data and protein abundance from venom proteome data. Coloured bars represent the summation of net change in expression/abundance of all toxins encoded by each gene family for each of the species.

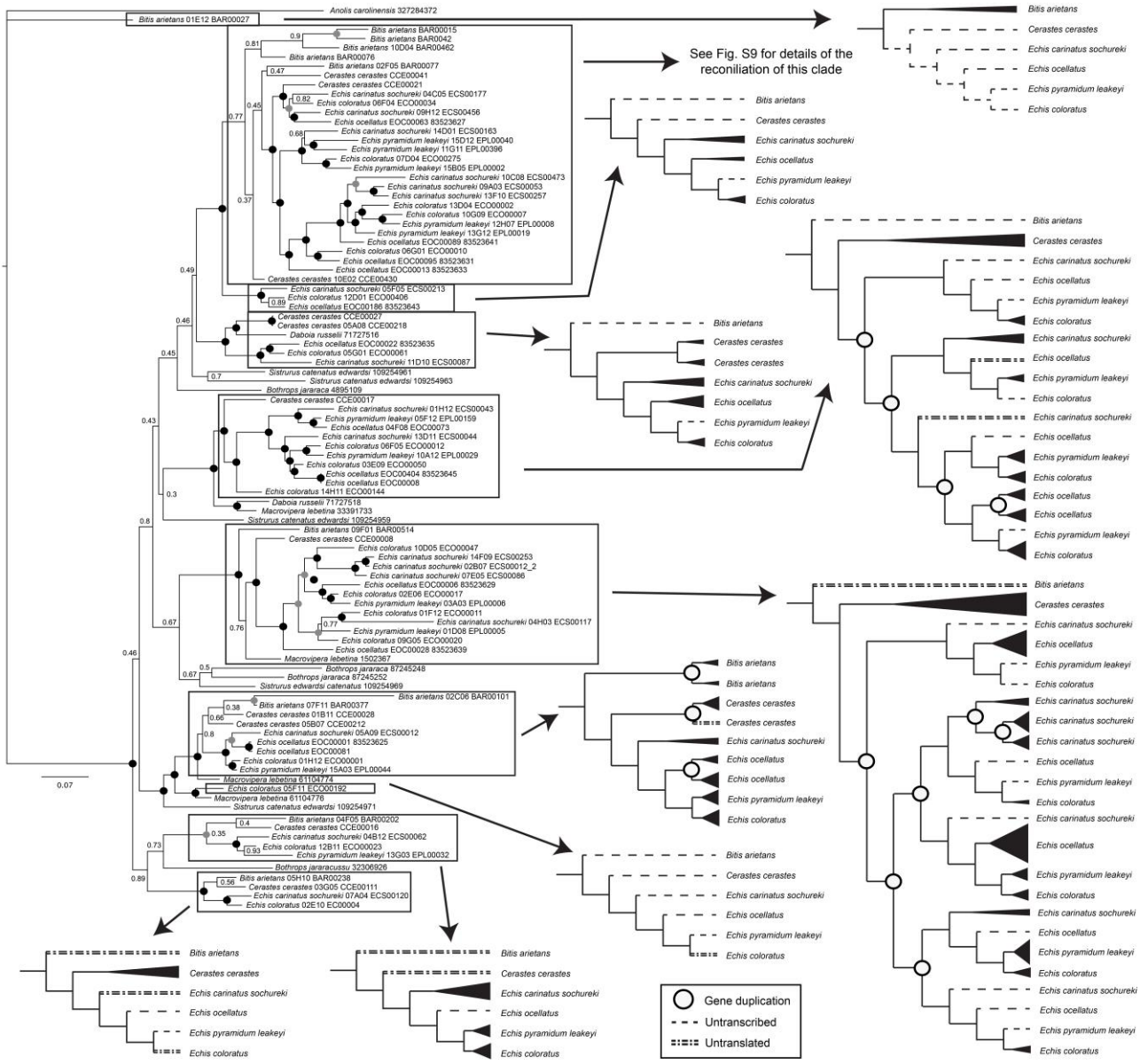

Fig. S8. The evolutionary history of the snake venom metalloproteinase (SVMP) toxin family. The gene tree on the left displays the reconstructed evolutionary history for the gene family. Black dots represent Bayesian posterior probabilities of 1.00; grey dots represent >0.95. Black boxes highlight clades of the gene tree containing *Bitis*, *Cerastes* and/or *Echis* toxins. Arrows next to each clade show the gene tree reconciled to the species phylogeny. The width of each branch in the reconciled tree represents the proteomic abundance of each toxin detected in each species. Bordered white circles represent inferred gene duplication events; dotted branches indicated inferred untranscribed loci; double dotted branches indicate genes that were expressed but untranslated.

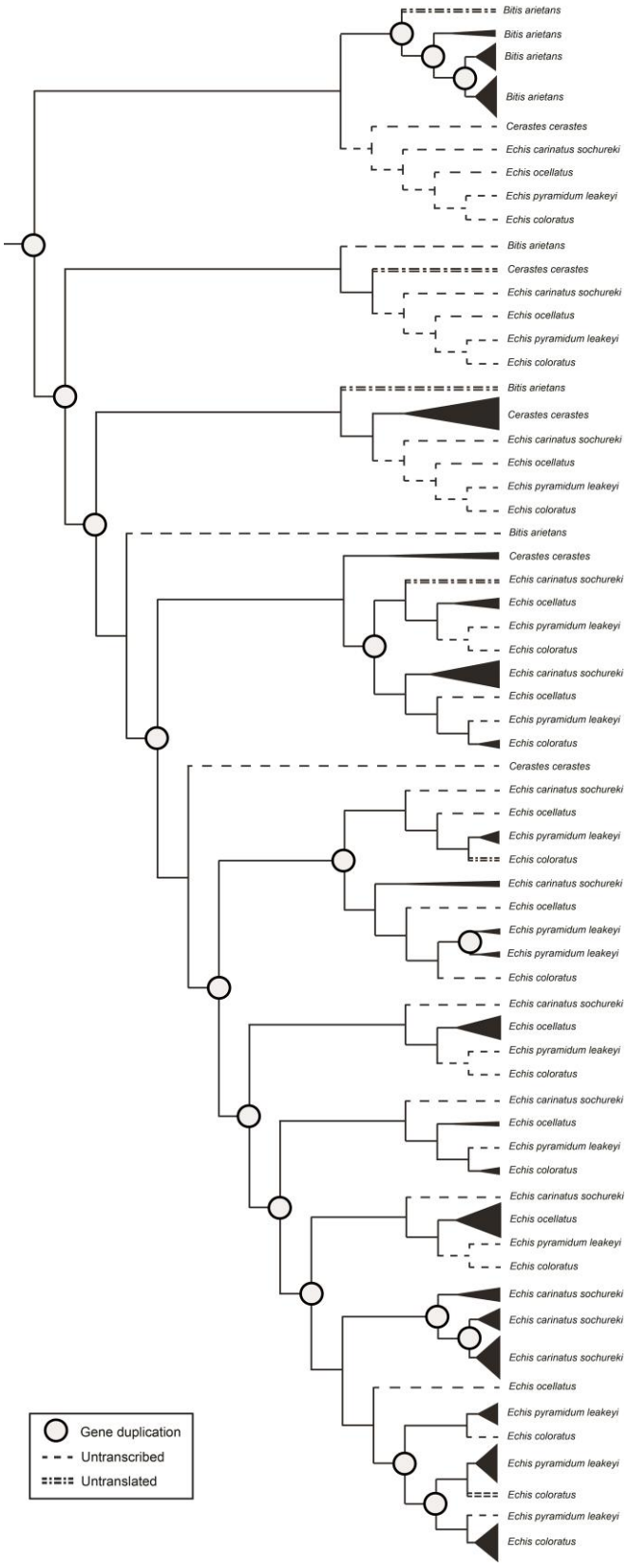

Fig. S9. Gene tree/species tree reconciliation for a major clade of snake venom metalloproteinases displayed in Fig. S8. The width of each branch in the reconciled tree represents the proteomic abundance of each toxin detected in each species. Bordered white circles represent inferred gene duplication events; dotted branches indicated inferred untranscribed loci; double dotted branches indicate genes that were expressed but untranslated.

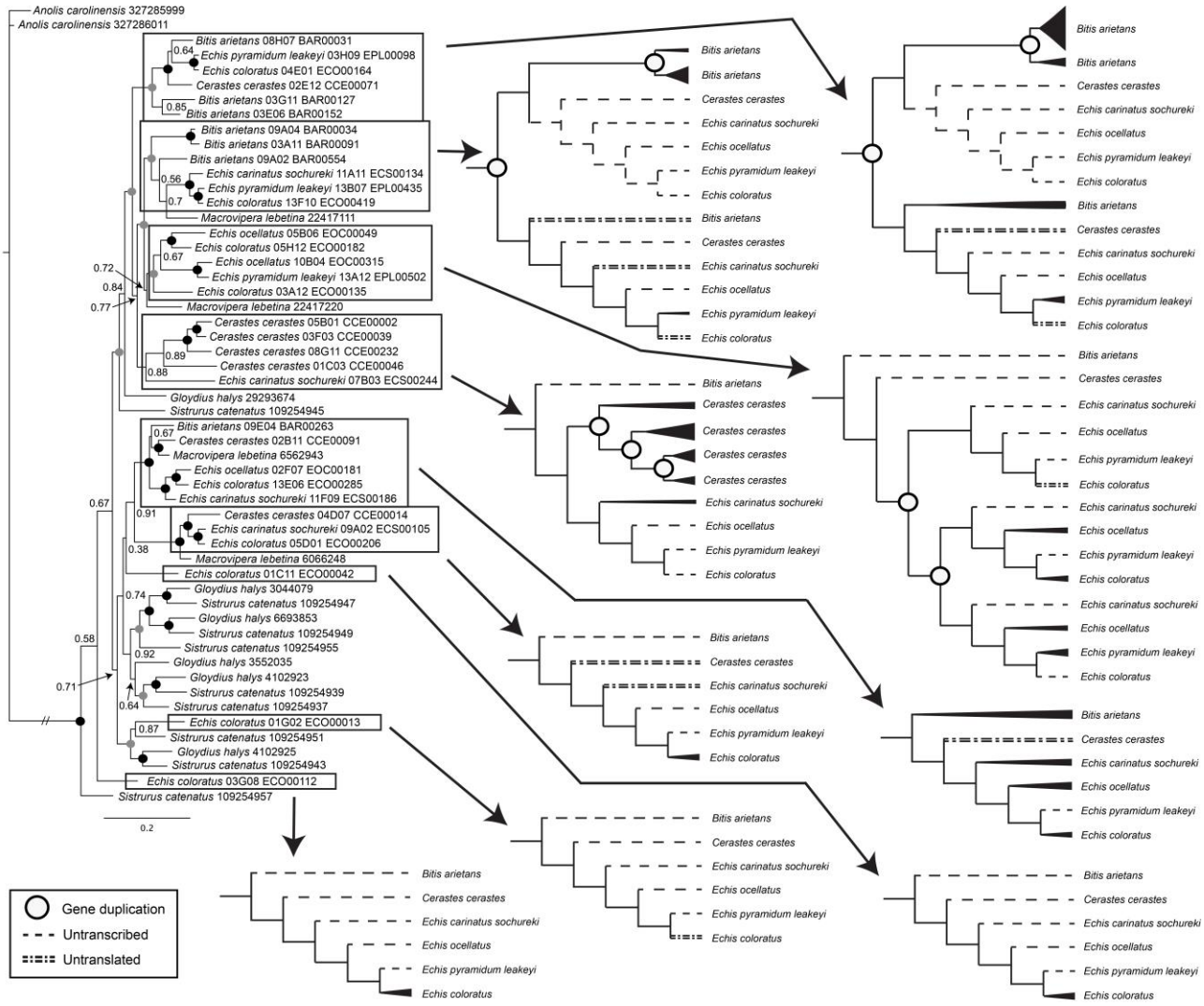

Fig. S10. The evolutionary history of the serine protease (SP) toxin family. The gene tree on the left displays the reconstructed evolutionary history for the gene family. Black dots represent Bayesian posterior probabilities of 1.00; grey dots represent >0.95. Black boxes highlight clades of the gene tree containing *Bitis*, *Cerastes* and/or *Echis* toxins. Arrows next to each clade show the gene tree reconciled to the species phylogeny. The width of each branch in the reconciled tree represents the proteomic abundance of each toxin detected in each species. Bordered white circles represent inferred gene duplication events; dotted branches indicated inferred untranscribed loci; double dotted branches indicate genes that were expressed but untranslated.

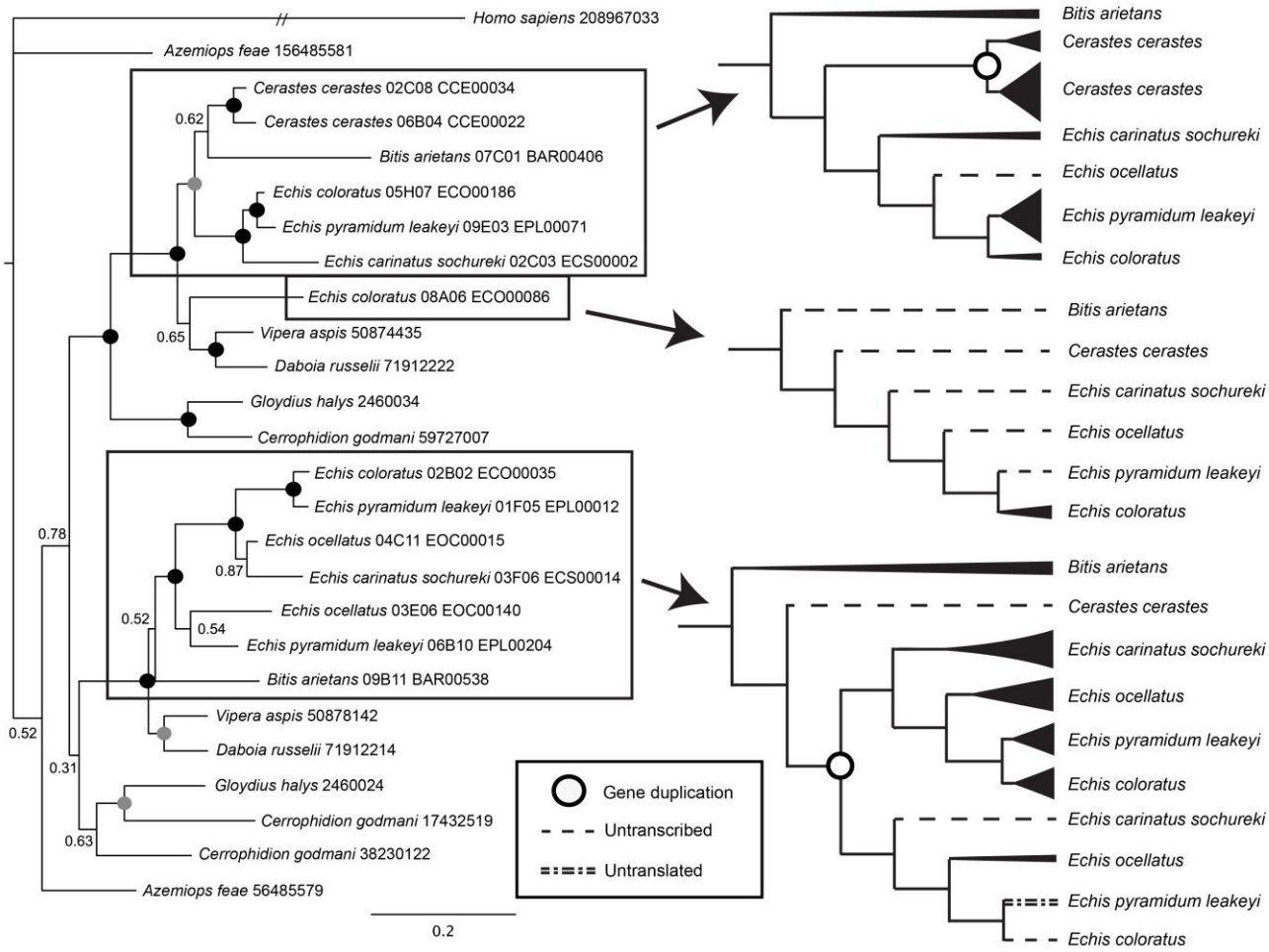

Fig. S11. The evolutionary history of the phopholipase A_2 (PLA_2) toxin family. The gene tree on the left displays the reconstructed evolutionary history for the gene family. Black dots represent Bayesian posterior probabilities of 1.00; grey dots represent >0.95. Black boxes highlight clades of the gene tree containing *Bitis*, *Cerastes* and/or *Echis* toxins. Arrows next to each clade show the gene tree reconciled to the species phylogeny. The width of each branch in the reconciled tree represents the proteomic abundance of each toxin detected in each species. Bordered white circles represent inferred gene duplication events; dotted branches indicated inferred untranscribed loci; double dotted branches indicate genes that were expressed but untranslated.

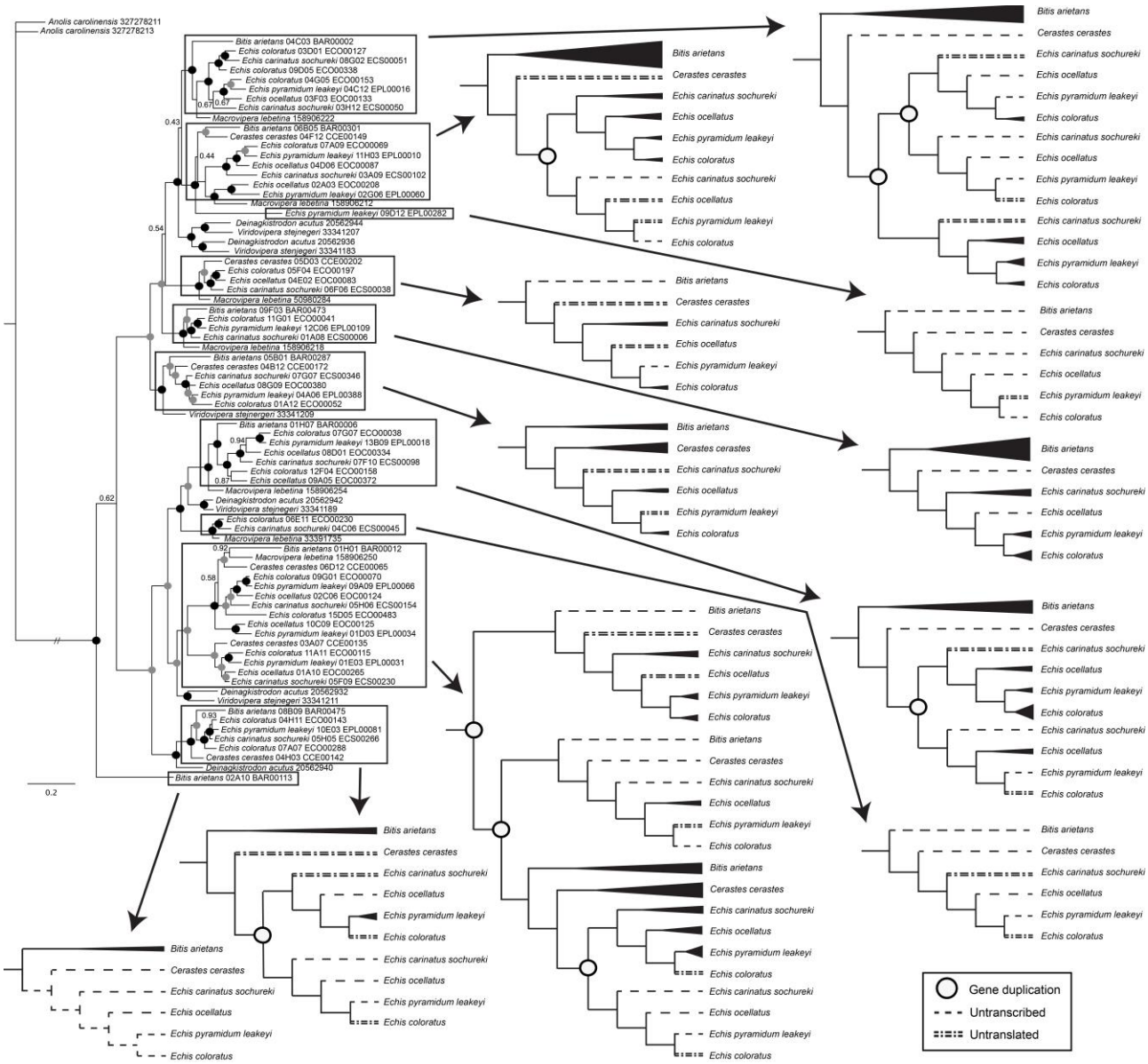

Fig. S12. The evolutionary history of the C-type lectin (CTL) toxin family. The gene tree on the left displays the reconstructed evolutionary history for the gene family. Black dots represent Bayesian posterior probabilities of 1.00; grey dots represent >0.95. Black boxes highlight clades of the gene tree containing *Bitis*, *Cerastes* and/or *Echis* toxins. Arrows next to each clade show the gene tree reconciled to the species phylogeny. The width of each branch in the reconciled tree represents the proteomic abundance of each toxin detected in each species. Bordered white circles represent inferred gene duplication events; dotted branches indicated inferred untranscribed loci; double dotted branches indicate genes that were expressed but untranslated.

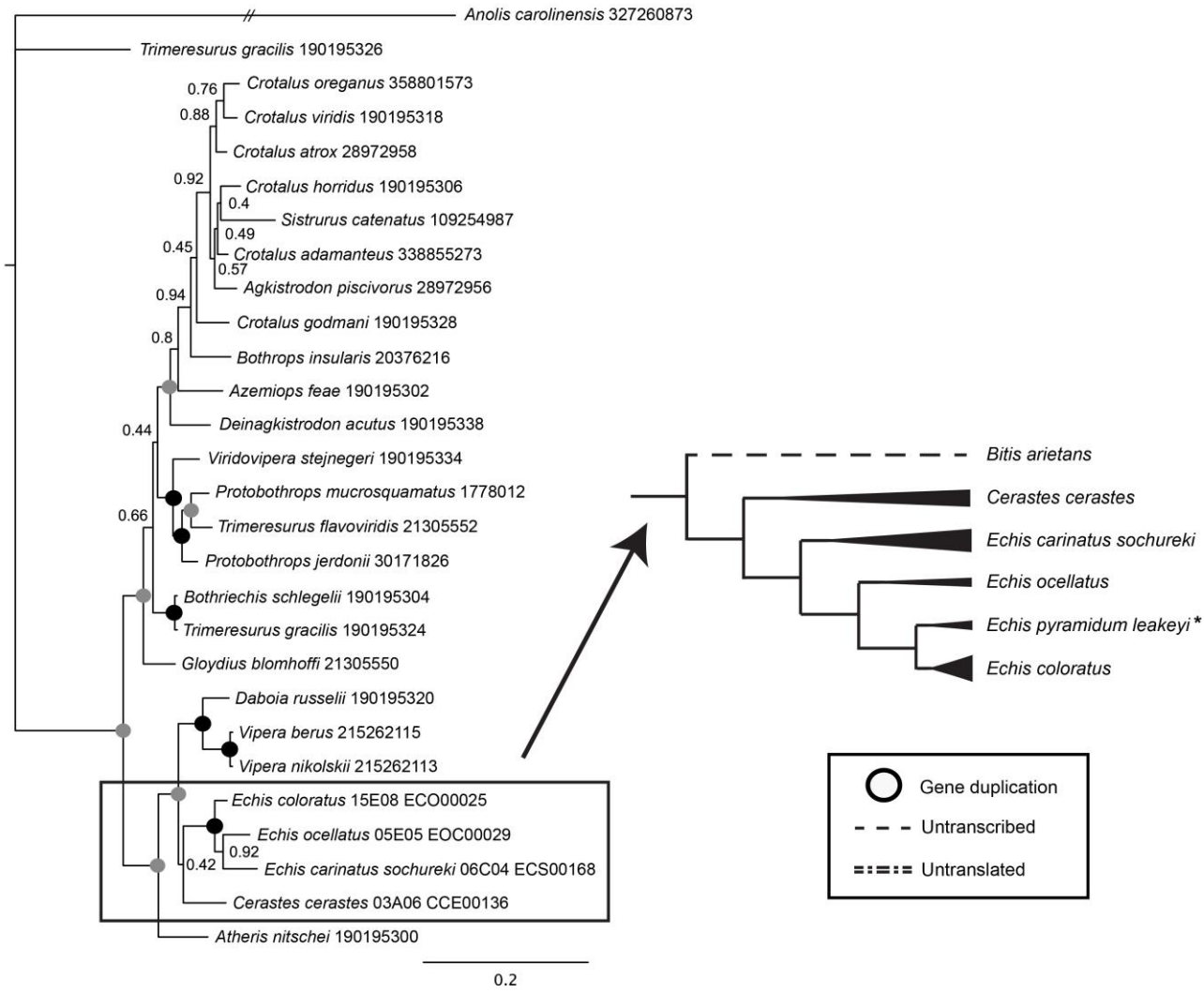

Fig. S13. The evolutionary history of the cysteine rich secretory protein (CRISP) toxin family. The gene tree on the left displays the reconstructed evolutionary history for the gene family. Black dots represent Bayesian posterior probabilities of 1.00; grey dots represent >0.95. Black boxes highlight clades of the gene tree containing *Bitis*, *Cerastes* and/or *Echis* toxins. Arrows next to each clade show the gene tree reconciled to the species phylogeny. The width of each branch in the reconciled tree represents the proteomic abundance of each toxin detected in each species. Bordered white circles represent inferred gene duplication events; dotted branches indicated inferred untranscribed loci; double dotted branches indicate genes that were expressed but untranslated. * Note that this gene was not expressed in the venom gland transcriptome but was detected in the venom proteome. Its placement in the reconciled tree is based on proteomic similarity with the other *Echis* CRISPs and the absence of any paralogous genes recovered.

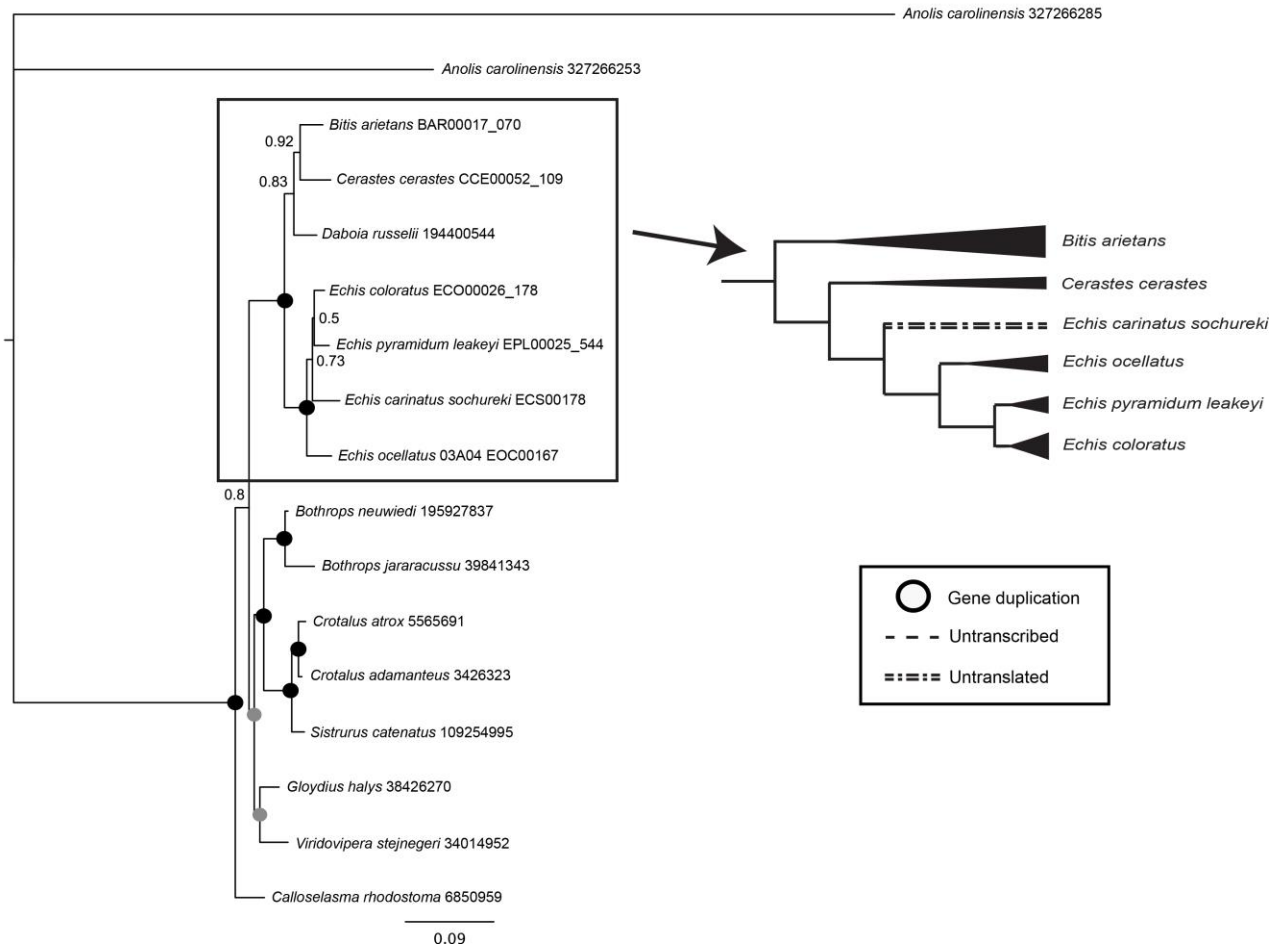

Fig. S14. The evolutionary history of the L-amino acid oxidase (LAAO) toxin family. The gene tree on the left displays the reconstructed evolutionary history for the gene family. Black dots represent Bayesian posterior probabilities of 1.00; grey dots represent >0.95. Black boxes highlight clades of the gene tree containing *Bitis*, *Cerastes* and/or *Echis* toxins. Arrows next to each clade show the gene tree reconciled to the species phylogeny. The width of each branch in the reconciled tree represents the proteomic abundance of each toxin detected in each species. Bordered white circles represent inferred gene duplication events; dotted branches indicated inferred untranscribed loci; double dotted branches indicate genes that were expressed but untranslated.

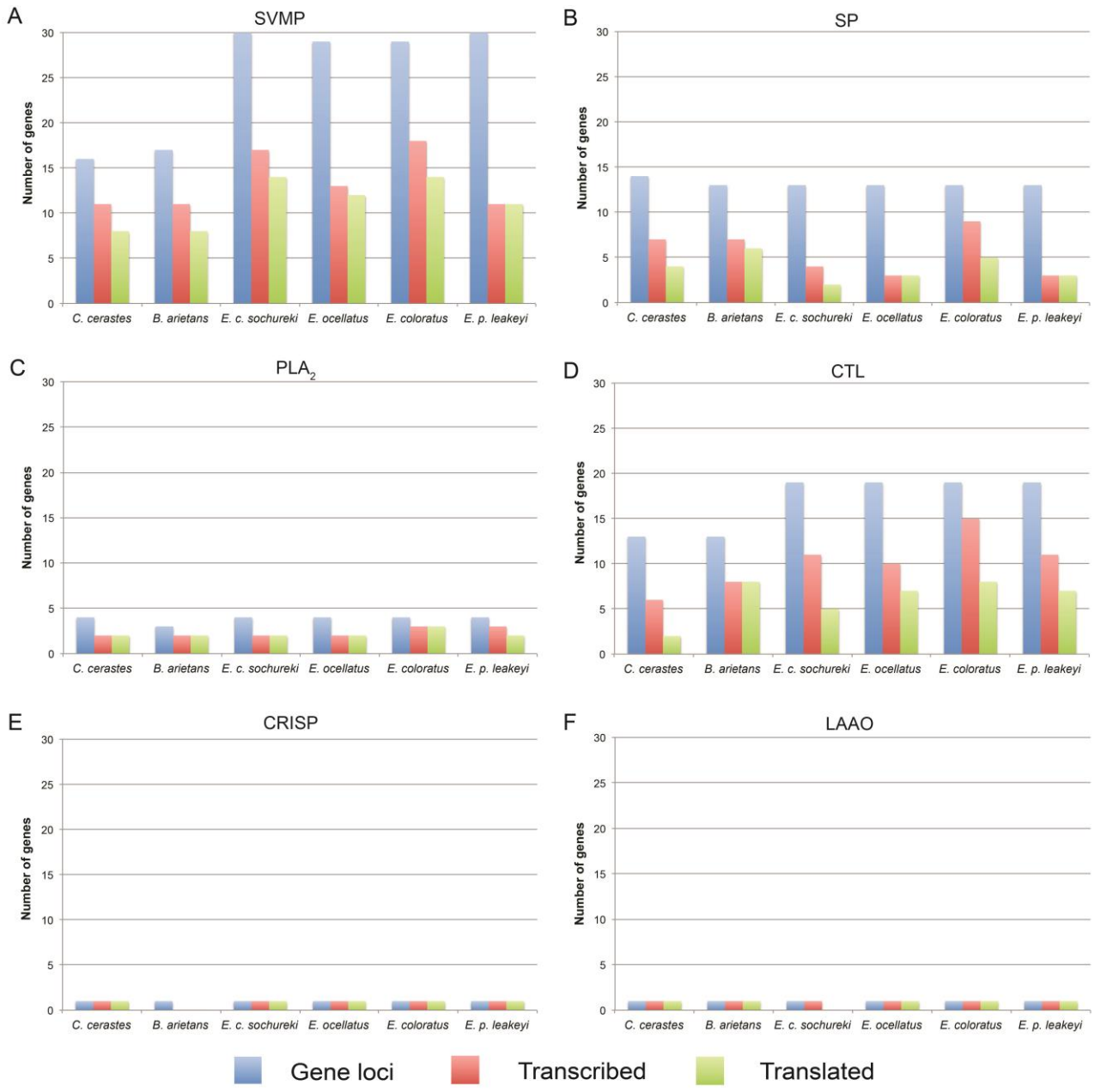

Fig S15. Comparisons of the number of gene loci, genes transcribed and genes translated encoded by the different toxin families in each species. A) snake venom metalloproteinases (SVMP); B) serine proteases (SP); C) phospholipase A2 (PLA2); D) C-type lectins (CTL); E) cysteine-rich secretory proteins (CRISP); F) L-amino acid oxidase (LAAO). Note that the y–axis has been standardised to emphasise the difference between large multi-locus and single-locus gene families. Single loci gene families exhibit a largely 1:1:1 ratio of gene:transcript:protein, whereas major discrepancies are observed in the multi-locus gene families.

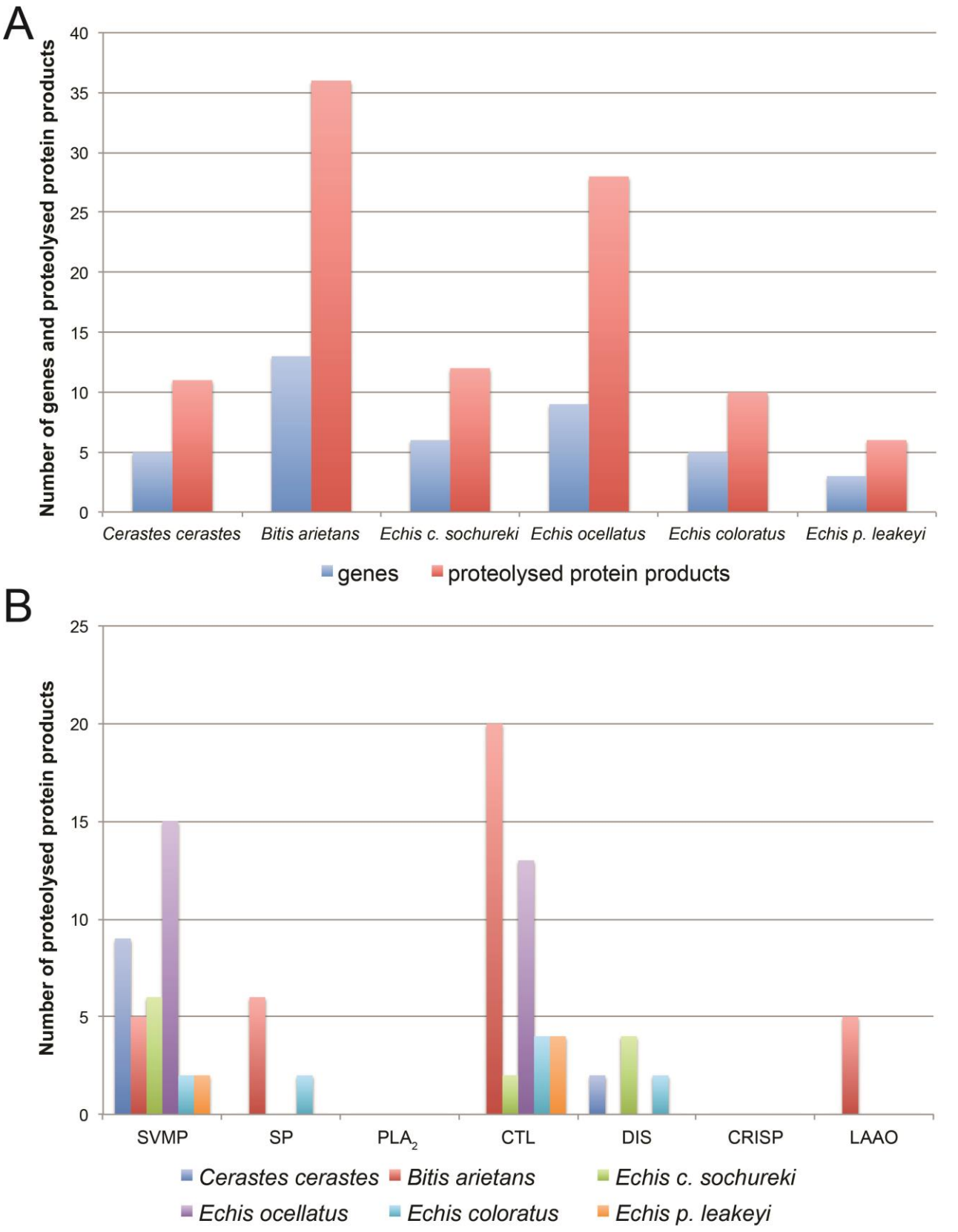

Fig. S16. Quantification of post-translationally proteolysed toxins. A) A comparison of the number of venom toxin genes detected in each species that show evidence of producing post-translationally proteolysed protein products (blue bars) and the number of proteolysed protein products that these genes produced (red bars). B) A comparison of the number of proteolysed proteins produced by each toxin gene family for each species.

3. SI TABLES

Table S1. Assignment of reverse-phase fractions from the venom of *Cerastes cerastes* (Cce) (Egypt) to known protein families.

Spot ID	Spot %	Mass	m/z	z		MS/MS or N-terminal (Nt-)-derived sequences	Protein ID NCBI	Protein ID CCE-	Protein family
CCE-									
2	9.5		444.4	1+		ZKW		00297 00012 00063	SVMPi
4	3.5	8 kDa	654.9	3+		GDDMNDYCTGISSDCPR	P83041	00007	Dimeric disintegrin CC5
					Nt	NSAHPCCDPVTCKPK			
5	2.8	8 kDa	654.9	3+		GDDMNDYCTGISSDCPR	P83041	00007	Dimeric disintegrin
					Nt	NSAHPCCDPVTCKPK			
7	12.9	15 kDa	991.5	2+		FENQDIICGDEDPCNR	ACO92622	00022	D49-PLA_2
			2968.3	1+		SALLSYSAYGCYCGWGGQGKPQDATDR			
			1533.1	1+		CCFVHDCCYGR			
					Nt	NLYQFGKMIKHKTGK			
8	4.7	15.5 kDa			Nt	NLFQFGKMIKHKTGK	ACO92623	00034	D49-PLA_2
			1828.9	1+		VAAICFGENVNTYDKK			
			1533.1	1+		CCFVHDCCYGR			
			2968.3	1+		SAILSYSAYGCYCGWGGQGKPQDATDR			
			1700.8	1+		VAAICFGENVNTYDK			
9	1.1	7 kDa			Nt	NSAHPCCDPVTCKPK	P83043	00007	Dimeric disintegrin CC8

10	3.2	33 kDa			Nt	VIGGAECNINEHRSL	~Q7SYF1	00232	Serine proteinase
			734.8	2+		VIGGAECNINEHR			
			578.3	2+		IYLGLHHFR			
			607.8	2+		VFDYTDWIR		00002	Serine proteinase
			1160.6	1+		IMGWGAITSPK			
11,12	4.9	33 kDa	3687.9	1+		GDSGGPLICNGQIQGIVSWGDEVCGKPNKPGVYTK			
			734.8	2+		VIGGAECNINEHR	~Q7SYF1	00039	Serine proteinase
			578.3	2+		IYLGLHHFR			
13	1.7	24 kDa			Nt	SVDFDSESPRKPEIQ		00136	CRISP
14	6.3	32 kDa			Nt	VIGGAECNINEHPFL	Q75YF1	00232	Serine proteinase
			1155.3	1+		IYLGLHHFR			
			1184.3	1+		SLVFLYNSSR			
			1468.2	1+		VIGGAECNINEHR			
			3687.7	1+		VLSAAHCDGENMKIYLGLHHFRLPNKDRQIR			
15	2.6	31.5 kDa			Nt	VVGGDECNINEHQSL		00046	Serine proteinase
16	1.8	31 kDa			Nt	VVGGDECNINEHRSL		00051	Serine proteinase
17	9.1	15 kDa			Nt	DQDCLPGWSYEEK		00065	C-type lectin-like
			502.7	2+		TADNQWXR			
			575.3	2+		YELVWIGLR			
		17 kDa	749.9	3+		DFDCPSDWSAYDQYCYR		00172	C-type lectin-like
19	2.6	55 kDa	727.3	3+		LYEIVNTVNEMYLPLNVR		00017	PIII-SVMP
20	0.9	56 kDa	626.8	2+		SAGQLYQESLR		00052	LAO
	2.9	52 kDa	913.8	3+		DECDXPEHCTGQSAECPSDXTR		00017	PIII-SVMP
			505.8	2+		GSYFGYCR			PIII-SVMP

21	1.0	66 kDa	691.8	2+	VPLVGLEIWSNR	00021	PIII-SVMP
	1.3	52 kDa	691.6	3+	NLLVAITMAHELGHNLGLR	00218	PIII-SVMP
22	3.0	58 kDa	702.8	2+	VALIGLEFWSNR	00041	PIII-SVMP
			514.9	2+	IPACEPDVK		
23	3.5	59 kDa	557.8	2+	ZFILTPEQR	00111	PIII-SVMP
			652.8	2+	LVIVVDNVMFR		
	0.8	56 kDa	667.1	2+	YSTGIVQDHSTR	~00017	PIII-SVMP
			678.7	2+	VALVGLEIWSNR	00027	PIII-SVMP
24	14.1	58 kDa	702.8	2+	VALIGLEFWSNR	00041	PIII-SVMP
			651.1	2+	NEYQTYLTNR	00008/61	PII-SVMP
			1663.8	1+	FLTAGTVCLPARGDW		
			3868.7	1+	VAATMAHEIGHNLGMNHDGNQCNCGANGCVMAGMLR		
25	2.1	56 kDa	764.3	2+	ATVAQDACFQFNR	00212	PIII-SVMP
			580.1	2+	CECCDQCK		
			521.8	2+	IPCAPEDIK		
	0.2	23 kDa	651.1	2+	NEYQTYLTNR	00008/61	PI(PII)-SVMP
			1707.6	1+	GEHCISGPCCENCK		
			1352.7	1+	NPQCILNQPLR		
			1701.9	1+	YIELVIVADHAMVTK		
			1417.6	1+	AQPSYQFSDCSK		
			1301.6	1+	NEYQTYLTNR		
26	3.5	48 kDa	764.3	2+	ATVAQDACFQFNR	00212	PIII-SVMP
			566.3	2+	LGNQYGYCR	00212	PIII-SVMP
			651.3	2+	NEYQTYLTNR	00008/61	PII-SVMP

Protein assignment undertaken by MALDI-TOF-TOF- and nESI- CID-MS/MS of peptide ions obtained from in-gel digested protein bands separated by SDS-PAGE. X = Ile or Leu; Z, pyroglutamic acid; M_{ox}, methionine sulphoxide. Cysteine residues are carbamidomethylated. Nt, N-terminal sequence determined by automated Edman degradation. Apparent molecular masses (in kDa) were estimated from SDS-PAGE analysis of reduced samples.

Table S2. Assignment of reverse-phase fractions from the venom of *Bitis arietans* (Bar) (Nigeria) to known protein families.

Spot ID	Spot %	Mass	m/z	z	MS/MS or N-terminal (Nt-)-derived sequences	Mascot score	Protein ID NCBI	Protein ID BAR-	Protein family
BAR-									
1, 2	0,2		568.4	2+	ZRPPGHHIP	de novo		BA00003	BPP/pGpH
3,4	0,7		649.0	2+	(223.1)RPPRPQIPP	de novo		BA00003	BPP/pGpH
5	1,3		585.0	2+	(268.3)PPMPMG(291.1)	de novo			unknown
6	1.2		431.1	1+	ZDW	de novo		BA00003	SVMP inhibitor
7	2,3	8891			Nt: SPPVCGNKILEQGED		P17497	BA00042	SVMP disintegrin
8,9,10	1,1	8-9 kDa			Nt: SPPVCGNKILEQGED		P17498	BA00042	SVMP disintegrin
12	2.2	14▪ kDa	522,3 494,2	2+ 2+	EMIDHVSGR GKPQDATDR		AAR06850	BA00406	D49-PLA2
13	3,1	14▪ kDa	404,7 412,7 486,8	2+ 2+ 2+	ICECDK AAAICFR GKPIDATDR		AAX86635	BA00538	D49-PLA2
14,15,16	0,9	60▪/150▪	494,7 466,7	2+ 2+	ATYWYER YFCLSSR		CBM40647	BA00034	Serine protease
14	1,1	60▪/150▪	607,8	2+	VFDYTDWIR				

15	0,4	12■ kDa	532,8	2+	SKNDYYFK	234	P08935	BA00172	Cystatin
			567,8	2+	NDYYFKER				
			425,2	2+	NDYYFK				
			743,4	2+	RVVEAQSQVVSGVK				
			665,4	2+	VVEAQSQVVSGVK				
			902,9	2+	DVTDPDVQEAAAFAVEK				
16	3,5	60■ kDa	607,8	2+	VFDYTDWIR		CBM40647	BA00091	Serine protease
			807,4	2+	STYWYELLPAQSR				
	3,5	28■ kDa	987,5	2+	DGHLISIDSQEEADFVAK		Q6T7B6	BA00002	C-type lectin
			459,3	2+	LVSENVEK				
			576,3	2+	GVDVFWIGMK				
			582,3	2+	SEILWMGLSK		Q6X5T5	BA00006	C-type lectin
17	2	60■ kDa	459,3	2+	LVSENVEK		Q6T7B6	BA00002	C-type lectin
			576,3	2+	GVDVFWIGMK				
			454,2	2+	KEADFVAK		Q6X5T5	BA00006	C-type lectin
			607,8	2+	VFDYTDWIR		CBM40647	BA00091	Serine protease
			807,4	2+	STYWYELLPAQSR				
			494,7	2+	ATYWYER		CBM40647	BA00034	Serine protease
			607,8	2+	VFDYTDWIR				
	1,1	31■ kDa	454,2	2+	KEADFVAK		Q6X5T5	BA00006	C-type lectin
			511,3	2+	AAHQKLPEK	92	Q6T6S7	BA00263	Serine Protease
			797,4	2+	TLCAGILEGGIDSCK				
	3,2	16■/31■	987,5	2+	DGHLISIDSQEEADFVAK		Q6T7B6	BA00002	C-type lectin
			459,3	2+	LVSENVEK				
18	2,6	60▼/40▼	571,8	2+	VFDYGDWIK		CBM40646	BA00031	Serine protease
			552,8	2+	TLCAGVLEGGK	de novo		BA00127	Serine protease

			757,3	2+	VIGGDECDINEHR	110	Q9PRY9		Serine protease
			505,3	3+	VIGGDECDINEHR				
	0,5	33▼ kDa	785,4	2+	NVQNEDEEIRVPK	211	Q6T6S7	BA00263	Serine protease
			530,3	2+	LFDYSVCR	de novo			
			1044,	2+	FHCAGTLLNKEWVLTAAR				
			530,3	2+	LFDYSVCR				
			797,4	2+	TLCAGILEGGIDSCK				
	0,5	25▼ kDa	511,3	2+	AAHQKLPEK	108	Q6T6S7	BA00263	Serine protease
			595,3	2+	WDKDIMLIR				
			597,8	2+	IMGWGSITTTK				
			530,3	2+	LFDYSVCR				
19	0,2	100■ kDa	571,8	2+	VFDYGDWIK		CBM40646	BA00031	Serine protease
			534,3	2+	IKPNPDQQK				
	0,4	100■/50■	490,3	3+	VIGGAECNINEHR		CBM40647	BA00152	Serine Protease
			734,9	2+	VIGGAECNINEHR				
	1,5	100■/50■	607,8	2+	VFDYTDWIR		CBM40647	BA00091	Serine protease
			807,4	2+	STYWYELLPAQSR				
			552,8	2+	TLCAGVLEGGK	60	P86497	BA00152	Serine Protease
	0,1	28■ kDa	807,4	2+	STYWYELLPAQSR		CBM40647	BA00091	Serine protease
			466,7	2+	YFCLSSR				
	0,8	12■ kDa	602,3	3+	DVTDPDVQEAAAFAVEK	201	P08935	BA00172	Cystatin
			902,9	2+	DVTDPDVQEAAAFAVEK				
			567,8	2+	NDYYFKER				
			665,4	2+	VVEAQSQVVSGVK				
			532,8	2+	SKNDYYFK				

			775,4	2+	YNAHSKNDYYFK				
20	1,6	100▪ kDa	571,8	2+	VFDYGDWIK		CBM40646	BA00031	Serine protease
			534,3	2+	IKPNPDQQK				
			734,9	2+	VIGGAECNINEHR		CBM40647	BA00152	Serine Protease
			748,4	2+	DIMoxLIK				
	6,2	100▪/45▪	552,8	2+	TLCAGVLEGGK	69	P86497	BA00031/152	Serine Protease
	3	45▪ kDa	607,8	2+	VFDYTDWIR			BA00034/91	
			431,8	2+	AFDEPKR		Q6T7B6	BA00473	C-type lectin
			630,8	2+	CGDDYPFVCK				
	1,6	36▪ kDa	534,3	2+	IKPNPDQQK		CBM40646	BA00031	Serine protease
			612,3	2+	IKPNPDQQKR				
			504,3	2+	DAYGDLPEK				
			625,8	2+	DAYGDLPEKSR				
			571,8	2+	VFDYGDWIK				
21	1,2	40▪ kDa	807,4	2+	STYWYELLPAQSR		CBM40647	BA00091	Serine protease
			578,8	2+	VFDYADWIK		CBM40648	BA00152	Serine protease
			734,9	2+	VIGGAECNINEHR				
			552,8	2+	TLCAGVLEGGK	76	P86497		Serine Protease
	0,3	28▪ kDa	454,3	2+	KEADFVAK		Q6X5T5	BA00006	C-type lectin
			742,8	2+	EADYEEFLEIAR		G8XQX1	BA00017	LAO
			502,3	2+	VTVLEASER	57	P0DI84	BA00017	LAO
			987,4	2+	DGHLISIDSQEEADFVAK		Q6T7B6	BA00002	C-type lectin
			459,3	2+	LVSENVEK				
			728,3	2+	DEGCLPDWSSYK	47	Q7LZK8	BA00012	C-type lectin
	0,7	12▪ kDa	454,7	2+	EESAFVAR		I7ICN3	BA00113	C-type lectin
			525,8	2+	VAPDTCFLK		AAX86634	BA00027	PIII-SVMP

22	3,5	31▼/14▼	431,7	2+	AFDEPKR	de novo	Q6T7B6	BA00473	C-type lectin
	1,1	18▼ kDa	431,7	2+	AFDEPKR		Q6T7B6	BA00473	C-type lectin
			839,4	2+	EEADFVAQLVSENVK				
			860,9	2+	SSPDYVWIGLWNQR				
			401,7	2+	TWLNLR				
	1,5	14▼ kDa	742,8	2+	EADYEEFLEIAR		G8XQX1	BA00017	LAO
			502,3	2+	VTVLEASER				
			475,7	2+	EEGEFIVK		Q6X5T4	BA00465	C-type lectin
			502,3	2+	VTVLEASER				
			431,7	2+	AFDEPKR				
			434,2	2+	NCFGLEK				
23	0,7	33■ kDa	742,8	2+	EADYEEFLEIAR		G8XQX1	BA00017	LAO
			502,3	2+	VTVLEASER				
			1008,	2+	LNEFVQETENGWYFIK	221	Q4F867	BA00017	LAO
			849,9	2+	DPGLLKYPVKPSEAGK				
			538,3	2+	YPVKPSEAGK				
			422,8	2+	AVEELKR				
			747,3	2+	ADDKNPLEECFR				
			498,3	3+	ADDKNPLEECFR				
			1073,	2+	RFDEIVGGMoxDQLPTSMoxYR				
24	0,7	150■/31■	742,8	2+	EADYEEFLEIAR		G8XQX1	BA00017	LAO
	0,2	150■ kDa	502,3	2+	VTVLEASER		G8XQX1	BA00017	LAO
	0,2	55■ kDa	463,8	2+	FVFDLASK		Q6X5T4	BA00475	C-type lectin
			517,3	2+	IYIWIGLR		AAR06853		C-type lectin 3
			552,8	2+	TLCAGVLEGGK	52	P86497	BA00031/152	Serine Protease

	0,5	31▪ kDa	576,3	2+	VTVTYNTPEK		G8XQX1	BA00070	LAO
			425,2	2+	FDYNAYTR		Q6X5T4	BA00475	C-type lectin
			1046,	2+	FCMEQANDGHLVSIQSIK	49	Q6X5T3	BA00287	C-type lectin 5
			469,3	2+	VTYVNWR				
			475,3	2+	VIYVNWR	116	AAR06853		C-type lectin 3
			691,3	2+	EGESQMCQALTK				
			699,3	2+	EGESQMoxCQALTK				
			517,3	2+	IYIWIGLR				
			652,9	2+	IYIWIGLRDR				
24,25	0,3	28▪ kDa	438,7	2+	STTDLPSR		G8XQX1	BA00070	LAO
			459,3	2+	LVSENVEK		Q6T7B6	BA00002	C-type lectin
			764,9	2+	ELVNGGHLMoxSVNSR	52	Q7LZK8	BA00012	C-type lectin
25	1	28▪ kDa	582,3	2+	SEILWMGLSK		Q6X5T5	BA00006	C-type lectin
			469,7	2+	EEADFVTK		Q6X5T3	BA00301	C-type lectin
			620,3	2+	FVYDAWIGLR				
			712,8	2+	DPGCLPDWSSYK	101	Q7LZK5	BA00301	C-type lectin (Bitiscetin α)
			469,7	2+	EEADFVTK				
			620,3	2+	FVYDAWIGLR				
			849,9	2+	FVYDAWIGLRDESK				
			603,8	2+	TWTDLPCGEK				
			983,5	2+	TWTDLPCGEKNPFICK				
26	0,9	50▪ kDa	713,3	2+	EISEYGMVDPGTK		AAX86634	BA00027	PIII-SVMP
	0,5	28▪ kDa	517,7	2+	VGTWEDAEK		Q6X5T3	BA00301	C-type lectin
			469,7	2+	EEADFVTK				
			431,3	2+	LASQTLTK				
			620,3	2+	FVYDAWIGLR				
			813,9	2+	IVLVWIGLSHFWR	103	Q7LZK8	BA00012	C-type lectin (Bitiscetin□□)

			985,5	2+	ALSDEPICFVAESFHNK				
27	1,4	50▪ kDa	765,4	2+	IIALGHSGFSEDQR		F8S0Z7	BA00083	5'-Nucleotidase
			979,1	2+	YLGYLNVVFDDKGNVIK				
			513,8	2+	ASGNPILLNK				
			529,8	2+	DIPEDQDVK				
			555,8	2+	QAFEHSVHR				
			628,9	2+	GRGELLQVSGIK				
			522,3	2+	GELLQVSGIK				
			779,1	3+	GDSSNHSSGDLDISIVGDYIKR				
			437,7	2+	VFPAVEGR				
			982,5	2+	ALSDEPICFVAESFHNK	56	Q7LZK8	BA00012	C-type lectin
	0,4	33▪ kDa	742,9	2+	EADYEEFLEIAR		G8XQX1	BA00017	LAO
			931,5	3+	TSNPQHVVIVGAGMSGLSAAYVLAGAGH				
			502,3	2+	VTVLEASER				
			538,3	2+	YPVKPSEAGK				
			611,8	2+	SAGQLYQASLGK				
			422,8	2+	AVEELKR				
			469,7	2+	EEADFVTK	30	Q6X5T3	BA00301	C-type lectin
			431,3	2+	LASQTLTK				
			983,5	2+	TWTDLPCGEKNPFICK				
	0,4	30▪ kDa	502,3	2+	VTVLEASER		G8XQX1	BA00017	LAO
			576,3	2+	VTVTYNTPEK		G8XQX1	BA00070	LAO
			431,3	2+	LASQTLTK			BA00298	
			502,3	2+	VTVLEASER	49	P0DI84	BA00070	LAO
			438,7	2+	STTDLPSR				
			575,4	2+	INFKPPLPPK	48	B5U6Y8		LAO
			422,8	2+	AVEELKR	48	B5U6Y8	BA00017/70	LAO
27,28,29	3,5	25▪ kDa	431,3	2+	LASQTLTK		Q6X5T3	BA00301	C-type lectin

			849,9	2+	FVYDAWIGLRDESK				
			832,4	2+	FCVENSGHLASIDSK	236	Q7LZK5	BA00301	C-type lectin
			469,7	2+	EEADFVTK				
			861,7	3+	FCVENSGHLASIDSKEEADFVTK				
			431,3	2+	LASQTLTK				
			849,9	2+	FVYDAWIGLRDESK				
			698,8	2+	CFGLDVHTEYR				
			603,8	2+	TWTDLPCGEK				
28	0,5	150▪ kDa	ND						
	0,6	55▪ kDa	513,8	2+	ASGNPILLNK		F8S0Z7	BA00083	5'-Nucleotidase
			792,4	2+	IQLQNYYSQEIGK				
			628,9	2+	GRGELLQVSGIK				
			522,3	2+	GELLQVSGIK				
			765,4	2+	IIALGHSGFSEDQR	143	F8S0Z7	BA00083	5'-Nucleotidase
			653,9	2+	QVPVVQAYAFGK				
			555,8	2+	QAFEHSVHR				
			437,7	2+	VFPAVEGR				
			764,9	2+	ELVNGGHLMoxSVNSR	91	Q7LZK8	BA00012	C-type lectin
			728,8	2+	DEGCLPDWSSYK				
28,29	6,8	25▪ kDa	517,7	2+	VGTWEDAEK		Q6X5T3	BA00301	C-type lectin
			469,7	2+	EEADFVTK				
			620,3	2+	FVYDAWIGLR				
			712,8	2+	DPGCLPDWSSYK	214	Q7LZK5	BA00301	C-type lectin
			517,7	2+	VGTWEDAEK				
			620,3	2+	FVYDAWIGLR				
			919,1	3+	TQQCSPQWTDGSSVVYENVDEPTK				
			983,5	2+	TWTDLPCGEKNPFICK				
			764,9	2+	ELVNGGHLMoxSVNSR	195	Q7LZK8	BA00012	C-type lectin
			410,7	2+	TWADAEK				

			628,3	2+	TWADAEKFCK				
			728,8	2+	DEGCLPDWSSYK				
			701,3	2+	DEGCLPDWSSYKGHCYK				
29	1,8	66▪ kDa	595,3	2+	VVNNVNVIYR		ADI47619	BA00462	PIII-SVMP
			492,7	2+	CCNAATCK		ADI47619	BA00042	PIII-SVMP
			1052,	2+	LTPGSQCNYGECCDQCR				
	1,3	25▪ kDa	581,8	2+	KVGTWEDAEK	250	Q7LZK5	BA00301	C-type lectin
			582,3	2+	WIQWTCNR	195	Q7LZK8	BA00012	C-type lectin
30	0,1	100▪ kDa	742,8	2+	EADYEEFLEIAR		G8XQX1	BA00017	LAO
			502,3	2+	VTVLEASER				
			611,8	2+	SAGQLYQASLGK				
			645,4	2+	YINVIVVADQR		ADI47619	BA00042	PIII-SVMP
			611,8	2+	SAGQLYQASLGK		G8XQX1	BA00070	LAO
			576,3	2+	VTVTYNTPEK				
			595,3	2+	VVNNVNVIYR		ADI47619	BA00462	PIII-SVMP
			431,7	2+	AFDEPKR		Q6T7B6	BA00473	C-type lectin
			860,9	2+	SSPDYVWIGLWNQR				
			742,9	2+	EADYEEFLEIAR		G8XQX1	BA00017	LAO
	1,1	66▪ kDa	1052,	2+	LTPGSQCNYGECCDQCR		ADI47619	BA00042	PIII-SVMP
30,31	0,9	50▪ kDa	538,3	2+	YPVKPSEAGK		G8XQX1	BA00017/70	LAO
			611,8	2+	SAGQLYQASLGK				
			422,8	2+	AVEELKR				
			502,3	2+	VTVLEASER		G8XQX1	BA00017	LAO
			438,7	2+	STTDLPSR		G8XQX1	BA00070	LAO
30	0,7	50▪ kDa	422,8	2+	AVEELKR				
			438,7	2+	STTDLPSR				

			535,3	2+	YPVKPSEAGK				
			575,4	2+	IIFKPPLPPK				
			728,3	2+	DEGCLPDWSSYK	60	Q7LZK8	BA00062	C-type lectin
	0,4	28▪ kDa	459,3	2+	LVSENVEK		Q6T7B6	BA00002	C-type lectin
			683,9	2+	SKGVDVFWIGMK				
			576,3	2+	GVDVFWIGMK				
			517,7	2+	VGTWEDAEK		Q6X5T3	BA00301	C-type lectin
			469,7	2+	EEADFVTK				
			431,3	2+	LASQTLTK				
			620,3	2+	FVYDAWIGLR				
			849,9	2+	FVYDAWIGLRDESK				
			756,9	2+	ELVNGGHLMSVNSR	132	Q7LZK8	BA00012	C-type lectin
			764,9	2+	ELVNGGHLMoxSVNSR				
			701,3	3+	DEGCLPDWSSYKGHCYK				
			728,8	2+	DEGCLPDWSSYK				
			431,6	2+	LASQTLTK	92	Q7LZK5	BA00301	C-type lectin
			469,7	2+	EEADFVTK				
			517,7	2+	VGTWEDAEK				
			620,3	2+	FVYDAWIGLR				
			849,9	2+	FVYDAWIGLRDESK				
			983,5	2+	TWTDLPCGEKNPFICK				
31	0,2	100▪ kDa	645,4	2+	YINVIVVADQR		ADI47619	BA00042	PIII-SVMP
			576,3	2+	VTVTYNTPEK		AAZ08620	BA00070	LAO
			438,7	2+	STTDLPSR				
			860,9	2+	SSPDYVWIGLWNQR		Q6T7B6	BA00473	C-type lectin
			422,8	2+	AVEELKR	47	G8XQX1	BA00017/70	LAO
	0,2	50▪ kDa	742,9	2+	EADYEEFLEIAR		G8XQX1	BA00017	LAO
			576,3	2+	VTVTYNTPEK		AAZ08620	BA00070	LAO
			575,3	2+	IIFKPPLPPK				

			1008,	2+	LNEFVQETENGWYFIK	97	Q4F867	BA00017	LAO
			538,3	2+	YPVKPSEAGK				
			422,7	2+	AVEELKR				
			441,7	2+	IFLTCTK		AAZ08620	BA00070	LAO
			438,7	2+	STTDLPSR				
			747,3	2+	ADDKNPLEECFR		G8XQX1	BA00017/70	LAO
	0,2	28▪ kDa	576,3	2+	GVDVFWIGMK		Q6T7B6	BA00002	C-type lectin
			469,3	2+	VTYVNWR		Q6X5T3	BA00287	C-type lectin
			463,8	2+	FVFDLASK		Q6X5T4	BA00475	C-type lectin
			525,2	2+	FDYNAYTR				
			728,8	2+	DEGCLPDWSSYK	44	Q7LZK8	BA00012	C-type lectin
			620,3	2+	FVYDAWIGLR	40	Q7LZK5	BA00301	C-type lectin
			849,9	2+	FVYDAWIGLRDESK				
32	0,1	55▼ kDa	611,8	2+	SAGQLYQASLGK		G8XQX1	BA00017	LAO
	0,1	45▼ kDa	514,3	2+	IPCAPQDVK	39	Q4VM08	BA00101	PIII-SVMP
			718,3	2+	LYCFDNLPEHK				
	0,3	24▼ kDa	639,7	3+	SSDPIKYINVIVVADQR		ADI47619	BA00042	PIII-SVMP
			645,4	2+	YINVIVVADQR				
			664,4	2+	LVTYYKGELNK				
			753,8	2+	SASDTLHSFAEWR				
			444,7	2+	ERDLLSR				
	0,8	18▼ kDa	463,8	2+	FVFDLASK		Q6X5T4	BA00475	C-type lectin
			525,2	2+	FDYNAYTR				
	0,8	10▼ kDa	469,3	2+	VTYVNWR	28	Q6X5T3	BA00287	C-type lectin
33	0,9	95▪ kDa	1177,	1+	VNADHVGFYR		CBJ34330	BA00089	Aminopeptidase A
			1266,	1+	AGFIDDAFALAR				

			501,3	1+	AGLLK			
			1709,	1+	YLQNEAEYIPWQR			
			1005,	1+	AIVAISYIR			
			1067,	1+	QCISLFGSR	AGL45259	BA00101	PIII-SVMP
			1527,	1+	ATVAQDACFQFNR			
			1131,	1+	LANDYGYCR			
			1259,	1+	LANDYGYCRK			
			1435,	1+	LYCFDNLPEHK			
			651,4	1+	YLIDK	Q4VM07	BA00377	PIII-SVMP
			2893,	1+	TDIVSPPVCGNFLVELGEDCDCGSPR			
			1629,	1+	DCQNQCCNAATCK			
			764,6	1+	TAGTVCR			
			1694,	1+	LSCEASYLFSDCSR			
			1326,	1+	MPQCLLNIPLK			
	5,3	31▪/20▪	959,1	2+	SSDPIKYINVIVVADQR	ADI47619	BA00042	PIII-SVMP
			753,9	2+	SASDTLHSFAEWR			
	3,8	20▪ kDa	639,7	3+	SSDPIKYINVIVVADQR	ADI47619	BA00042	PIII-SVMP
			645,4	2+	YINVIVVADQR			
			664,4	2+	LVTYYKGELNK			
			502,3	2+	GELNKITTK			
			857,9	2+	VHQYFNTLNEMYR			
			414,2	2+	YFYIRPILR			
			796,9	2+	SAAVVMNYQPEIDR			
			1122,	3+	SAAVVMNYQPEIDRAVAAIMAHEMGHNLG			
			597,6	3+	AVAAIMAHEMGHNLGIR			
			895,9	2+	AVAAIMAHEMGHNLGIR			
34	0,5	95▼ kDa	764,3	2+	ATVAQDACFQFNR	de novo	AGL45259	BA00101	PIII-SVMP
	0,3	45▼ kDa	718,3	2+	LYCFDNLPEHK			

			534,3	2+	QCISLFGSR			
	0,3	20▼ kDa	ND					
35,36	9,6	50■/28■	680,4	2+	KHDNAQLFTGTK	ADW54354	BA00015	PIII-SVMP
			616,3	2+	HDNAQLFTGTK			
			403,2	2+	FNGGIIGK			
			742,9	2+	EADYEEFLEIAR	G8XQX1	BA00017	LAO

Protein assignment undertaken by MALDI-TOF-TOF- and nESI- CID-MS/MS of peptide ions obtained from in-gel digested protein bands separated by SDS-PAGE. X = Ile or Leu; Z, pyroglutamic acid; M_{ox}, methionine sulphoxide. Cysteine residues are carbamidomethylated. Nt, N-terminal sequence determined by automated Edman degradation. Apparent molecular masses (in kDa) were estimated from SDS-PAGE analysis of reduced samples.

Table S3. Assignment of reverse-phase fractions from the venom of *Echis carinatus sochureki* (Ecs) (UAE) to known protein families.

Spot ID	%	Mass	m/z	z		MS/MS or N-terminal (Nt-)-derived sequences	Mascot score	Protein ID NCBI	ECS-	Protein family
ECS-										
2,3	0.1		795.3	3+		VDHDHD$(H)_6$PGSSV$(G)_7$		159232571		pHpG
	1.0	5246			Nt	ECESGPCCRNCKFLKEGTI		159162948	00086	SVMP disintegrin
4	3.6		444.1	1+		ZKW		159232571		SVMPi
5	0.9	5416			Nt	ECESGPCCRNCKFLKEGTI		159162948	00086	SVMP disintegrin
6	0.8	14762			Nt	NSVHPCCDPVKCEPREGE		6014970	00036	Disintegrin EC3A
					Nt	NSVHPCCDPVKCEPREGE		6014971		Disintegrin EC3B
7	0.9	14807			Nt	NSVHPCCDPVTCEPREGE		17375443	00035	Disintegrin EC6A
					Nt	NSVHPCCDPVTCKPKRGK		17375444		Disintegrin EC6B
8-11	1.1	23-26 kDa					142	297593852	00056/ 00213	SVMP (frag)
12,13	1.3	23 kDa					208	297593828	00012	SVMP (frag)
14,15	0.5	26 kDa	1425.7	1+		XFCEXXKNTCK	de novo	297593852	00056/ 00213	SVMP (frag)
			1222.5	1+		AESYFYCRK	de novo			
			1094.5	1+		AESYFYCR	de novo			
16	4.9	14014					207	2499430	00014	S49-PLA_2
		13834	866.4	1+		NLNTYNK	276	2499430	00014	S49-PLA_2
			471.2	2+		GPPLDATDR				

			1441.6	1+		YTYYPNFWCK				
			1014.9	2+		SPFPSYTSYGCFCGGGER				
			2243.9	1+		CCLAHSCCYDTLPDCSPK				
			820.8	2+		ENGEIICENSTSCK				
17	1.9	13834			Nt	SVVELGKMIIQETGK	117	2499430	00014	S49-PLA_2
		13654			Nt	NLYQFGRMIWNRTGKPA	202	~27734436	00002	D49-PLA_2
18	3.8	13851			Nt	SIVELGKMIIQETGK	221	2499430	00014	S49-PLA_2
		13 kDa					282	2499430	00014	S49-PLA_2
19	0.8	14848					202	297594088	00036	Disintegrin
20	0.9	14808					117	17375443	00035	Disintegrin EC6A
21,22	3.1	24642			Nt	NVDFDSESPRKPEIQN	189	215262114	00168	CRISP
			582.7	2+		NVDFDSESPR				
			640.6	3+		KPEIQNEIIDLHNSLR				
			764.9	2+		WAFQCILDHSPR				
			835.8	2+		SNCAASCFCHSEIK				
23	1.0	38 kDa				VIGGAECNINEHRFLAFVY	98	116113	00244	Serine proteinase
			552.6	2+		IYFGLHNLK				
			963.6	3+		LDKPVTSSTHIAPISLPSSPPSVGSVCR				
	0.8	34 kDa	734.6	2+		VIGGAECNINEHR	91	297593764	00186	Serine proteinase
			552.6	2+		IYFGLHNLK				
24,25	0.6	35 kDa					92	116113	00244	Serine proteinase
26	2.2	16 kDa			Nt	DFDCPPEWSTYDQYCYKA	32		00006	C-type lectin-like
			1196.6	1+		WTDGSNLIYK	de novo			

			1243.5	1+		FTKNCFGLEK	de novo			
			1260.5	1+		CGDDYPFVCK	de novo			
		12 kDa			Nt	AAFCCPIGWSSYDQNCYKA	39		00230	C-type lectin-like
			1030.5	1+		MANHWSR	de novo			
			1391.6	1+		TDCSGTHNIVCK	de novo			
			1500.8	1+		AWNVEANCFVYK	de novo			
			2008.1	1+		DSHLVSLHNIAEADFVIK	de novo			
27	0.8	15 kDa					91	40889261		C-type lectin-like
		13 kDa	1353.6	1+		GYCYFVFNQR	92	~82131555	00154	C-type lectin-like
			1477.7	1+		AWDNDLHCFTAK	de novo			
			1908.8	1+		DCHWEWSDGAQLDYK	de novo			
28	1.4	56 kDa					282	297593794	00053	PIII-SVMP
	3.2	48 kDa					282	297593794	00053	PIII-SVMP
	0.8	15 kDa			Nt	DQDCLSGWSFYEGHCYK	135	40889261	00102	C-type lectin-like
	0.8	13 kDa			Nt	DCPQDWFVYKGYCYFVF	225	~82131555	00154	C-type lectin-like
29,30	7.1	52 kDa					159	297593826	00257	PIII-SVMP
31	4.4	48 kDa					134	297593798	00053	PIII-SVMP
		18 kDa					104	38493055	00006	C-type lectin-like
32	8.6	53 kDa					180	297593818	00053	PIII-SVMP
33	3.8	52 kDa					98	297303482	00473	PIII-SVMP
35	8.9	55 kDa					213	297593804	00177	PIII-SVMP
	0.2	52 kDa					114	297593804	00177	PIII-SVMP
	0.2	17 kDa					112	~73621141	00038	C-type lectin-like

35-43	0.2	13,15 kDa	915.5	1+	BFSFFVVK	de novo			C-type lectin-like
			511.3	2+	TADNBWMR	de novo			C-type lectin-like
			1040.5	1+	TNWYEAEK	de novo			C-type lectin-like
			1353.7	1+	GYCYFVFNBR	de novo			C-type lectin-like
			1477.7	1+	AWDNDLHCFTAK	de novo			C-type lectin-like
36	2.8	56 kDa				167	297593820	00043	PIII-SVMP
37	0.3	58 kDa				102	297593828	00062	PIII-SVMP
	0.5	42 kDa				289	297593828	00062	PIII-SVMP
38	5.2	42 kDa				293	297593828	00062	PIII-SVMP
39	2.4	48 kDa				322	320579395	00087	PIII-SVMP
	2.3	23 kDa				91	297594080	00117	PI(PII)-SVMP
		13 kDa				86		00154	C-type lectin-like
40	4.1	48 kDa				318	297594078	00012	PII-SVMP (188-390)
	0.1	23446				102	297594078	00012	PI(PII)-SVMP
41	0.1	48 kDa				179	297594072	00253	PIII-SVMP
	0.6	26 kDa				79	297594078	00012	PI(PII)-SVMP
42	0.6	52 kDa				56		00163	PIII-SVMP
	0.3	48 kDa				45	297593828	00062	PIII-SVMP
	0.2	26 kDa	638.6	2+	NINFDNNVIGR	79	297594078	00012	PI-SVMP
	2.1	23 kDa	638.6	2+	NINFDNNVIGR	102	297594078	00012	PI-SVMP
43	2.1	42 kDa				80	297594072	00253	PIII-SVMP
44	1.9	23196	753.1	2+	DLINVVSSSSDTLR	111	297594070	00253	PI(PII)-SVMP

45	2.1	23196	753.1	2+	DLINVVSSSSDTLR	106	297594070	00253	PI(PII)-SVMP
	1.9	23184				80	297594072	00253	PI(PII)-SVMP

Protein assignment undertaken by MALDI-TOF-TOF- and nESI- CID-MS/MS of peptide ions obtained from in-gel digested protein bands separated by SDS-PAGE. X = Ile or Leu; Z, pyroglutamic acid; M_{ox}, methionine sulphoxide. Cysteine residues are carbamidomethylated. Nt, N-terminal sequence determined by automated Edman degradation. Apparent molecular masses (in kDa) were estimated from SDS-PAGE analysis of reduced samples.

Table S4. Assignment of reverse-phase fractions from the venom of *Echis ocellatus* (Eoc) (Nigeria) to known protein families.

Spot ID	Spot %	Mass	m/z	z		MS/MS or N-terminal (Nt-)-derived sequences	Protein ID NCBI	Protein ID EOC-	Protein family
EOC									
2,3	0.4		573.4	2+		AKKKDEAPK	CAJ01681		
			538.1	2+		KKKDEAPKM	CAJ01681		
			750.3	3+		DEGTKCGEGKVCSNGYCVDL	CAJ01685	00024	SVMP fragment
4	5.7		444.2	1+		ZKW		00007	SVMPi
			859.7	3+		DCESGPCCDNCKFLKEGTICK	Q3BER1	00006	SVMP disintegrin
			927.4	3+		DCESGPCCDNCKFLKEGTICKMA	Q3BER1	00006	SVMP disintegrin
6	2.3	5591.7			Nt	DCESGPCCDNCKFLKEGTICK	Q3BER1	00006	SVMP disintegrin
		5494.3			Nt	DCESGPCCDNCKFLKEGTICK	Q3BER1	00006	SVMP disintegrin
		5112.8			Nt	DCESGPCCDNCKFLKEGTICK	Q3BER1	00006	SVMP disintegrin
		4765.1			Nt	DCESGPCCDNCKFLKEGTICK	Q3BER1	00006	SVMP disintegrin
10	1.5	14489.4			Nt	NSAHPCCDPVTCQPKQGEHCI	Q3BER6		Disintegrin EO4B
		14584.6					Q3BER3		Disintegrin EO4B
		14684.6					Q3BER6/R3		Disintegrin EO4B
11	0.6	14520.5			Nt	NSAHPCCDPVTCQPKQGEHCI	Q3BER4		Disintegrin EO5
12	0.1	14502.3	721.3	2+		QGEHCISGPCCR			Disintegrin
			520.8	2+		FLNSGTICK			
			886.4	2+		NSAHPCCDPVTCQPK			
13	0.5	24 kDa	533.7	2+		FKPAGTECR		00063	SVMP DC-frag
			540.7	2+		GKSYFYCR			

14	<0.1	23 kDa	533.7	2+		FKPAGTECR		00086/95	SVMP DC-frag
			826.4	2+		AVVGQDVCFEENKR			
			805.6	3+		NECDLPEYCTGQSAECPIDR			
15	<0.1	24220	533.7	2+		FKPAGTECR		00063	SVMP DC-frag
			540.7	2+		GKSYFYCR			
			762.9	3+		LHSWIECEFGECCDQCR			
			810.3	3+		NTCKYDYSEDPDYGMVGQGTK			
			941.7	3+		SECDLPEYCTGQSVDCPIDHFHR			
16	<0.1	21515	548.7	2+		AESYFYCR		00186	SVMP DC-frag
			714.3	2+		LFCEIIENTCK			
			641.6	3+		NGQPCLNNYGYCYNGK			
			661.9	3+		YDYSEDPNYGMVDEGTK			
			918.8	3+		SECDLPEYCTGQSADCPTDHFHK			
17	<0.1	23638	559.8	2+		LGNTYAYCR		00022/24	SVMP DC-frag
			606.3	2+		CPLTLYQCR			
			850.9	2+		DVVGVQESCFQYNR			
			819.9	2+		SFGDYISCLPCYR			
			738.3	3+		LHSWVECESGECCDQCR			
18	0.3	23 kDa	591.8	2+		NQCISLFGSR		00081	SVMP DC-frag
			648.8	2+		NPCQIFYTPR			
			718.8	2+		LYCFDNLPEHK			
			857.4	3+		AIVAEDACFQFNSLGIDYGYCR			
19	8.8	13825.6			Nt	SVVELGKMIIQETGKSPFPS	P48650	00015	S49-PLA_2
			2000.8	1+		SPFPSYTSYGCFCGGGEK			
			1368.8	1+		YTYYPNFLCK			

			1535.6	1+		CCFVHSCCYDK			
20	1.7	13861.6			Nt	SVIEFGTMIIEETGRSPFPF	AAN77204	00079	D49-PLA_2
21	0.4	29627.6	505.8	2+		FLTQYNPK	CAJ01685	00022/24	SVMP DC-frag
			516.3	2+		CMINKPLR			
22	0.5		906.9	2+		TDIISPPVCGNELLER	CAJ01685	00022/24	SVMP DC-frag
			501.3	2+		AGTLCRPAR			
			606.3	2+		CPLTLYQCR			
			850.9	2+		DVVGVQESCFQYNR			
			819.9	2+		SFGDYISCLPCYR			
			559.7	2+		LGNTYAYCR			
23	0.7	23818.4	582.8	2+		GETYLIEPLK	CAJ01681	00006	PI(PII)-SVMP
			844.3	2+		SYQFSDCSMNEYR			
24	1.2	28 kDa	797.2	2+		TXCAGXXEGGXDSCK	AAR24534		Serine proteinase
			598.6	2+		XMGWGSXTTTK		00181	Serine proteinase
25	0.1	24753.4	769.8	2+		MEWYPEAAANAER		00029	CRISP
			582.7	2+		NVDFDSEGPR			
			640.6	3+		KPEIQNEIIDLHNSLR			
26	2.6	29610.8							
		16 kDa	611.5	3+		DCXPGWSSHEGHCYK		00125	C-type lectin-like
			808.3	2+		QSGXDFVWTGXTYK		00124	C-type lectin-like
			634.6	2+		WVNXYCEER		00521	C-type lectin-like
		14 kDa	598.3	2+		CDWGWTNGAK		00372	C-type lectin-like
			700.8	2+		DCSSGWTAYGK			
			636.0	3+		LTAQTLESQIVWMGLSK			
27	0.4	31 kDa	452.1	2+		TLCAGILR		00049	Serine proteinase

			755.6	2+	LETPWAAKNVQPR			
			629.2	2+	SSELVIGGAECDINEHR		00315	Serine proteinase
			649.6	4+	VPHCANLEILDYSVCR			
	<0.1	16 kDa	808.3	2+	QSGXDFVWTGXTYK		00124	C-type lectin-like
			634.6	2+	WVNXYCEER		00521	C-type lectin-like
	<0.1	14 kDa	598.3	2+	CDWGWTNGAK		00372	C-type lectin-like
			700.8	2+	DCSSGWTAYGK			
28	0.9	29451.3	505.8	2+	FLTQYNPK	CAJ01685	00022/24	SVMP DC-frag
			516.3	2+	CMINKPLR			
			906.9	2+	TDIISPPVCGNELLER			
			501.3	2+	AGTLCRPAR			
			606.3	2+	CPLTLYQCR			
			850.9	2+	DVVGVQESCFQYNR			
			819.9	2+	SFGDYISCLPCYR			
			559.7	2+	LGNTYAYCR			
29	0.2	33 kDa	614.8	2+	VFDYDDWXR	CAD86932		Serine proteinase
			715.8	2+	SXPSSPPSVGSTCR		00181	
			735.8	2+	VXGGAECNXNEHR		00315	
	0.1	16 kDa	511.2	2+	YNVWIGLR		00038	C-type lectin-like
			632.2	2+	AEFLTQLVSQK			
			683.4	3+	FCNQWDGGHLVSIESTAK			
	0.1	14 kDa	517.7	2+	TTDNQWLR		00124	C-type lectin-like
			741.8	2+	AWNNELNCFVSK			
30	0.6	18 kDa	597.8	2+	CDWAWSNGAK		00056	C-type lectin-like
			629.3	3+	FCSEQANGGHLVSVHSR			

	0.7	17 kDa	743.8	2+	WVNFYCEKPSR	00133	C-type lectin-like
			424.8	2+	TWVDAEK		
			382.7	2+	VFNQEK		
			360.2	2+	CIGLEK		
31-32	0.8	18 kDa	629.3	3+	FCSEQANGGHLVSVHSR	00334	C-type lectin-like
			870.4	3+	EAGLVGVLAYQTLESEIIWMGLSK		
			737.9	3+	YEAWAEESYCIYIASNNK		
			686.9	3+	EWNSRPCEMFGHFACK		
			796.3	2+	EQQCNPEWNDGSK	00380	C-type lectin-like
			468.3	2+	IIYVNWK		
			630.9	3+	FCSEQANGGHLVSIQSR		
			410.6	2+	TWADAEK		
		17 kDa	424.8	2+	TWVDAEK	00133	C-type lectin-like
			743.8	2+	WVNFYCEKPSR		
		16 kDa	352.7	2+	WVPFR	00087	C-type lectin-like
			632.4	2+	AEFLTQLVSQK		
			511.3	2+	YNVWIGLR		
			575.4	2+	KYNVWIGLR		
			683.6	3+	FCNQWDGGHLVSIESTAK		
		14 kDa	517.7	2+	TTDNQWLR	00124	C-type lectin-like
			741.8	2+	AWNNELNCFVSK		
			829.8	2+	VDFVWIGMSDFWR		
33	2.1	52 kDa	665.4	2+	VTVLEASQLVGGR	00167	LAO
			647.8	2+	EGWYANLGPMR		
			690.8	2+	KFWEDDGIQGGK		
			748.7	3+	DIVVVGAGMSGLSAAYVLAGAGHK		
			729.4	2+	EADYEEFLEIAK		
			606.4	2+	NPLEECFR		

34	3.2	98 kDa	556.3	2+	TLDSFGEWR	CAJ01685	00022/24	PIV-SVMP
			691.4	2+	YVEFIIVVDQR			
			605.3	3+	TDIISPPVCGNELLER			
			738.3	3+	LHSWVECESGECCDQCR			
			819.8	2+	SFGDYISCLPCYR			
			851.4	2+	DVVGVQESCFQYNR			
		16 kDa	511.3	2+	YNVWIGLR		00087	C-type lectin-like
			575.4	2+	KYNVWIGLR			
			683.6	3+	FCNQWDGGHLVSIESTAK			
		14 kDa	517.7	2+	TTDNQWLR		00124	C-type lectin-like
			741.8	2+	AWNNELNCFVSK			
	0.7	56 kDa	556.3	2+	TLDSFGEWR		00022/24	PIII-SVMP
			691.4	2+	YVEFIIVVDQR			
			605.3	3+	TDIISPPVCGNELLER			
			738.3	3+	LHSWVECESGECCDQCR			
			819.8	2+	SFGDYISCLPCYR			
			851.4	2+	DVVGVQESCFQYNR			
35	2.2	56 kDa	483.2	2+	SYYDNFR		00089	PIII-SVMP
			710.8	2+	SVGVVEDYSPIVR			
			730.8	2+	AEDTLYSFGDWR			
			766.8	3+	IYEMLNTVNEIYLYLHIR			
			899.9	3+	HYYQNFLTDYKPDCTLIRPPR			
			547.8	2+	FRPAGTECR			
			738.2	3+	LHSWVECESGECCDQCR			
			915.6	3+	SECDLPEYCTGQSADCPTDVFHR			
			540.3	2+	YNNDLTAIR			
	0.5	52 kDa	591.8	2+	NQCISLFGSR	Q2UXQ0	00008	PIII-SVMP
			564.8	2+	LYCLDNSSR			
			810.3	2+	LGYCYNGDCPIMR			

			694.4	2+	VPLVGIVFWSNR			
	2.1	48 kDa	715.8	2+	APLVGIEFWNQR	Q2UXQ9	00063	PIII-SVMP
			762.9	3+	LHSWIECEFGECCDQCR			
			836.9	3+	TWIYeMVNTVNeIYLPLNIR			
			533.8	2+	FKPAGTECR			
			941.4	3+	SECDLPEYCTGQSVDCPIDHFHR			
36	10.2	62 kDa	533.8	2+	FKPAGTECR		00089	PIII-SVMP
			541.9	2+	GESYFYCR			
			826.4	2+	AVVGQDVCFEENKR			
			746.3	3+	LHPWVECETGECCDQCR			
			805.6	3+	NECDLPEYCTGQSAECPIDR			
37	4.2	56 kDa	799.6	2+	FSVGVVEDYS(514.1)			PIII-SVMP
			512.4	2+	GTGYFYCR			
	5.3	52 kDa	767.1	3+	LHSWIECEFGECCEQCR		00013	PIII-SVMP
			665.4	2+	HIELVIVVDHR			
			710.9	2+	LFCEIVPNICR			
			533.8	2+	FKPAGTECR			
			501.9	2+	GEGDFYCR			
38	5.6	62 kDa	591.8	2+	NQCISLFGSR		00081	PIII-SVMP
			605.8	2+	CPLTLYQCR			
			648.9	2+	NPCQIFYTPR			
			718.8	2+	LYCFDNLPEHK			
39	2.2	56 kDa	857.4	3+	AIVAEDACFQFNSLGIDYGYCR			
			640.9	2+	VAIVADYLIFR		00001	PIII-SVMP
			672.4	2+	MPQCILIKPSR			
			819.9	2+	IYEILNILNEIYK			

			829.5	2+	HIKVAIVADYLIFR			
	3.0	48 kDa	745.9	3+	ITHDNAQLLTAVNLNGDTIGR		00028	PIII-SVMP
39a	2.7	52 kDa	710.9	2+	LFCEXVPNXCR		00013	PIII-SVMP
41-43	8.5	110/56 kDa	743.8	2+	GYCYNGNCPXXR		00008	PIII-SVMP
			836.9	3+	TWIYeMVNTVNeIYLPLNIR		00404	PIII-SVMP
42,43	7.8	23 kDa	745.9	3+	ITHDNAQLLTAVNLNGDTIGR	Q2UXQ3	00028	PI-SVMP
			626.6	2+	EILNSFGEWR			
			722.8	2+	AQDSYHFSDCSK			
			455.2	2+	ZQHFDPR			
42b	1.4	56 kDa	852.4	2+	YVQLVIVADHSMVTK		00071	PIII-SVMP
43	6.5	56 kDa	852.4	2+	YVQLVIVADHSMVTK		00071	PIII-SVMP
			844.4	2+	SYQFSDCSMNEYR			
			598.9	2+	(154.1)GECCDNCK			
			743.8	2+	GYCYNGNCPXXR			

Protein assignment undertaken by MALDI-TOF-TOF- and nESI- CID-MS/MS of peptide ions obtained from in-gel digested protein bands separated by SDS-PAGE. X = Ile or Leu; Z, pyroglutamic acid; M_{ox}, methionine sulphoxide. Cysteine residues are carbamidomethylated. Nt, N-terminal sequence determined by automated Edman degradation. Apparent molecular masses (in kDa) were estimated from SDS-PAGE analysis of reduced samples.

Table S5. Assignment of reverse-phase fractions from the venom of *Echis coloratus* (Eco) (Egypt) to known protein families.

Spot ID		Mass				MS/MS or N-terminal	Mascot	Protein ID		Protein family
ID	%		m/z	z		(Nt-)-derived sequences	score	NCBI	ECO-	
ECO-										
3,4	0.5	5631			Nt	GEECDCGSPADCQNP			00144	SVMP disintegrin
	0.8	5561			Nt	CDCGSPADCQNPCCD			00144	SVMP disintegrin
5	2.9		444.1	1+		ZKW		159232571		SVMPi
6	1.8	5395.6			Nt	GEECDCGSPADCQNP			00144	SVMP disintegrin
7	0.3		559.0	2+		(226.3)VSDAPVVPP	de novo			unknown
8-12	2.1	22-26 kDa					142	297593852	00406	SVMP
13	3.7	14485					171	182705261	00024	Disintegrin
14,15	1.8	13994			Nt	NLYQFGKMIKNKTGKPAMFSY	153	25453159	00186	D49-PLA_2
			1854.5	1+		FENEDIICGDDDPCR				
16	0.9	14328			Nt	NSAHPCCDPVTCQPRE	131	297594068	00024	Disintegrin
17	4.2	13692			Nt	SVIELGKMIVQLTNKTPAS	288	2499430	00035	S49-PLA_2
			946.5	1+		MIVQLTNK				
			524.8	2+		IYPNILCR				
			670.4	2+		YKIYPNILCR				
			401.7	2+		AVAICLR				
			1286.7	1+		AVAICLRENLK				
			1607.7	1+		ICECDKAVAICLR				

			1826.7	1+		WENGEIICENSTSCK				
			927.4	2+		TPASYVSYGCFCGGGDR				
			1059.5	2+		YKWENGEIICENSTSCK				
			2305.8	1+		CCFVHSCCYDTLPDCSPK				
18	3.3	13901	945.5	1+		TDQYKYK	220	~2499430	~00035	S49-PLA$_2$
		13720	946.5	1+		MIVQLTNK				
			980.4	1+		RICECDK				
			524.8	2+		IYPNILCR				
			1339.7	1+		YKIYPNILCR				
			941.4	2+		TPVSYVSYGCFCGGGDR				
			2117.8	1+		YKWENGEIICENSTSCK				
			2305.8	1+		CCFVHSCCYDTLPDCSPK				
19	3.8	13798			Nt	HLLQFENMIYQKTGKFA	254	25453160	00086	D49-PLA$_2$
			1563.8	1+		HLLQFENMIYQK				
			2086.9	1+		FAIIAYSNYGCYCGWGGK				
			1533.5	1+		CCFVHDCCYGR				
			1236.6	1+		VAANCFAENLK				
			1284.6	1+		YWLSSIIDCK				
20,21	4.2	24683			Nt	NVDFDSESPRKPEIQNEIID	348	~215262114	00025	CRISP
			1537.7	1+		MEWYPEAAANAER				
			1571.7	1+		CGENIYMSPYPIK				
			1462.6	1+		GGCAAAYCPSSAYK				
			1495.7	1+		EDEFINCNDLVK				
			1670.7	1+		SNCAASCFCHSEIK				
22	0.4	33 kDa					198	297593754	00285	Serine proteinase
23	0.4	41 kDa					113	297593768	00112	Serine proteinase
	0.3	39 kDa	1104.5	1+		TLCAGVLEGGK	91	297593768	00112	Serine proteinase

			1194.6	1+	IMGWGTISTTK				
			1284.6	1+	TMKIHFGVHSK				
24	0.9	38 kDa				125	297593768	00112	Serine proteinase
25	0.5	42 kDa				216	297593738	00182	Serine proteinase
26	0.7	36 kDa				144	297593782	00206	Serine proteinase
27	0.1	37 kDa				132	297593766	00042	Serine proteinase
	1.2	16 kDa				178	~82174837	00131	C-type lectin-like
	1.2	13 kDa				144	~33243080	00115	C-type lectin-like
28	0.8	37 kDa				289	297593766	00042	Serine proteinase
	1.6	16 kDa	1832.8	1+	GGHLASIESSEEGDFVAK	78	~82174837	00041	C-type lectin-like
			1538.8	1+	SSADYVWIGLWNK				
	1.6	13 kDa				142		00038	C-type lectin-like
29	0.1	42 kDa				81	~297593750		Serine proteinase
	0.1	16 kDa				76	~82174837	00041	C-type lectin-like
	0.1	13 kDa				106	~82090788	00038	C-type lectin-like
30,31	0.1	48 kDa				69	~297593766	~00042	Serine proteinase
	0.1	16 kDa	1832.8	1+	GGHLASIESSEEGDFVAK	76	~82174837	00041	C-type lectin-like
			1538.8	1+	SSADYVWIGLWNK				
			1260.6	1+	CGDDYPFVCK				
	0.1	13 kDa				106	~82090788	00038	C-type lectin-like
32	0.3	17 kDa				131		00127	C-type lectin-like
	1.6	16 kDa				264	~82090788	00038	C-type lectin-like
	1.6	13 kDa				279	~82090788	00038	C-type lectin-like

33	0.4	16 kDa			105		00153	C-type lectin-like
	0.4	13 kDa			264	~82090788	00038	C-type lectin-like
35,36	4.6	56 kDa			151	~347602327	00544	LAO
37	1.8	58 kDa			192	297593872	00034	PIII-SVMP
38	1.3	65 kDa			193	297593876	00010	PIII-SVMP
	6.0	58 kDa			125	297593890	00023	PIII-SVMP
	0.1	19 kDa	935.5 1+	IIYVNWK	42		00052	C-type lectin-like
			1569.7 1+	EQQCTSEWNDGSK				
			1903.9 1+	FCTEQANGGHLVSIQSR				
	0.1	16 kDa			109		00197	C-type lectin-like
39	4.6	65 kDa			133	297593904	00050	PIII-SVMP
	3.5	58 kDa			193	297593902	00012	PIII-SVMP
40	3.1	65 kDa			219	297593904	00050	PIII-SVMP
	0.8	58 kDa			236	297593874	00012	PIII-SVMP
41	0.4	58 kDa			112	297593848	00002	PIII-SVMP
42	0.6	58 kDa			183	297593880	00002	PIII-SVMP
43	2.6	58 kDa			174	297593942	00061	PIII-SVMP
	1.5	42 kDa			90	297593880	00002	PIII-SVMP
	1.3	40 kDa			218	297593880	00009	PIII-SVMP
	0.4	28 kDa	1433.8 1+	SVGVIEDYSPIVR	68		00009	PIII-SVMP (fragm)
			2270.9 1+	LHSWVECEFGQCCDQCR				
			2682.2 1+	SECDLPESCTGQSAECPTDVFHR				
	0.1	17 kDa	1279.8 1+	HITHFWIGLR	38		00197	C-type lectin-like
			1317.6 1+	EHMTWEEAER				

			1966.9	1+	SEWSDGSSVSYDNLHKR				
	0.1	14 kDa				78		00069	C-type lectin-like
44	0.8	58 kDa				122	297593866	00002	PIII-SVMP
	1.6	28 kDa	1701.9	1+	YIELVIVADHAMVTK	78	297594018	00011	PI(PII)-SVMP
			1892.8	1+	TWVHQIVNDMTVMYR				
			2239.1	1+	DLITVTSSAEDTLNLFGTWR				
45	4.6	58 kDa				89	297593954	00009	PIII-SVMP
	1.4	28 kDa				78	297594018	00011	PI(PII)-SVMP
	1.8	26 kDa	1653.8	1+	NDADSTASISTCNGLK	61		00020	PI-SVMP
			1701.9	1+	YIELVIVADHAMVTK				
46	6.4	58 kDa				111	297593956	00001	PIII-SVMP
	0.3	24 kDa				129	297594002	00047	PI-SVMP
47	1.1	58 kDa				110	297593880	00002	PIII-SVMP
48	1.6	58 kDa				140	297593954	00009	PIII-SVMP
	3.8	24 kDa	915.5	1+	ERDLLNR	de novo		00017	PI-SVMP
			1066.6	1+	SFGEWRER	de novo			
49	2.1	58 kDa				150	297593866	00002	PIII-SVMP

Protein assignment undertaken by MALDI-TOF-TOF- and nESI- CID-MS/MS of peptide ions obtained from in-gel digested protein bands separated by SDS-PAGE. X = Ile or Leu; Z, pyroglutamic acid; M_{ox}, methionine sulphoxide. Cysteine residues are carbamidomethylated. Nt, N-terminal sequence determined by automated Edman degradation. Apparent molecular masses (in kDa) were estimated from SDS-PAGE analysis of reduced samples.

Table S6. Assignment of the reverse-phase fractions from the venom of *Echis pyramidum leakeyi* (Epl) (Kenya) to known protein families

Spot ID		Mass				MS/MS or N-terminal	Mascot	Protein ID		Protein family
ID	%		m/z	z		(Nt-)-derived sequences	score	NCBI	EPL-	
EPL										
4	1.6	5556			Nt	DCASGPCCRDCKFLKEGTI		82207847	00006	SVMP disintegrin
5	3.9		444.1	1+		ZKW		159232571		SVMPi
6	2.7	5435 5296			Nt	CASGPCCRDCKFLKEGTI		82207847	00006	SVMP disintegrin
7	1.4		559.0	2+		(226.3)VSDAPVVPP	de novo			unknown
8	0.4	12 kDa					103	~82194569		SVMP disintegrin
9	0.5	23 kDa				NLYQFGKMIKNKTGKPAM	107	~27734436	00001	D49-PLA_2
10,11	12.3	14103				NLYQFGKMIKNKTGK	358	27734438	00071	D49-PLA_2
12	0.4	22 kDa				SVIELGKMIIQLTNKTPAS	186		00012	S49-PLA_2
	7.2	13697				SVIELGKMIIQLTNKTPAS	334		00195	S49-PLA_2
			1825.9	1+		TPASYVSYGCFCGGGDK				
			1082.7	1+		IYPNFLCR				
			2305.7	1+		CCFVHSCCYDTLPDCSPK				
13	0.8	36 kDa					132	~297593760		Serine proteinase
14	0.2	28 kDa					101	~71041900		CRISP
15	0.5	38 kDa					270	~297593784		Serine proteinase

16	0.4	35 kDa				151	~297593732	00435	Serine proteinase
17	1.2	42 kDa				175	297593750	00502	Serine proteinase
18,19	3.0	46 kDa				118		00098	Serine proteinase
	0.6	16 kDa				219		00053	C-type lectin-like
	0.6	13 kDa				81		00018	C-type lectin-like
23,24	0.3	56 kDa				204	~347602327	00544	LAO
	1.2	16 kDa	1311.7	1+	CVGLEEQTGYR	68		00016	C-type lectin-like
			1568.9	1+	VFNQEMTWADAEK				
			1949.1	1+	SSIYYVWIGLSYEGPSK				
	1.2	13 kDa				103		00066	C-type lectin-like
25	1.8	56 kDa				204	~347602327	00544	LAO
	0.2	16 kDa				97		00016	C-type lectin-like
	0.1	15 kDa				137		00010	C-type lectin-like
	2.4	13 kDa				115		00066	C-type lectin-like
26	1.5	20 kDa				133	~82131629	00081	C-type lectin-like
27	1.3	48-50 kDa				176	297593964	00008	PIII-SVMP
	1.4	17 kDa				145		00109	C-type lectin-like
	1.3	12 kDa				368	~82090802	00031	C-type lectin-like
28	8.9	46 kDa				306	297593962	00008	PIII-SVMP
29	2.0	52 kDa				137	~27805465	00159	PIII-SVMP
	1.8	46 kDa				80	297593962	00008	PIII-SVMP
30	4.9	55 kDa				297	297593966	00029	PIII-SVMP

31	6.2	56 kDa				261	297593996	00019	PIII-SVMP
	1.3	28 kDa	1701.9	1+	YIELVIVADHAMVTK	de novo	~297594018		PI-SVMP
32	3.5	63 kDa				164	297593958	00002	PIII-SVMP
33	0.6	67 kDa				121	297593982	00040	PIII-SVMP
	2.1	48 kDa				133	~297593936	00008	PIII-SVMP
	0.5	24 kDa				91	~297594080	00005	PI-SVMP
34	0.6	63 kDa				150	297593984	00396	PIII-SVMP
	4.1	23 kDa				91	~297594080	00005	PI-SVMP
35	0.6	63 kDa				118	297593996	00019	PIII-SVMP
	5.6	24 kDa				88	~297594080	00005	PI-SVMP
36,37	5.5	55 kDa				198	297593974	00044	PIII-SVMP
35-37	0.1	16 kDa	1469.6	1+	TWEEAEKFCNR	de novo		00010	C-type lectin-like
			1476.7	1+	CFVLNQYTEFR				
		15 kDa				148	~205275155	00066	C-type lectin-like
38	0.5	56 kDa				147	297593996	00019	PIII-SVMP
	4.2	42 kDa				92	~297593828	00032	PIII-SVMP

Protein assignment undertaken by MALDI-TOF-TOF- and nESI- CID-MS/MS of peptide ions obtained from in-gel digested protein bands separated by SDS-PAGE. X = Ile or Leu; Z, pyroglutamic acid; M_{ox}, methionine sulphoxide. Cysteine residues are carbamidomethylated. Nt, N-terminal sequence determined by automated Edman degradation. Apparent molecular masses (in kDa) were estimated from SDS-PAGE analysis of reduced samples.

Table S7. Functional characterisations of *Echis*, *B. arietans* and *C. cerates* venom.

	LD_{50}		
	Mouse: µg/mouse (95% confidence intervals)	**Scorpion:** µg/µl (95% confidence intervals)	**Locust:** µg/µl (95% confidence intervals)
B. arietans	4.26 (2.53-5.13)	82.15 (49.62-120.31)	597.30 (376.80-921.00)
C. cerastes	7.71 (4.38-13.37)	Not determined	Not determined
E. c. sochureki	15.10 (6.49-19.70)	9.55 (6.91-11.93)	305.40 (193.20-468.00)
E. ocellatus	12.43 (9.00-20.45)	39.07 (27.34-50.26)	377.5 (236.30-593.40)
E. p. leakeyi	13.55 (8.98-38.33)	9.55 (6.91-11.93)	147.60 (92.70-218.00)
E. coloratus	9.81 (6.06-19.25)	139.89 (97.19-211.59)	394.40 (249.60-609.00)

	Haemorrhage	**Coagulation**		**ED_{50}**
	Lesion size in mm (SEM)	µg to clot 200µl **human** plasma	µg to clot 200µl **mouse** plasma	µl/mouse (95% confidence intervals)
B. arietans	12.10 (+/- 0.422)	No clotting at 100µg dose	No clotting at 100µg dose	No protection at 150µl dose[1]
C. cerastes	12.82 (+/- 0.465)	No clotting at 100µg dose	No clotting at 100µg dose	No protection at 150µl dose[2]
E. c. sochureki	10.28 (+/- 0.831)	0.35	0.52	No protection at 150µl dose[3]
E. ocellatus	10.82 (+/- 0.455)	0.09	0.25	58.46 (35.32-90.92)
E. p. leakeyi	9.33 (+/- 0.271)	0.44	0.25	64.87 (23.86-129.65)
E. coloratus	11.99 (+/- 0.732)	17.5	2.51	44.25 (21.90-58.29

SEM – standard error of the mean. EchITAbG is a monospecific anti-*E. ocellatus* antivenom. [1] 4/5 mice survived at this dose. [2] 0/5 mice survived at this dose. [3] 1/5 mice survived at this dose.

Table S8. Estimated models of sequence evolution determined by MrModelTest.

Toxin family	Model at each codon position		
	1	**2**	**3**
SVMP	GTR + I + Γ	GTR + I + Γ	GTR + Γ
SP	GTR + Γ	GTR + I + Γ	GTR + Γ
PLA_2	GTR + Γ	GTR + Γ	GTR + Γ
CTL	GTR + I + Γ	GTR + I + Γ	GTR + I + Γ
CRISP	GTR + Γ	HKY + Γ	GTR + Γ
LAAO	GTR + Γ	GTR + Γ	HKY + Γ

The models of sequence evolution implemented in MrBayes for each dataset were selected using the Akaike information criterion (AIC) in MrModelTest (http://www.abc.se/~nylander/mrmodeltest2/mrmodeltest2.html).

4. SI REFERENCES

1. Casewell NR, Harrison RA, Wüster W, Wagstaff SC (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.
2. Wagstaff SC, Harrison RA (2006) Venom gland EST analysis of the saw-scaled viper, Echis ocellatus, reveals novel α9β1 integrin-binding motifs in venom metalloproteinases and a new group of putative toxins, renin-like aspartic proteases. *Gene* 377:21–32.
3. Paine MJI, Desmond HP, Theakston RDG, Crampton JM (1992) Gene expression in Echis carinatus (Carpet Viper) venom glands following milking. *Toxicon* 30:379–386.
4. Calvete JJ (2011) Proteomic tools against the neglected pathology of snake bite envenoming. *Expert Rev Proteomics* 8:739–758.
5. Casewell NR et al. (2011) Domain loss facilitates accelerated evolution and neofunctionalization of duplicate snake venom metalloproteinase toxin genes. *Mol Biol Evol* 28:2637–2649.
6. Sanz-Soler R et al. (2012) Recombinant expression of mutants of the Frankenstein disintegrin, RTS-ocellatusin. Evidence for the independent origin of RGD and KTS/RTS disintegrins. *Toxicon* 60:665–675.
7. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797.
8. Casewell NR, Wagstaff SC, Harrison RA, Wuster W (2011) Gene tree parsimony of multilocus snake venom protein families reveals species tree conflict as a result of multiple parallel gene loss. *Mol Biol Evol* 28:1157–1172.
9. Ronquist F et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542.
10. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
11. Vernot B, Stolzer M, Goldman A, Durand D (2008) Reconciliation with non-binary species trees. *J Comput Biol* 15:981–1006.

12. Casewell NR et al. (2010) Pre-clinical assays predict pan-African Echis viper efficacy for a species-specific antivenom. *PLoS Negl Trop Dis* 4:e851.
13. Finney DJ (1947) *Probit analysis; a statistical treatment of the sigmoid response curve.* (Macmillan, Oxford, England).
14. Barlow A, Pook CE, Harrison RA, Wüster W (2009) Coevolution of diet and prey-specific venom activity supports the role of selection in snake venom evolution. *Proc Biol Sci* 276:2443–2449.
15. Richards DP, Barlow A, Wüster W (2012) Venom lethality and diet: differential responses of natural prey and model organisms to the venom of the saw-scaled vipers (Echis). *Toxicon* 59:110–116.
16. Theakston RD, Reid HA (1983) Development of simple standard assay procedures for the characterization of snake venoms. *Bull World Heal Organ* 61:949–956.