Venomics and antivenomics of Indian spectacled cobra (*Naja naja*) from the Western Ghats

Muralidharan Vanuopadath [1], Dileepkumar Raveendran [2], Bipin Gopalakrishnan Nair [1], Sudarslal Sadasivan Nair [1,*]

[1]School of Biotechnology, Amrita Vishwa Vidyapeetham, Kollam-690 525, Kerala, India

[2]Indriyam Biologics Pvt. Ltd., SCTIMST-TIMed, BMT Wing-Poojappura, Thiruvananthapuram 695 012, Kerala, India

*** Correspondence:** Sudarslal Sadasivan Nair, Ph.D.,
School of Biotechnology, Amrita Vishwa Vidyapeetham,
Clappana PO., Kollam-690 525, Kerala, India
E-mail: sudarslal@am.amrita.edu
Tel: 91-476-2803119
Fax: +91-476-2899722

Running Title: Venom proteomics of *Naja naja* from the Western Ghats

Abstract

Venom proteome profiling of *Naja naja* from the Western Ghats region in Kerala was achieved through SDS-PAGE and RP-HPLC followed by Q-TOF LC-MS/MS analysis, incorporating PEAKS and Novor assisted *de novo* sequencing methodologies. A total of 115 proteins distributed across 17 different enzymatic and non-enzymatic venom protein families were identified through conventional and 39 peptides through homology-driven proteomics approaches. Fourteen peptides derived through *de novo* complements the Mascot data indicating the importance of homology-driven approaches in improving protein sequence information. Among the protein families identified, glutathione peroxidase and endonuclease were reported for the first time in the Indian cobra venom. Immunological cross-reactivity assessed using Indian polyvalent antivenoms suggested that VINS showed better EC_{50} (2.48 µg/mL) value than that of PSAV (6.04 µg/mL) and Virchow (6.03 µg/mL) antivenoms. Western blotting experiments indicated that all the antivenoms elicited poor binding specificities, especially towards low molecular mass proteins. Second-generation antivenomics studies revealed that VINS antivenom was less efficient to detect many low molecular mass proteins such as three-finger toxins and Kunitz-type serine protease Inhibitors. Taken together, the present study enabled a large-scale characterization of the venom proteome of *Naja naja* from the Western Ghats and emphasized the need for developing more efficient antivenoms.

Key Words: Snake venom proteomics; *Naja naja*; Mass Spectrometry; *De novo* sequencing; Immunological cross-reactivity

1. Introduction

Snakebite envenoming is a major threat to the people living in the tropical regions of the world. It is noted that approximately half of global snakebite deaths are reported from India. Studies also showed that around 1.2 million people died in India due to snakebites, that too within a short period of about two decades (2000-2019), resulting in approximately 58,000 deaths per year [1]. It is equally disturbing to note that apart from the massive mortality rate, a three-fold increase in morbidity cases accounts for several limb amputations and other permanent disabilities [2]. Due to the degree of severity, chances of occurrence, lack of proper treatment, high costs involved in treating envenomed victims and increased mortality and morbidity rates, snakebite envenoming has been listed under the neglected tropical disease category by the World Health Organization (WHO) in June 2017 [3]. In India, spectacled cobra (*Naja naja*), common krait (*Bungarus caeruleus*), saw-scaled viper (*Echis carinatus*) and Russell's viper (*Daboia russelii*), commonly known as the 'Big Four' venomous snakes, account for majority of the morbidity and mortality cases [4].

The Indian cobra belongs to the genus Naja and family Elapidae is a highly venomous snake species. This species is native to countries of the Indian subcontinent including India, Sri Lanka, Bangladesh and Nepal [5]. They are seen in wide habitats like plains, dense or open forests, rocky terrains, wetlands, agricultural lands, outskirts of villages and even in highly populated urban areas. The species is also known under different names such as spectacled cobra, Asian cobra and binocellate cobra due to the presence of hoods behind the neck [6]. Snake venoms belonging to Elapidae family are rich in proteins that include three-finger toxins (3FTxs) and phospholipase A_2s (PLA_2s). Though these proteins are smaller in size, they are highly toxic and are responsible for most of the morbidity and mortality cases associated with the envenomation. The major contributory proteins in the three-finger toxins

family include post-synaptic neurotoxins, cardiotoxins/cytotoxins and muscarinic toxins [7]. Since *Naja naja* is distributed throughout India, studies were conducted to understand constituent proteins of *Naja naja* venom collected from different geographical locations and the data clearly indicate intra-species variation in the venom composition [5, 8-14]. The observed difference can be correlated to multiple factors including the geographical and prey-specific variations [15, 16]. Irrespective of all these, it is interesting to note that cobra venom from all the geographical areas are predominant with proteins belonging to three-finger toxin class.

Currently, mass spectrometry follows two principal approaches, top-down and bottom-up, for the structural characterization of venom components. The former deals with the in-depth analysis of proteins through gas phase fragmentation to establish their identity, identify post-translational modifications and characterise various isoforms [17, 18]. Whereas the latter involves proteolytic digestion of proteins into peptide fragments before mass spectrometric analysis. Since the bottom-up proteomics approaches use the information stored in public database repositories including NCBI and Uniprot, the success rate of protein identification through this strategy totally relies on the protein sequence information present in the database. However, due to the lack of inadequate protein sequence information in public repositories, in-depth characterisation of snake venom proteome remains a challenging one. Previous studies from our laboratory on venom proteome of various snake species had clearly demonstrated the importance of homology-assisted sequencing strategies in elucidating the venom profiles of snake species, whose sequence information is under-represented in the databases [19, 20]. In addition to this, the advantages of *de novo* sequencing followed by MS-BLAST and PEAKS analysis over conventional proteomics approaches have been substantiated by multiple reports [21-23].

Snake venomics enable characterising the venom arsenal of various snake species and antivenomics on the other hand utilise proteome information for extending the possibilities in improving antivenom preparations. Antivenomics could be also used for assessing the immunological cross-reactivity between homologous toxins present within various snake species [24]. Since antivenoms are the only choice available for treating envenomed victims, there is a need in improving the efficacy of the available antivenoms [25, 26]. Polyvalent antivenoms in India are manufactured using the pooled venom collected from the 'big four' snake species which includes *Naja naja*. Still the commercial antivenoms from the major Indian manufacturers showed differences in immunological reactivity profiles towards *Naja naja* venom collected from different geographical locations in India [8, 10, 11, 13, 14]. For example, though Indian polyvalent antivenoms contain enough antibodies against antigenic epitopes present in the cobra venom, they often fail to neutralize the low-molecular mass toxins such as 3FTxs and PLA_2s [8, 10]. Similarly, cobra venom from different geographical locations exhibited varied immunological cross-reactivity profiles to the polyvalent antivenoms used [8, 11].

We report here, new observations on the venom proteome of *Naja naja* venom from the Western Ghats region in Kerala through orthogonal separation and identification strategies. Also, we have evaluated the specificity and binding potential of commercially available antivenoms towards *Naja naja* venom through WHO approved in vitro preclinical protocols.

2. Materials and methods

2.1 Reagents and chemicals

Formic acid (FA), methanol, acetonitrile (ACN), trifluoroacetic acid (TFA), acrylamide, bisacrylamide, ammonium persulphate, coomassie brilliant blue R-250 (CBB R-250), beta-mercaptoethanol (β-ME), bromophenol blue, glycerol, glycine, iodoacetamide (IAm), ammonium bicarbonate, Tetramethyl benzidine/Hydrogen Peroxide (TMB/H_2O_2), HRP-

conjugated secondary antibody, tetramethylethylenediamine, dithiothreitol (DTT), trypsin, chymotrypsin and V8 proteases were purchased from Sigma-Aldrich. ECL substrate, Tris, SDS and PVDF membrane were procured from Bio-Rad. NHS-activated Sepharose 4 Fast Flow was obtained from GE Healthcare and protein ladder from Thermo Fisher Scientific.

2.2 Venom collection

Venom from two snakes belong to *Naja naja* species from the Western Ghats region in Kerala, one of the states in India, (NNwg) was collected in captive. Venom milking was done in-house (as per the Government order, GO (Rt) No. 94/2009/F & WLD dated 25/02/2009) in the Government serpentarium (Thiruvananthapuram, Kerala, India) by an experienced herpetologist. Followed by lyophilization, venom was stored at -80 ^{o}C until analysis.

2.3 Antivenoms

Lyophilized equine antivenoms raised against the 'Big Four' snakes of India (*Echis carinatus*, *Naja naja*, *Bungarus caeruleus*, and *Daboia russelli*) obtained from VINS (batch number: 01AS16040, expiry date: 07/2020), Virchow (batch number: PAS00116, expiry date: 01/2020) and Premium Serums and Vaccines (PSAV; batch number: ASVS (I) Ly-010, expiry date: 08/2022) were used for the entire immunological cross-reactivity assays.

2.4. Protein estimation using BCA assay

Protein estimation of crude venom and antivenom samples was performed using standard BCA assay [27]. For this, BSA with known concentrations was serially diluted from 2 mg/mL stock to get working concentrations of 1.5, 1.0, 0.75, 0.5, 0.25, 0.125, 0.025, and 0.03125 mg/mL. 10 µL from each of these standards was then mixed with 200 µL of BCA reagent (A and B solutions in a 1:50 ratio) and kept for incubation (30 min at 37 °C). The values were recorded at 562 nm using a plate reader.

2.5 SDS-PAGE and In-gel digestions of crude NNwg venom

12% SDS gel was used to resolve the crude NNwg venom (100 µg) proteins. The bands obtained were stained and visualized using CBB R-250 stain followed by scanning using a gel doc system (Bio-Rad). The gels bands separated under reducing conditions were then processed for in-gel digestion procedures. For this, the gel bands were cut into small pieces and excess stain was removed through 25 mM ammonium bicarbonate buffer and 50% ACN washing. This step was repeated until the gel bands were clear. Subsequently, the gel pieces were subjected to reduction using 100 mM DTT. The reaction was performed at 56 °C for 45 min. After incubation, the tube was allowed to cool at room temperature and excess of DTT was aspirated. Further, alkylation was performed using IAm (250 mM in ammonium bicarbonate buffer, pH 8.0). This was kept for incubation in the dark at room temperature for 30 min. After this, 25 mM ammonium bicarbonate buffer solution was used to wash the gel pieces for removing any leftover IAm. Subsequently, the gel pieces were dehydrated using 100% ACN for removing water content. Followed by this, in-gel digestion was performed using trypsin, chymotrypsin and V8 protease on three different gel preparations. For trypsin digestion, 25 µL of enzyme (from 12.5 ng/µL stock) was added to the gel pieces after dehydration (using 100% ACN). Approximately, 200 ng of chymotrypsin or V8 protease were added on to the individual gel pieces for the respective digestion processes. All the three enzymatic digestions were allowed to proceed at 37 °C for 16 h. The extraction of the digested peptides was carried-out using ACN and water containing 0.1 % TFA. The extraction step was repeated twice, and the resulting supernatants were pooled and were completely dried using a vacuum concentrator. The samples resuspended in 5% ACN solution were subjected to mass spectrometric analysis or stored at -80 ^{o}C until analysis.

2.6 Reversed-phase HPLC of NNwg venom

3.2 mg of crude venom was dissolved in 80 µL of Milli-Q water and was centrifuged at 12000 rpm for 15 min at 4 ^{o}C. From this, 5 µL was loaded on to a reversed-phase column

(Agilent Zorbax SB C8, 150 x 2.1 mm, 3.5 µM) through an Agilent 1260 HPLC system for the separation of various venom components. Solvent (A: water + 0.1% FA; B: ACN + 0.1% FA) flowrate was maintained at 200 µL/min. The following time program was used for the fraction collection: 5% B for 5 min, 5-15 % for 10 min, 15-45 % B for 60 min, 45-70 % B for 12 min, 70-95% B for 5 min and equilibration for 13 min. Absorbance was monitored at 210 nm and the peaks were manually collected and stored.

2.7 In-gel and In-solution enzymatic digestions of HPLC separated protein fractions

In-gel digestion of RP-HPLC separated fractions were performed as mentioned above (section 2.4). The fractions collected after reversed-phase HPLC was subjected to in-solution enzymatic digestions also. pH of the samples was adjusted to 8.0 using 50 mM ammonium bicarbonate buffer prior to in-solution digestion. Initially the samples were subjected to reduction using 100 mM DTT. This was then kept for incubation at 56 °C for 45 min. Alkylation was performed on the resulting solution using IAm (250 mM in ammonium bicarbonate buffer). The tubes were incubated at room temperature in the dark for 30 min. Followed by this, trypsin, chymotrypsin or V8 protease were added and kept for overnight incubation at 37 °C. The resultant solutions were acidified and dried completely using a vacuum concentrator.

2.8 Tandem mass spectrometry

The digested samples were reconstituted in 5% ACN containing 0.1% FA. Sample analysis was performed using an Agilent 6540 Q-TOF LC-MS/MS equipped with a chip cube source. The samples were infused to the MS through an HPLC system containing capillary and nano pumps that deliver solvents to the enrichment (40 nL) and analytical columns (Zorbax 300 SB-C18, 150 mm x 75 µm, 5 µ) present in the chip, respectively. The solvent (A: water with 0.1% formic aid; B: 90% acetonitrile with 0.1% formic acid) flowrate was maintained at 3 µL/min for the capillary and 0.3µL/min for the nano pump. The gradient conditions used for

the separation of the peptide samples were as mentioned: initial equilibration using 5% B for 2 min, 5-45 % B in 30 min, 45-90% B in 45 min, 90% B for 2 min. All the MS data were collected in a data-dependent and positive ionisation mode using nitrogen as the collision gas. The MS parameters used for the data acquisition were as follows: MS range- 250-3000 m/z; MS/MS-50-3000 m/z; Capillary voltage: 2000 V; Dry gas temperature: 325 ^{o}C; Dry gas flow: 4 L/min; Fragmentation voltage: 150 V; MS and MS/MS scan speeds were set at 4 spectra/sec and 3 spectra/sec, respectively. A maximum of five precursors were selected for the fragmentation. The charge states for tryptic digested samples were selected as 2, 3, >3 and for chymotrypsin and V8 digests: 1, 2, 3, >3. Active exclusion was enabled after 3 spectra with a releasing time of 0.5 min. MassHunter software (Version 6.01) from Agilent was used for the data acquisition and analysis.

2.9 Identification and validation of proteins using Mascot and Scaffold

The mgf files generated from the raw mass spectrometric data were analysed using Mascot (version 2.5.1, Matrix Science) through Mascot Daemon (version 2.5.0) interface with the following parameters: Database: Serpentes (Taxonomy ID:8570); peptide tolerance: 10 ppm, MS/MS tolerance: 0.5 Da; amino acid modifications: cysteine (carbamidomethylation as fixed modification), glutamine, asparagine (deamidation) and methionine (oxidation) as variable modifications.; number of missed cleavages: 2; Enzymes (selected charge states), trypsin (2^+, 3^+, 4^+), chymotrypsin and V8 protease (1^+, 2^+ and 3^+); the database searches were decoy enabled. The obtained data were further validated using Scaffold software (version 4.5.0, Proteomes Software Inc.) by setting the protein and peptide identification probabilities to 99% and 95%, respectively [28, 29]. The data were filtered using the inbuilt peptide and protein prophet algorithms and those proteins which were represented by a minimum of 2 peptides were included for the protein identification and classification process.

2.10 Protein identification through PEAKS

The obtained MS data was also subjected to protein identification using PEAKS Studio (version 8.5, Bioinformatics Solutions Inc., Canada) software. For the analysis, the precursor mass tolerance was set at 10 ppm and the fragment mass tolerance was set at 0.5 Da. Carbamidomethylation of cysteine was set as the fixed modification and deamidation of asparagine and glutamine along with the oxidation of methionine as variable modifications. Trypsin, chymotrypsin or V8 protease digestion options were selected according to the enzymes used. The confidently identified sequence tags (achieved by fixing the average local confidence to ≥30%) were then searched against Serpentes database. The peptides derived through *de novo* analysis were further validated through BLAST searches and were then manually verified [19].

2.11 *De novo* sequencing through Novor and MS-BLAST

In order to identify peptides with possible amino acid mutations, Novor followed by MS-BLAST sequence similarity searches were performed. For this, the raw data converted to mgf format was subjected to Novor analysis using the DeNovoGUI (Version 1.15.11) user interface [30]. Parameters used for the identification of proteins and peptides are: 10 ppm and 0.5 Da error tolerance for precursor and fragment masses, respectively; amino acid modifications: cysteine (carbamidomethylation as fixed modification), glutamine, asparagine (deamidation) and methionine (oxidation) as variable modifications. The obtained data after Novor analysis were pre-processed manually and subjected to MS-BLAST [31] analysis using Serpentes database as described previously [20].

2.12 BLAST search

Peptides with mutations obtained through *de novo* approaches using PEAKS and Novor followed by MS-BLAST were validated through BLAST search. For this, the obtained sequences were searched against the non-redundant (nr) database from NCBI (blastp suite at www.ncbi.nlm.nih.gov/BLAST) using the taxonomy, Serpentes (taxid: 8570) [32]. Following

BLAST, the identified peptide (s) representing various snake venom proteins were listed using the score and expectation (E) values obtained. Ideally, peptides with the lowest E-values were taken further. For matches with similar E-values and hit scores, proteins belonging to the same family (though they are having different accession numbers) and different snake species, the top hit was considered for the representation purpose. Also, if a *de novo* derived peptide represent a protein that was already identified through conventional database assisted proteomic analysis, the peptide was assigned to that particular protein for improving the sequence coverage. All the obtained sequences were then manually verified from the tandem mass spectra.

2.13 ClustalW

For confirming the sequences obtained for newly reported proteins from NNwg venom, the obtained peptide sequences were aligned from N-C terminus and subjected to BLAST search as mentioned above. Multiple sequence alignment of the top four protein hits along with the peptide sequences obtained through our proteomic characterization were represented through ClustalW.

2.14 Data availability

The raw mass spectrometry data and processed files have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) [33] via the PRIDE partner repository [34] with the dataset identifier PXD018936.

2.15 End point titration ELISA

ELISA was performed as described previously [19]. For this, 10 ng of NNwg venom added on to high bind ELISA plate (Costar) was kept for overnight incubation at 4□. The unbound venom was removed from the wells by inverting the plates and flick drying. Followed by this, any non-reactive sites in the wells were blocked by adding 2.5% BSA in phosphate buffered saline (PBS) and incubated for 2 h at room temperature. Primary antibodies were serially diluted using 2.5% BSA by a factor of three (1:100, 1:300, 1:900, 1:2700, 1:8100, 1:24300,

1:72900, 1:218700, 1:656100, 1:1968300, 1: 5004900) and was incubated for 2 h at room temperature. The unbound antibodies were removed through phosphate buffered saline with 0.1% tween-20 (PBST) wash and this was repeated thrice. Followed by this, secondary antibody (HRP conjugated) at a dilution of 1:32,000 in 2.5 % BSA was added on to each well and kept for incubation at room temperature for 2 h. Excess of secondary antibodies were removed through 0.1% PBST wash as mentioned above. To the wells, 100 µL of HRP substrate solution, TMB/H_2O_2, was added and incubated in the dark for 30 min at room temperature. The end point readings were taken by terminating the reaction by adding 100 µL of 0.5 M sulphuric acid to each well. The optical density values were recorded at 450 nm using a plate reader (Bio-Rad). The EC_{50} values were computed from the obtained OD values through non-linear regression analysis using Prism (version 6.01) software [19].

2.16 Western blotting

For immunoblotting experiments, 50 µg of crude NNwg venom resolved on a 15% SDS gel was transferred (25 V for 1 hr) to a PVDF membrane. To minimize non-specific binding, the membrane was blocked using 5% BSA at room temperature for 1 hr. Antivenoms obtained from Virchow, PSAV and VINS were used as primary antibodies (dilution used, 1:500) and was kept for overnight incubation at 4 □. The membrane was washed thrice for 10 min using 0.1% PBST solution. To this, secondary antibody (HRP conjugated) at 1:5000 dilution in 5% BSA was added and kept at room temperature for 1 hr. The excess of secondary antibody was then removed through 0.1% PBST wash as mentioned above. Followed by this, ECL substrate was added on to the membrane and signals indicating the venom-antivenom interaction were captured using a plate reader.

2.17 Second generation antivenomics

To determine the venom-antivenom interaction, second generation antivenomics protocol was adapted from a previous study [24]. For this, 600 µL of NHS-activated Sepharose 4 fast flow

was packed on to a column. The column was washed using 15 column volumes of ice cold 1 mM HCl for removing isopropanol (used for storing column material) and made sure that the pH of the column was maintained between 7.0 and 8.0. After this, the column was pre-equilibrated using coupling solution (0.2 M $NaHCO_3$ and 0.5 M NaCl, pH 8.3). To this 30 mg of anti-venom dissolved in the coupling buffer was added and kept for overnight incubation at 4 °C. Subsequently, unbound anti-venom was eluted out and the non-reacted groups were blocked using 0.1M Tris-HCl at room temperature for 4 hr. The column was then alternatively washed using buffers with high (0.1 M Tris-HCl; pH 8.5) and low pH (0.1 M acetate buffer containing 0.5 M NaCl; pH 4.5). After this, the column was equilibrated using 5 column volumes of working buffer solution (1X PBS, pH 7.4). Followed by, 200 µg of venom dissolved in PBS was added. After keeping it for incubation (4 h at room temperature), the unbound venom was washed out using PBS. The bound venom proteins were eluted using 0.1 M glycine-HCl, pH 2.0 and the resultant solution was immediately neutralised using 1M Tris buffer, pH 9.0. The bound and unbound fractions after desalting was subjected to RP-HPLC analysis as mentioned above.

2.18 Statistical analysis

Prism (GraphPad software Inc., San Diego, CA; version 5) was used for performing statistical comparisons. The data obtained from triplicate experiments were plotted on to the software and is represented as mean ± standard deviation or as stated, otherwise.

3. Results and discussion

3.1 Proteomics of *Naja naja* venom from the Western Ghats

The crude NNwg venom components were resolved on a 12% SDS gel. Bands observed after CBB staining indicated that the venom components were distributed across a wider mass range (Fig.1). All the visible bands excised from the samples run under reducing conditions

were subjected to various enzymatic digestion protocols followed by tandem mass spectrometric data collection. In parallel, a portion of the crude NNwg venom was fractionated on a reversed-phase column as mentioned in the methodology section. As indicated in Fig. 2a, a total of 12 fractions were collected. Each of these fractions were further separated under reducing conditions on a 12% SDS gel. From the results (Fig. 2b), it is evident that most of the fractions (labelled 1—8) fall around 10 kDa molecular weight region. All these peaks were eluted within 42 min (26% B) of the chromatographic run and rest of the peaks (labelled 9-12) found to contain proteins with molecular weights range between 10-315 kDa.

The proteins and peptides identified through orthogonal separation and multiple enzymatic digestion procedures are shown in Supplementary Table 1. From the Table, it is evident that NNwg venom contains proteins belonging to enzymatic and non-enzymatic families. A total of 115 proteins representing 17 different snake venom protein families were identified. As seen in Fig. 3a, 29% of the identified proteins belongs to 3FTx family followed by PLA_2 (10%) and SVMP (9%). The identified proteins classified according to their molecular masses indicated that almost 29% of them falls under 10 kDa region (Fig.3b). 3FTxs, the highly abundant proteins commonly found in elapidae snake species are seen in this mass range. Peptides that failed to appear in conventional database searches were subjected to homology-driven proteomic approaches using PEAKS and Novor based *de novo* sequencing followed by MS-BLAST analysis and manual validation. *De novo* analysis of a peptide representing cytotoxin 3a identified from NNwg is shown in Fig. 4. Manual inspection of the doubly charged peptide at m/z 730.4047 indicated that NNwg derived peptide from our study has an arginine instead of lysine (K12R) of the reference peptide. In a similar way, we could identify a total of 39 peptides representing thirteen snake venom protein families through *de novo* analysis. The identified peptides harbouring one or more amino acid mutations are

shown in Table 1. All the proteins identified from NNwg venom were then grouped under different snake venom protein families depending on the extent to which they mediate various biological functions and are outlined in detail below [7].

3.1.1 Dominant protein families identified from NNwg venom

The major enzymatic and non-enzymatic protein families identified from NNwg venom includes three-finger toxins, phospholipase A_2 (PLA_2s), snake venom metalloproteases (SVMP) and snake venom serine proteases (SVSP).

Snake species belonging to the Elapidae family are abundant with proteins representing three-finger toxins [7]. In agreement with these, our analysis resulted in the identification of 33 proteins representing 3FTxs (Supplementary Table 1) and contributed to 29% of the total proteins identified (Fig. 3a). Three finger toxins present in cobra venom belong to low molecular mass proteins (comprise of 57 to 82 amino acids), that include neurotoxins, cytotoxins/cardiotoxins and muscarinic toxins [35, 36]. Neurotoxins and muscarinic toxins bind to muscle nicotinic acetylcholine receptors (nAChRs) and muscarinic acetylcholine receptors, respectively [35, 37, 38]. As seen in Supplementary Table 1, almost all subclasses of 3FTxs including long and short neurotoxins, weak toxins, muscarinic toxin-like protein, cytotoxins/cardiotoxins were identified from NNwg venom. Most of the NNwg venom derived 3FTxs proteins showed sequence similarity with the proteins reported from various Naja species. Though Mascot searches enabled the determination of peptides representing long neurotoxin proteins belonging to 3FTx family, our *de novo* approaches helped in the identification of four additional peptides from the same category (Table 1). The peptide sequence with m/z, 1071.961, was observed to contain one amino acid substitution at the 49th position (K49R) when compared to the sequence present in the database (Suppl.Fig.1a (i)).

The peptide, DVLGCAATCPTVK, representing an Alpha-elapitoxin was also obtained. Comparing to the *Austrelaps labialis* counterpart, doubly charged NNwg peptide with m/z, 696.3388, was having V58D mutation (Suppl.Fig.1a (ii)). Similarly, through tandem mass spectrometric analysis and homology searches, the peptide at m/z, 484.2621 (z=3) was having asparagine in the position 8 instead of aspartate (Suppl.Fig.1a (iii)) and the doubly charged peptide at m/z, 679.7821 was contained to have a K23G substitution in the reference sequence (Suppl.Fig.1a (iv)).

Similarly, as seen in Supplementary Table1, our analysis resulted in the identification of 18 proteins representing different cytotoxins such as cytotoxin-1, 2, 7, 8, 9, 13 and 14. Additionally, three peptides representing cytotoxins were derived through *de novo* approaches (Table 1). The peptide at m/z, 789.8619, was found to have four amino acid mutations (K23R, V27W, A28N, and T29K) (Suppl.Fig.1a (vi)). Similarly the peptides, FPLRFPVKR and LIPLAYRTCPAGK, were having amino acid mutation at the 54th (K54R) and 12th (K12R) positions, respectively (Suppl.Fig.1a (v, vii)).

PLA_2s, the collection of proteins belonging to enzymatic venom protein family, are mainly involved in releasing lysophosphoglycerides and fatty acids by the catalytic cleavage of phosphoglycerides. They mediate multiple pharmacological effects such as platelet aggregation, anticoagulation, neurotoxic and cardiotoxic activities [39]. Most importantly, the myotoxic, neurotoxic and cytotoxic potential of several purified PLA_2s from the Indian cobra venom has been studied in detail [40-42]. As seen in Supplementary Table 1, twelve proteins representing several PLA_2s are identified through the present study. Additionally, a doubly charged tryptic peptide, NSGTPVDDLDR, with m/z 594.779, representing a PLA_2 protein was determined through *de novo* analysis (Table 1). On manual inspection, it was found that NNwg derived peptide has asparagine instead of glycine in the 25th position (Suppl.Fig.1b).

Altogether, the identified PLA_2s constitutes to 10% of the total proteome of NNwg venom (Fig. 3a).

Snake venom proteases, another major venom proteins belonging to the enzymatic family, are known to cleave fibrinogen and affect haemostasis through diverse anti- and pro-coagulant mechanisms [43]. Unlike viper venoms, both serine and metalloproteases are less abundant in elapidae species (12-19%) [7]. In accordance with this, our analysis indicated that SVMPs (9%) and SVSPs (6%) contribute only 15% of the total NNwg venom proteome identified (Fig.3a). SVMPs are zinc-dependent high molecular mass proteins with multidomain organization, divided into four main classes, PI, PII, PIII and PIV. They possess fibrinolytic, haemorrhagic and platelet aggregation inhibitory activities [44]. For example, kaouthiagin, a metalloprotease that cleaves and binds to von Willebrand factor (vWF) was isolated and characterized from *Naja kaouthia* venom. Cleavage of vWF results in collagen mediated platelet aggregation inhibition [45]. As seen in Supplementary Table 1, several peptides representing two kaouthiagin proteins were identified from the present study. Also, it is reported that NN-PF3, a non-toxic anticoagulant SVMP protein inhibits collagen induced plated aggregation through binding to α2β1 integrin receptors [46]. Like cobra venom SVMPs (NN-PF3 and kaouthiagin), several other PIII-$SVMP_S$ from different snake species were characterized for their platelet-aggregation inhibitory activities [47-49]. As seen in Supplementary Table 1, most of the SVMPs identified through our study belong to the PIII class. Though peptides representing various metalloproteases were derived through Mascot-enabled database search, twelve more additional peptides representing the same protein family were obtained through homology-assisted *de novo* searches (Table 1). One of the identified peptides, LYCTEKNTMSCR (L367R), was representing a kaouthiagin protein present in the database (Suppl.Fig.1c (i)). Similarly, the peptide, TAAAVVQDYSR, has Ala-130 and Arg-139 instead of Ser-130 and Ser-139 when compared to the *N.atra* counterpart

(Suppl. Fig.1c (ii)) and the peptide CPIMTNQCIALRGPRK (m/z, 644.3293), was found to have two amino acid substitutions at the C-terminus (G524R, V525K) compared to the original sequence present in the database (Suppl.Fig.1c (iii)). Another metalloproteinase peptide sequence, VTLDLFGEWKER, was shown to have an amino acid mutation, R43K (Suppl.Fig.1c (iv)). Apart from this, eight more peptides representing multiple SVMPs were identified through *de novo* approach. The details of all the *de novo* derived peptides are provided in Table 1 and Suppl.Fig.1c (v-xii).

SVSPs, on the other hand, are mainly composed of thrombin-like and fibrinolytic proteins. Thrombin-like proteins are primarily used as defibrinogenating agents due to its ability to remove fibrin (formed during clotting) from the blood circulation [50]. As seen in Supplementary Table 1, all the NNwg venom derived thrombin-like proteins showed sequence similarity to peptides reported from *Naja naja* and *Gloydius ussuriens* species. In addition to this, tissue-type plasminogen activator proteins were also present in NNwg venom. Since plasmin is involved in degrading clots formed by fibrin, tissue type plasminogen activators can be used as thrombolytic agents for treating deep vein thrombosis [51]. Apart from the proteins reported in Supplementary Table 1, our *de novo* analysis resulted in the identification of two additional peptides, GLCYQDLGDTYR, LFPDSFCTSAR, harbouring various amino acid mutations representing SVSP proteins (Table 1, Suppl.Fig.1d). In the former peptide sequence, three mutations were identified (A103G, M104L, G109L) (Suppl.Fig.1d (i)). Whereas in the latter sequence, tyrosine has been replaced by a phenylalanine (Y449F) (Suppl.Fig.1d (ii)).

3.1.2 Secondary protein families identified from NNwg venom

Proteins belonging to L-amino acid oxidase (LAAO), Kunitz-type serine protease inhibitors (KSPI), C-type lectins (CTLs) and Cysteine-rich secretory proteins (CRISP) falls under this category [7].

LAAOs are flavoenzymes with flavin adenine dinucleotide as cofactors. Flavins are accountable for the yellow colour of snake venom and they mediate cell toxicity by inducing oxidative stress through hydrogen peroxide release [52]. They are also involved in arbitrating platelet aggregation activation or inhibition, anticoagulation, myonecrotic and haemorrhagic activities [52, 53]. From the LC-MS characterization, it is evident that 7% of the total proteins identified from NNwg venom represents LAAO family (Fig.3a). All the NNwg venom derived LAAO peptides showed sequence similarity to the peptides reported from *Bungarus multicinctus*, *Oxyuranus scutellatus* and several other Naja species (Supplementary Table 1). On the other hand, *de novo* sequencing yielded two more additional peptides harbouring different amino acid mutations. As seen in Table 1, the peptides, FWEADGLHNK and IAETVHLNAR, representing an LAAO, showed sequence similarity with *Bungarus multicinctus* and *Naja atra* proteins, respectively. Manual inspection of the MS/MS data of FWEADGLHNK confirmed that the NNwg venom derived peptide has asparagine and lysine at the 360th and 361st positions, respectively (G362N and G363K) (Suppl.Fig.1e (i)). Similarly, the triply charged peptide (IAETVHLNAR) at m/z, 375.2134, was found to contain threonine instead on methionine (M273T) (Suppl.Fig.1e (ii)).

Kunitz-type serine protease inhibitors are known to act on several serine proteases and MMP-2s [54]. They inhibit activity of coagulation factors including thrombin, protease C, factors Xa and XIIa and also found to regulate blood pressure [55]. Considerable number of KSPIs are identified through proteomics analysis of *Naja naja* venom from different parts of India [5, 8-12]. Our study also resulted in the identification of eight different peptides representing two KSPI proteins (Supplementary Table 1). The *de novo* sequencing methods on the other hand has resulted in the identification of an additional peptide representing a kunitz-type inhibitor protein reported from *Walterinnesia aegyptia* venom (Table 1 and Suppl.Fig.1f). On

manual inspection, the doubly charged peptide with m/z, 927.9134, was contained to have four amino acid mutations (P43T, S44A, H50D and H52R).

Snaclecs, the C-type lectins found in snake venom are known to activate and inhibit platelet aggregation mechanisms [56]. Though venom proteome of Indian cobra has been studied from various geographical locations, only one study has revealed the presence of a CTL that showed sequence similarity to a C-type lectin BfL-1 from *Bungarus fasciatus* [12]. This study also identified peptides representing the same protein and are included in Supplementary Table 1. Through *de novo* approach, two peptides representing the above mentioned protein (Table 1) were identified. The peptides, YNCPLDWLPK and NNEFCVELVK, were contained to have one amino acid mutation (T25N, N134K) each (Suppl.Fig.1g (i & ii)).

Our proteomic analysis resulted in the identification of proteins representing CRISP family too. The identified peptides showed sequence similarity to CRISP proteins reported from two Naja species (Supplementary Table 1). Apart from this, three doubly charged peptides, LTNCDSLVK, NGDFNSQSTR and QNSCQDDWLK exemplifying CRISP proteins harbouring single amino acid mutations were identified from NNwg venom through *de novo* approaches (Table 1). For the peptide with m/z 525.272 (z=2), leucine in the 196th position has been substituted by valine (Suppl.Fig.1h (i)). Whereas the other two doubly charged NNwg venom derived peptides with m/z, 563.7368 and 647.2816 have V20G and S217N mutations, respectively (Suppl.Fig.1h (ii &iii)).

3.1.3 Minor protein families identified from NNwg venom

The minor proteins families present in snake venom includes 5'-nucleotidase (5-NU), Phosphodiesterase (PDE), Pospholipase B (PLB), Acetylcholinesterase (ACE), Nerve growth factor (NGF) and Vespryn/ohanin (OLP) [7].

The nuclease activity seen in snake venom is due to the presence of phosphodiesterases and 5-nucleotidases [57, 58]. Several of these enzymes with potent nuclease activity was isolated from various Naja species. In our study too, peptides representing seven PDEs and six 5-NUC have been identified (Supplementary Table 1). In addition, a peptide representing a PDE harbouring two amino acid mutations were also identified through the *de novo* analysis (Table 1). The peptide, AATYFWPGTDVK was contained to have S26T and I28V mutations with respect to the reference protein present in the database (Suppl.Fig.1I).

Since ohanin like proteins are present only in smaller amounts, they are not isolated or characterized in detail from snake venoms. Nevertheless, ohanin identified from king cobra has shown to induce hyperalgesia and hypolocomotion in mice [59]. In addition to six proteins identified through conventional approaches (Supplementary Table 1), our *de novo* analysis yielded a doubly charged tryptic peptide, FNGSPCVLGSPGFR, having a single amino acid mutation (S101G) representing ohanin protein family (Table 1 & Suppl.Fig.1j).

Further, the peptides identified through our analysis showed sequence similarity to cobra venom factor proteins reported from *Ophiophagus hanna*, *Naja kaouthia*, *Pseudonaja textilis* and *Bungarus fasciatus* (Supplementary Table 1). Apart from this, *de novo* characterization resulted in the identification of five additional peptides with various amino acid mutations representing several CVF proteins. As seen in Suppl.Fig.1k (i &ii) and Table 1, two of the *de novo* derived peptides, CPLPNLENGK and LILNTPLNAQSLPLTVR, were having one amino acid mutation. The doubly charged peptides, at m/z 571.2884 and 932.0644, were found to contain V96L and T409A mutations, respectively (Suppl.Fig.1k (i &ii)). On a similar way, three more peptides with multiple amino acid mutations representing CVF proteins were obtained through the study (Suppl.Fig.1k (iii-v)).

Though Mascot searches enabled the determination of peptides representing proteins belonging to endonuclease family (Supplementary Table 1), our *de novo* study helped in the

identification of an additional peptide with two amino acid mutations (Table 1). The triply charged peptide sequence with m/z, 664.0157, was observed to contain two amino acid substitutions at the 241st (N241L) and 245th (R245Q) positions when compared to the sequence present in the database (Suppl.Fig.1l).

Peptides representing glutathione peroxidase proteins were also identified from this study (Supplementary Table 1). In addition, a doubly charged peptide, SITGNYIAFR, with m/z 571.7955, representing the same family was identified through *de novo* analysis (Table 1). The obtained peptide showed sequence similarity to a GPX3 protein reported from *Thamnophis elegans* and through manual analysis it was confirmed that Glu-78 in the original sequence has been replaced by an asparagine (Suppl.Fig.1m).

Apart from the above-mentioned protein families, peptides representing acetylcholinesterases, phospholipase B and nerve growth factor were also identified from NNwg venom (Supplementary Table 1).

3.2 Newly identified protein families from the Indian Cobra venom

Though a total of 17 different snake venom protein families were identified, it is interesting to note that glutathione peroxidase and endonucleases were reported for the first time in Indian cobra venom. The peptide information representing these families of proteins are included in Supplementary Table 1 and Table 1. For confirming the authenticity of the derived peptides, sequences were aligned from N-C terminus and subjected to pBLAST using the non-redundant database from NCBI. The sequences obtained were further subjected to multiple sequence alignment using ClustalW. Figures 5a and b show the sequence similarity of the obtained sequences with respect to various snake species. As seen in Supplementary Table 1, the peptides representing GPrx 3 identified from NNwg venom showed 77% sequence similarity to a glutathione peroxidase reported in *Ophiophagus hannah*. Similarly,

the endonuclease identified in this study shown to have 53% sequence similarity with an *O.hannah* protein (accession no: ETE56723).

3.3. Importance of *de novo* sequencing strategies in elucidating NNwg venom proteome

Previous studies from our laboratory have shown the importance of *de novo* sequencing workflows in elucidating snake venom proteome profiles [19, 20]. It was observed that the use of complementary identification strategies improves the confidence of protein identification. In agreement with this, our analysis on *Naja naja* venom proteome through *de novo* approaches resulted in improving sequence information of proteins that were determined through conventional proteomic approaches. For example, though a SVMP protein, atrase-A was identified through conventional approaches (Supplementary Table 1, SI No-55), four additional peptides representing the same protein (accession no- ADF43026) were determined through *de novo* analysis. Likewise, a total of 14 peptides (Table1; accession numbers marked in bold) representing ten different snake venom proteins complements the conventional proteome data. Though venom gland transcriptomics has opened up the possibilities in exploring snake venom proteome profiles, due to its limitations, details on protein isoforms might be missed [60, 61]. For example, venom proteomics of *Sistrurus catenatus edwardsii* studied using different ionisation mechanisms using MALDI and ESI mass spectrometry enabled the identification of various protein isoforms that were not detected through the transcriptomics study [60]. Similarly, in the current study, several protein isoforms were derived through homology-driven approaches. As seen in Table 1, a peptide representing 3FTx (accession- 1CTX_A) protein was contained to have one amino acid mutation (K49R) in the *de novo* derived sequence and the original sequence without the mentioned mutation, presumed to belong to an isoform, was observed through Mascot analysis. Similarly, a total of five peptides representing various protein isoforms were derived through our analysis (Table 1; SI No-1, 3, 4, 5 &14). This is indicative of the fact that

homology approaches helps in the determination of protein isoforms too. In addition to this, 20 different *de novo* derived peptides representing various snake venom protein families were identified through *de novo* analysis (Table 1). These finding highlights the importance of following multiple data analysis approaches in elucidating the venom proteome of snake species whose protein sequence information is limited in the public repositories.

3.4 Commercial antivenoms showed varied immunological cross-reactivity profiles towards NNwg venom

From a clinical perspective, detailed information on the toxin components present in venom helps in streamlining antivenom production protocols. As recommended by World Health Organisation, various in vitro preclinical studies including ELISA, western blotting and immunoaffinity chromatography are widely used in determining the specificity and reactivity of commercially available antivenoms towards various snake venoms. In India, polyvalent antivenoms are manufactured by using pooled venom from 'Big Four' snake species that include *Naja naja, Bungarus caeruleus, Daboia russelii and Echis carinatus*. Still, commercial antivenoms from the major Indian manufacturers' present varied immunological reactivity towards cobra venom collected from different geographical locations in India [8, 10, 11]. In agreeing with the above observation, our results obtained through end-point titration ELISA indicate that the immuno-cross-reactivity of VINS antivenom (EC_{50}: 2.48 µg/mL) is better compared to Virchow (6.03 µg/mL) and PSAV antivenoms (6.04 µg/mL) (Fig. 6a). All the tested Indian antivenoms showed a dose-dependent increase in reactivity towards NNwg venom components. However, as seen in Fig. 6a, the binding potential seems to be varied across the tested antivenoms. Though the end point titration values of all the antivenoms were same (1:72,900), the strength of binding seen at the same dilution was higher in Virchow antivenom. At the same time, EC_{50} values computed from the graph

suggest that VINS (2.48 µg/mL) showed better reactivity towards NNwg venom compared to that of PSAV (6.04 µg/mL) and Virchow (6.03 µg/mL) antivenoms.

Similarly, the specificity of antivenom binding towards NNwg venom was assessed through immunoblot analysis. As seen in Fig. 6b (i), the resolved proteins were transferred on to a PVDF membrane and immunoblotting was performed using various polyvalent antivenoms as primary antibodies. From the results (Fig.6b (ii-iv)), it can be inferred that the binding specificities are varied across different antivenom preparations. Hence, there are notable differences in detecting the antigenic epitopes of NNwg venom by various antivenoms. It was evident that low molecular mass proteins were abundant in NNwg venom (Fig. 6b (i)). Irrespective of this, compared to low molecular mass proteins, all the antivenoms recognized high molecular mass proteins with better specificity. Previous reports indicate that high molecular mass proteins are less toxic but highly immunogenic compared to highly toxic low molecular mass proteins that are poorly immunogenic [62], and the same was observed from our findings. Hence, our observations are in agreement with the previous data that reveal the failures associated with Indian polyvalent antivenoms especially in recognizing small molecular mass proteins falling under 20 kDa region [8, 10]. The differences in the antivenom neutralisation efficiencies might also be contributed by the differences in the venom antigens that were used for immunizing horses while generating polyvalent antivenoms [63].

3.5 VINS antivenom was less efficient in detecting many low molecular mass proteins falling under 20 kDa

Though ELISA and western blotting can provide an overall immunoreactive profile of venom and antivenom interaction, they are not successful enough to distinguish the antivenom binding and non-binding protein components. To verify this, we have followed a second generation antivenomics approach combining immunoaffinity chromatography and LC-MS

based proteomic characterization [24]. Since the immuno-cross-reactivity assays showed a significant difference in binding potential of various antivenoms towards NNwg venom, second generation antivenomics approach was used in determining the identity of antivenom bound and unbound protein fractions. It was observed from the EC_{50} values that VINS antivenom showed better reactivity towards NNwg venom (Fig. 6a). Figure 7 (a-c) shows the UV chromatogram of VINS antivenom bound and unbound venom proteins. Though majority of the proteins are found in the bound fraction (Fig. 7c), some proteins (represented by the peaks, 1-4) are not getting recognized by the antivenom. From the MS analysis it was identified that 3FTxs, PLA_2s and KSPIs were present in those peaks, suggesting that VINS antivenom failed to detect few proteins falling under 20 kDa region (Fig. 7a). Altogether, our study reveals that the specificity of antivenoms varies across different manufactures. Hence, as recommended by WHO, the production and purification strategies need to be streamlined for improving the efficacy of antivenoms [64].

4. Conclusion

The study demonstrates the venom proteomics and antivenom immunological cross-reactivity of *Naja naja* venom collected from the Western Ghats region in Kerala. Peptides that failed to appear in conventional database searches were subjected to homology-driven proteomic approaches using PEAKS and Novor based *de novo* sequencing followed by MS-BLAST analysis and manual validation. This finding highlights the importance of following multiple data analysis approaches in elucidating the venom proteome of snake species whose protein sequence information is limited in the public repositories. Like other reports on cobra venom, our analysis also indicated that *Naja naja* venom from the Western Ghats is predominant with proteins belonging to three-finger toxin family. To date, there are a total of eight published reports on cobra venom proteome from different geographical locations of India (Table 3). Though the species remains the same, detailed venom proteome profiling of *Naja naja* venom

from different parts of India reports that venom composition varies according to the geographical location. The immunological cross-reactivity studies indicated that there are variations across Indian polyvalent antivenoms in detecting and binding to the antigenic epitopes present in *Naja naja* venom proteins. Altogether, our study reveals that the specificity of antivenoms varies across different manufacturers. This highlights the importance of improving antivenom preparations against *Naja naja* envenomation in India.

Acknowledgements

The authors would like to acknowledge Sri. Mata Amritanandamayi Devi, The Chancellor of Amrita Vishwa Vidyapeetham, for being the inspiration behind this study. We are grateful to the forest department of Kerala for their assistance during the field visits during the time of venom collection. The authors would like to thank Dr. Dileepkumar Raveendran and Indriyam Biologics for providing the venom, various chemicals and reagents including antivenoms, which were used for all the immunological cross-reactivity assays. MV thank Amrita Vishwa Vidyapeetham and Amrita Centre for Research and Development for providing funding in the form of fellowship. The authors also acknowledge Amrita Agilent Analytical Research Centre for the mass spectrometric data collection and analysis. The authors would like to thank Dr. Sanu K. Shaji, Ms.Meera Venugopal and Mr.Vishnu Raj for helping with the Western blotting experiments. Financial assistance in executing this project provided by Amrita Vishwa Vidyapeetham is also duly acknowledged.

Conflict of Interest

There are no potential conflict of interests.

Figure Legend

Figure 1. 12% SDS-PAGE analysis of crude NNwg venom. Lane 1,molecular weight ladder; Lane 2, Crude NNwg venom resolved without β-ME (non-reducing condition); Lane 3, Crude NNwg venom subjected to β-ME treatment (reducing condition). The bands

resolved under reducing conditions were further subjected to various enzymatic digestion (trypsin, chymotrypsin and V8 protease) strategies followed by Q-TOF LC-MS/MS analysis.

Figure 2a. Reversed-phase HPLC profile of crude NNwg venom. Crude *Naja naja* venom (500 µg) was subjected to reversed-phase HPLC separation using an Agilent 1260 HPLC system. X-axis and Y-axis corresponds to the retention time and absorbance observed at 210 nm, respectively.

Figure 2b. 12% SDS-PAGE of RP-HPLC separated fractions of NNwg venom. The 12 fractions collected after HPLC separation were resolved on a 12% SDS gel under reducing conditions. Each of these bands were further diced and subjected to trypsin, chymotrypsin and V8 protease digestion. The resulting peptides were further analysed by using a Q-TOF tandem mass spectrometer. (MW; molecular weight marker)

Figure 3. Relative abundance of protein families identified from NNwg venom. The relative abundance of NNwg venom proteins identified (a) after different separation and digestion protocols and (b) according to molecular weight. [Relative abundance = (Number of proteins identified representing a snake venom protein family/ Total number of proteins identified) X 100]. (PLA_2-Phospholipase A_2, PLB-Phospholipase B, SVMP- Snake Venom Metalloprotease, NGF- Nerve Growth Factor, CRISP- Cysteine Rich Secretory Protein, PDE-Phosphodiesterase, 5'-NUC- 5 Nucleotidase, SVSP- Snake Venom Serine Protease, 3FTx- 3 Finger Toxin, CVF- Cobra Venom Factor, LAAO- L-Aminoacid Oxidase, GPrx- Glutathione peroxidase, ACE- Acetylcholinesterase, OLP- Ohanin/ Vespryn family proteins, END-Endonuclease, CTL-C Type Lectin, GPrx-Glutathione Peroxidase)

Figure 4. *De novo* analysis of a peptide representing NNwg venom protein. Cobra venom derived peptide (shown in inset) has sequence similarity to Cytotoxin 3a (P86539; *Naja naja*) harbouring one amino acid mutation. The m/z values and corresponding ion intensities are labelled in the x- and y- axis, respectively. MS/MS spectrum representing b and y ions are

labelled along with the immonium ions (depicted by single letter amino acid code along with 'i' as subscript). The b7 and y7 ions representing the mutated amino acid arginine, R, are circled in the spectrum.

Figure 5. BLASTp and ClustalW analysis of endonuclease and glutathione peroxidase proteins identified from NNwg venom. (a) Endonuclease and (b) Glutathione peroxidase peptide sequences identified from NNwg venom after Mascot search were subjected to BLASTp analysis followed by multiple sequence alignment using known sequences from various snake species. Those highlighted in-boxes are the shared amino acid sequences between NNwg and other snake species.

Figure 6. Assessment of the immunological cross-reactivity of different polyvalent antivenoms towards NNwg venom. (a) Titration curves obtained through ELISA using various dilutions of VINS, PSAV, Virchow antivenoms. Results are represented as mean ± SD of triplicates. The EC_{50} values for different antivenoms obtained from this graph are shown in insets. (b) (i) 15% SDS-PAGE gel electrophoretic profile of crude NNwg venom under reducing conditions. Immunoblot of the SDS-PAGE with (ii) VINS (iii) PSAV and (iv) Virchow polyvalent antivenoms.

Figure 7. Immunoaffinity chromatography of NNwg using VINS antivenom (a) The protein peaks observed from crude NNwg venom after reversed-phase HPLC fractionation are labelled from 1-12 (b) the unbound protein fractions observed after PBS buffer wash and (c) bound protein fractions obtained after glycine buffer wash. Proteins present in the bound and unbound fractions were compared using the chromatogram obtained from the crude venom (shown in a). The proteins present in each peaks were then identified through proteomics approaches.

Figures

Figure 1. 12% SDS-PAGE analysis of crude NNwg venom

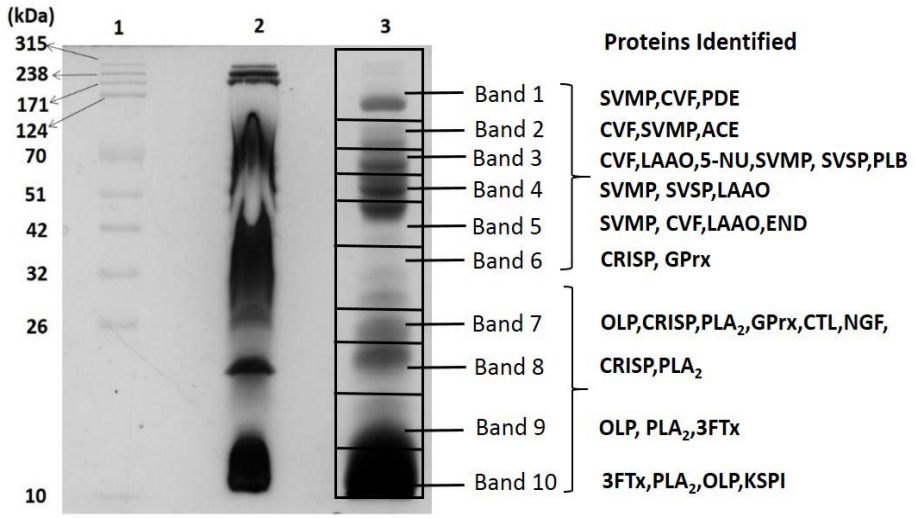


Figure 2a. Reversed-phase HPLC of crude NNwg venom

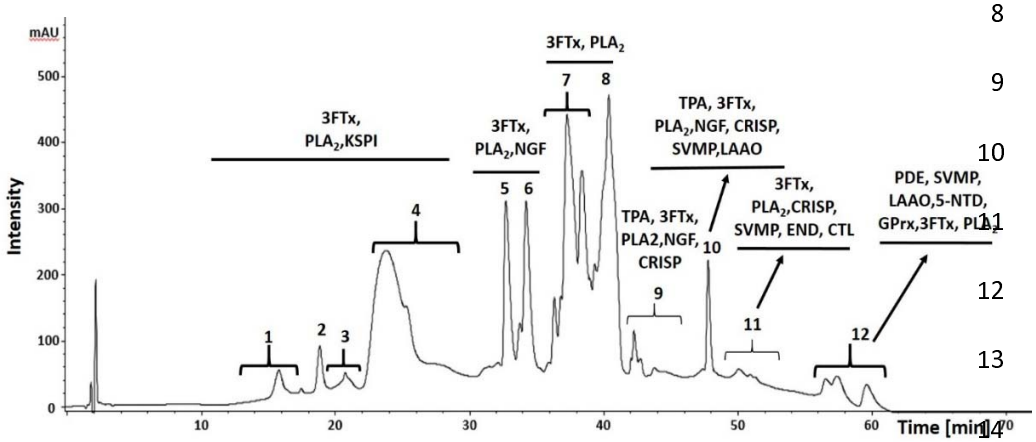


Figure 2b. 12% SDS-PAGE of RP-HPLC separated fractions of NNwg venom

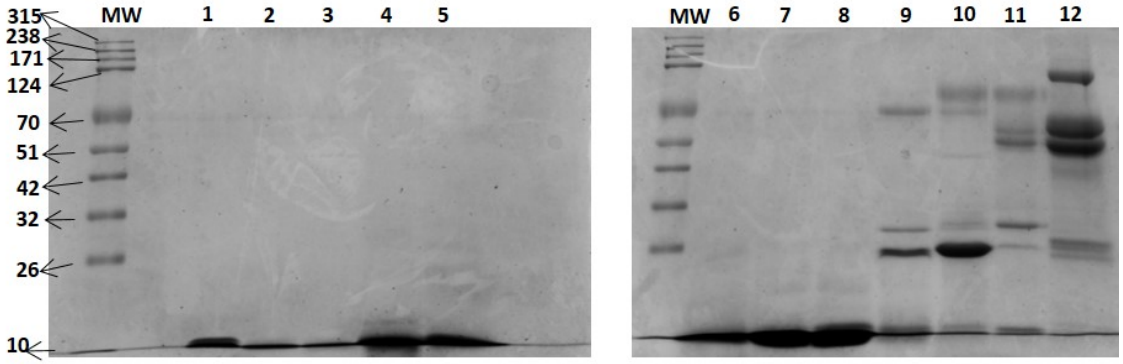


Figure 3. Relative abundance of protein families identified from NNwg venom

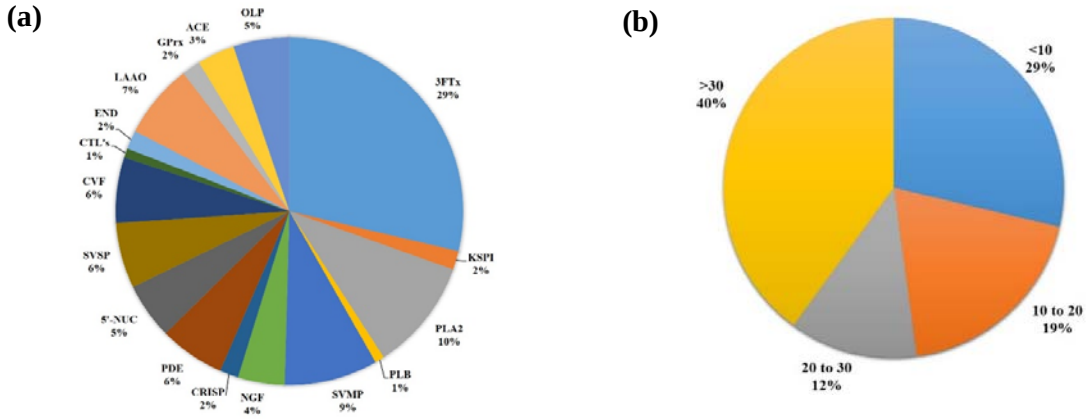


Figure 4. *De novo* analysis of a peptide representing NNwg venom protein

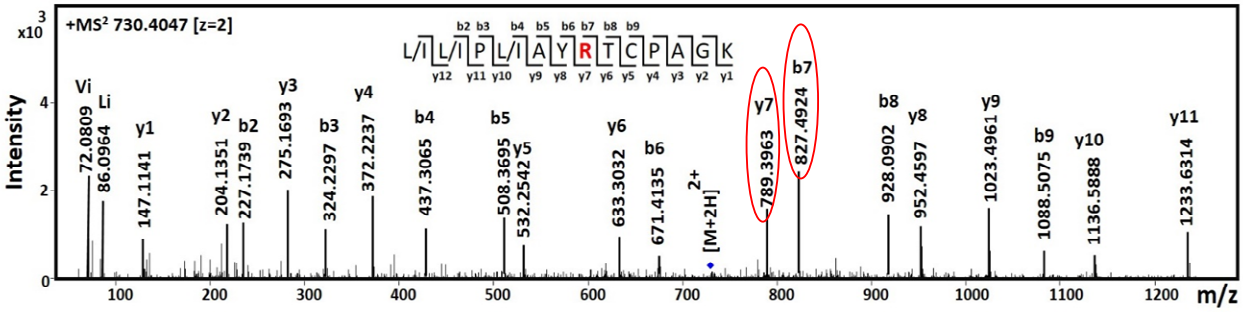


Figure 5. BLASTp and ClustalW analysis of endonuclease and glutathione peroxidase proteins identified from NNwg venom

(a)

```
Notechis scutatus      1 MLLYFILLSLSWSQMPTSGEVVDSFQDHCPQFFLRETPPVIGVPLSPVRICQRYQNLYRF
Pseudonaja textilis    1 MLLCFILLSLSWSQTPTSGEVVDSFQDHCPQFFLRETPPVIGVPLSPVRICQRYQNLYRF
Thamnophis elegans     1 MRLLFILLSVSWSQMPTSGEVVDSFQDTCPQFFLRETPPLIRIPPSPVRICQRYQNLYRF
Ophiophagus hannah     1 ------------------------------------------------------------
NNwg                   1 ------------------------------------------------------------

Notechis scutatus     61 ATLYDKHNRIPVYSAYVYNPGKAKRPGKWMVEPQLISSTFQSEMETERDFLNKKGPITAM
Pseudonaja textilis   61 ATLYDKHNRIPVYSAYVYNPGKAKRPGKWMVEPQLISSTFQSEMETERDFLNKKGPKAAM
Thamnophis elegans    61 ATLYDKDNRIPVYSAYVYNPGSAKRPKKWMVEPQLINSSFQPEMETEWDFLNKKGPKEAL
Ophiophagus hannah     1 -------------------------------------------METEWDFLNKKGSPKAL
NNwg                   1 ------------------------------------------------------------

Notechis scutatus    121 QNSQAILEDYKNLTDCNRGHLNPNGHQPDYSAKSSTFTLTNIVPQFIKLNGGAWNNYEQT
Pseudonaja textilis  121 QDSQAVLEDYKNLTDCNRGHLNPNGHQPDYSAKSSTFTLTNIVPQFIKLNGGAWNNYEQT
Thamnophis elegans   121 HVSQAVLEDYKNLSDYNRGHLNPNGHQPDYSAKSSTFTLTNIVPQFIKLNGGAWNNYEQT
Ophiophagus hannah    18 QDSQAVLEDYKNLADCNRGHLNPNGHQPDYSAKSSTFTLTNIVPQFIKLNGGAWNNYEQT
NNwg                   1 ------------------GHLNPNGHQPDYSAKSSTFTLTNIVPQFIKG-----------

Notechis scutatus    181 TMQQMTRGCHQTFAVVGAVPGDTYIARGRVNRPSHLWSAACCVIDNNHLRSWAILARNDQ
Pseudonaja textilis  181 TMQQMTRGCHQTFAVVGAVPGDTYISGGRVNRPSHLWSAACCVIDNNHLRSWAILARNDQ
Thamnophis elegans   181 TMVQMTKGCQETFAVVGAVPGDTYVAGGRVNKPSHMWSAACCMIDNNHLRSWAILARNDQ
Ophiophagus hannah    78 TMQQMTRGCQQTFAVVGAVPGDTYIARGRVNRPSHLWSAACCLIDNNHLRSWAILARNX-
NNwg                  32 --------CQQTFAVVGAVPGDTYIAR----------SAACCLIDNNHLRSWLAQLYN--

Notechis scutatus    241 NVVQRFTLGQLEEKLAQLYNVNHVSLFHSDCPRQ
Pseudonaja textilis  241 NVVQRFTLGQLEEKLAQLYNVNHVSLFHSDCPRQ
Thamnophis elegans   241 NVVEKLTLGQLEAHLAQLYNINQVSLFHSDCPRQ
Ophiophagus hannah   137 ----------LEEKLAQLYNVNHVSLFHSDCPRQ
NNwg                  72 --------------------VNHVSLFHSDCPRQ
```

(b)

```
Pseudonaja textilis    1 ----------------------------------------------MKGRLKGAWILSLL
Notechis scutatus      1 ----------------------------------------------MKGRLKGAWILSLL
P.mucrosquamatus       1 ----------------------------------------------MKGRLKGAWILSLL
Ophiophagus hannah     1 VELEGDNVLTQRQRERNTKEHLEQTQISVEARPGRGAEALSLLLSNHEGQTERSLDPFLA
NNwg                   1 ------------------------------------------------------------

Pseudonaja textilis   15 LAGLFQPSLGQEQQ----------------------------EVDCYDSVKGSAHDYGAF
Notechis scutatus     15 LAGLFQPSLGQEQQ----------------------------EVDCYDSVKGSAHDYGAL
P.mucrosquamatus      15 LAGLFQPNLGQEQQEILTGLQLSHQHDEARSCLHAQPIPDRAKVDCYDSVQGSAHDYGAL
Ophiophagus hannah    61 LGWFISAQPGTGTTGGLQLSHQHEEASSCSYAPNSQPIPDRAKVDCYDSVKGSAHDYGAL
NNwg                   1 ------------------------------------------AKVDCYDSVK---------

Pseudonaja textilis   47 SLSGEYIAFQKYVGKYILFVNVATYUGLTAQYLELNALQNSLRTDRLVILGFPCNQFGKQ
Notechis scutatus     47 SLTGEYIAFQNYVGKYILFVNVATYUGLTQQYLELNALQNSLRTDRLVILGFPCNQFGKQ
P.mucrosquamatus      75 SLSGEYIPFRNYAGKYILFVNVATYUGLTGQYLELNALQTSLGIDRLVILGFPCNQFGKQ
Ophiophagus hannah   121 SLT------------------------------ELNALQNSLKTDRLVILGFPCNQFGKQ
NNwg                  11 ------------------------------------------TDRLVILGFPCNQFGKQ

Pseudonaja textilis  107 EPGQNSEILQGIKYVRPGGGYVPNFQLFQKGDVNGENEQQIYTFLKNSCPPVVETFGDTT
Notechis scutatus    107 EPGQNSEILQGIKYVRPGGGYVPNFQLFQKGDVNGENEQQIYTFLKNSCPPVVETFGDTT
P.mucrosquamatus     135 EPGQNSEILQGIKHVRPGGGFVPNFQLFQKGDVNGENEQRIYTFLKNSCPPVVENFGDPA
Ophiophagus hannah   151 EPGQNSEILQGIKHVRPGGGFVPNFQLFQKGDVNGENEQKIYTFLKNSCPPVVETFGDTA
NNwg                  28 EPGQNSEILQGIK---------------QKGDVNGENEQKIYTFLKNSCPPVVETFGDTA

Pseudonaja textilis  167 RLFWSPLKIHDIKWNFEKFLVNPQGKPVMRWFHRTNVSTVKNDIIRYMRKNRNG
Notechis scutatus    167 RLFWSPLKIHDIKWNFEKFLVNPQGKPVMRWFHRTNVSTVKNDIIRYMRKNRNG
P.mucrosquamatus     195 KLFWSPLKTHDIKWNFEKFLVNPQGKPVMRWFQRTNVSSVKNDIIRYMRKNG--
Ophiophagus hannah   211 KLFWSPLKIHDIKWNFEKFLVNPQGKPVMRWFHRTNVSTVKNDIIRFMRKKRSG
NNwg                  73 KLFWSPLKIHDIKWNFEFLVN-PQGKPVMRWTN---VSTVKNDIIR--------
```

Figure 6. Assessment of the immunological cross-reactivity of different polyvalent antivenoms towards NNwg venom

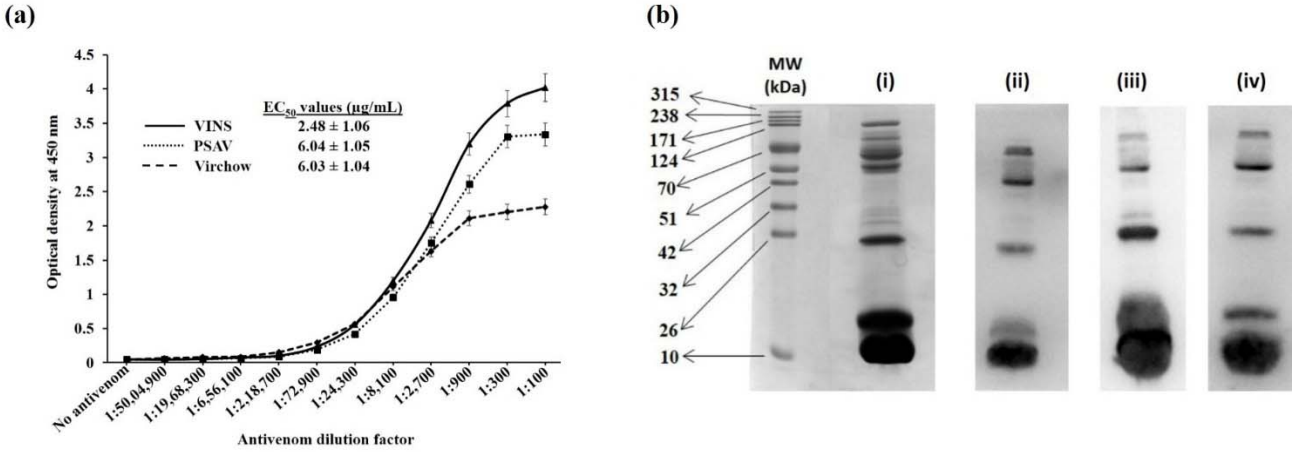
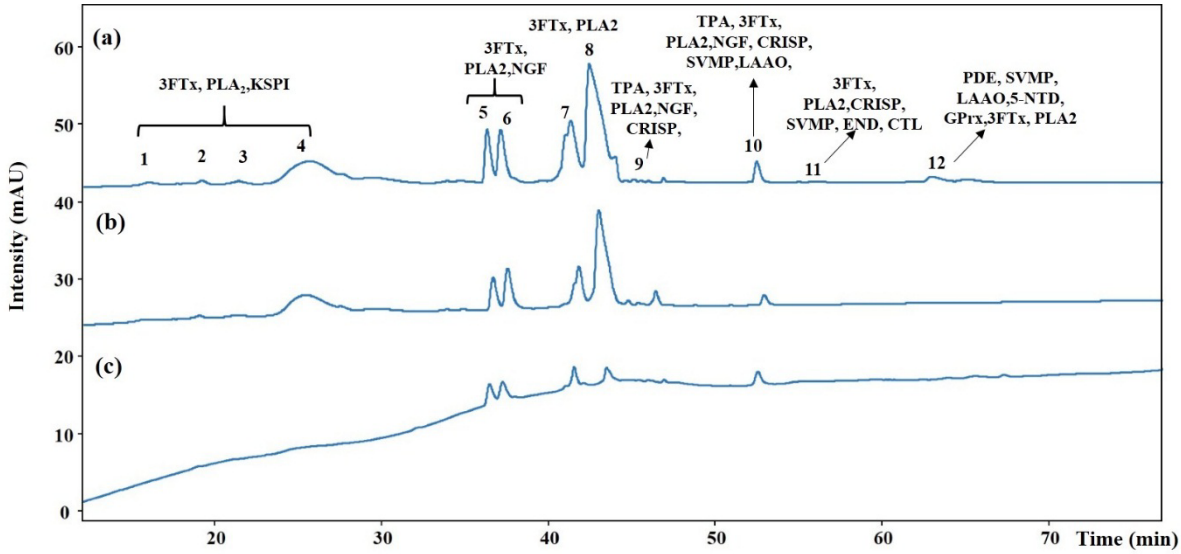


Figure 7. Immunoaffinity chromatography of NNwg using VINS antivenom



Tables

Table 1: Peptides representing various snake venom proteins identified through *de novo* analysis. *De novo* sequencing was performed using PEAKS and Novor softwares followed by

MS-BLAST. The mutated amino acids observed are highlighted in bold and underlined. Proteins marked in bold were identified through both conventional proteomics and *de novo* methods.

Table 2: Comparison of protein families identified in cobra venom collected from different geographical location of India.

Table 3: Summary of cobra venom proteomic studies from India.

Supplementary Table 1: Proteins identified from crude NNwg venom after SDS-PAGE and HPLC fractionation. The proteins separated through SDS-PAGE and HPLC were subjected to in-gel and in-solution digestion using trypsin, chymotrypsin and V8 protease. The digests were then subjected to Q-TOF LC-MS/MS data collection and the obtained proteins were assigned to different venom protein families through Mascot and PEAKS analysis. #; peptides identified through chymotrypsin digestion, ^; peptides identified through V8 protease digestion, P#; peptides derived through PEAKS analysis after chymotrypsin digestion, P^; peptides derived through PEAKS analysis after V8 protease digestion. All the other peptides were obtained through trypsin digestion.

Highlights

- Proteomics of cobra venom resulted in the identification of 115 proteins representing 17 snake venom protein families.
- *De novo* approaches exclusively yielded 39 peptides harbouring multiple amino acid mutations.
- Glutathione peroxidase and endonuclease were identified for the first time in Indian cobra venom.

- Indian polyvalent antivenoms showed varying cross-reactivity towards cobra venom.
- VINS antivenom was less efficient to detect many low molecular mass proteins (< 20 kDa).

References

[1] W. Suraweera, D. Warrell, R. Whitaker, G. Menon, R. Rodrigues, S.H. Fu, R. Begum, P. Sati, K. Piyasena, M. Bhatia, P. Brown, P. Jha: Trends in snakebite deaths in India from 2000 to 2019 in a nationally representative mortality study. *Elife* 2020, 9.

[2] World Health Organization. Snakebite envenoming.2019,https://www.who.int/news-room/fact-sheets/detail/snakebite-envenoming

[3] J.P. Chippaux: Snakebite envenomation turns again into a neglected tropical disease! *J Venom Anim Toxins Incl Trop Dis* 2017, 23:38.

[4] I.D. Simpson, R.L. Norris: Snakes of medical importance in India: is the concept of the "Big 4" still relevant and useful? *Wilderness Environ Med* 2007, 18:2-9.

[5] M. Choudhury, R.J.R. McCleary, M. Kesherwani, R.M. Kini, D. Velmurugan: Comparison of proteomic profiles of the venoms of two of the 'Big Four' snakes of India, the Indian cobra (Naja naja) and the common krait (Bungarus caeruleus), and analyses of their toxins. *Toxicon* 2017, 135:33-42.

[6] R. Whitaker, G. Martin, Diversity and Distribution of Medically Important Snakes of India, in: F.A. In: Gopalakrishnakone P., Fernando R., Gnanathasan C., Habib A., Yang CC. (eds) (Ed.) Clinical Toxinology in Asia Pacific and Africa, Springer, Dordrecht., 2015.

[7] T. Tasoulis, G.K. Isbister: A Review and Database of Snake Venom Proteomes. *Toxins (Basel)* 2017, 9.

[8] A. Chanda, B. Kalita, A. Patra, W. Senevirathne, A.K. Mukherjee: Proteomic analysis and antivenomics study of Western India Naja naja venom: correlation between venom composition and clinical manifestations of cobra bite in this region. *Expert Rev Proteomics* 2019, 16:171-184.

[9] S. Dutta, A. Chanda, B. Kalita, T. Islam, A. Patra, A.K. Mukherjee: Proteomic analysis to unravel the complex venom proteome of eastern India Naja naja: Correlation of venom composition with its biochemical and pharmacological properties. *J Proteomics* 2017, 156:29-39.

[10] A. Chanda, A. Patra, B. Kalita, A.K. Mukherjee: Proteomics analysis to compare the venom composition between Naja naja and Naja kaouthia from the same geographical location of eastern India: Correlation with pathophysiology of envenomation and immunological cross-reactivity towards commercial polyantivenom. *Expert Rev Proteomics* 2018, 15:949-961.

[11] K. Sintiprungrat, K. Watcharatanyatip, W.D. Senevirathne, P. Chaisuriya, D. Chokchaichamnankit, C. Srisomsap, K. Ratanabanangkoon: A comparative study of venomics of Naja naja from India and Sri Lanka, clinical manifestations and antivenomics of an Indian polyspecific antivenom. *J Proteomics* 2016, 132:131-143.

[12] K. Suryamohan, S.P. Krishnankutty, J. Guillory, M. Jevit, M.S. Schroder, M. Wu, B. Kuriakose, O.K. Mathew, R.C. Perumal, I. Koludarov, L.D. Goldstein, K. Senger, M.D. Dixon, D. Velayutham, D. Vargas, S. Chaudhuri, M. Muraleedharan, R. Goel, Y.J. Chen, A. Ratan, P. Liu, B. Faherty, G. de la Rosa, H. Shibata, M. Baca, M. Sagolla, J. Ziai, G.A. Wright, D. Vucic, S. Mohan, A. Antony, J. Stinson, D.S. Kirkpatrick, R.N. Hannoush, S. Durinck, Z. Modrusan, E.W. Stawiski, K. Wiley, T. Raudsepp, R.M. Kini, A. Zachariah, S. Seshagiri: The Indian cobra reference genome and transcriptome enables comprehensive identification of venom toxins. *Nat Genet* 2020, 52:106-117.

[13] A. Chanda, A.K. Mukherjee: Quantitative proteomics to reveal the composition of Southern India spectacled cobra (Naja naja) venom and its immunological cross-reactivity towards commercial antivenom. *Int J Biol Macromol* 2020, 160:224-232.

[14] R.R. Senji Laxme, S. Khochare, H.F. de Souza, B. Ahuja, V. Suranse, G. Martin, R. Whitaker, K. Sunagar: Beyond the 'big four': Venom profiling of the medically important yet neglected Indian snakes reveals disturbing antivenom deficiencies. *PLoS Negl Trop Dis* 2019, 13:e0007899.

[15] H.L. Gibbs, S.P. Mackessy: Functional basis of a molecular adaptation: prey-specific toxic effects of venom from Sistrurus rattlesnakes. *Toxicon* 2009, 53:672-679.
[16] S.P. Mackessy, J. Leroy, E. Mocino-Deloya, K. Setser, R.W. Bryson, A.J. Saviola: Venom Ontogeny in the Mexican Lance-Headed Rattlesnake (Crotalus polystictus). *Toxins (Basel)* 2018, 10.
[17] A.D. Catherman, O.S. Skinner, N.L. Kelleher: Top Down proteomics: facts and perspectives. *Biochem Biophys Res Commun* 2014, 445:683-693.
[18] R.D. Melani, F.C.S. Nogueira, G.B. Domont: It is time for top-down venomics. *J Venom Anim Toxins Incl Trop Dis* 2017, 23:44.
[19] M. Vanuopadath, S.K. Shaji, D. Raveendran, B.G. Nair, S.S. Nair: Delineating the venom toxin arsenal of Malabar pit viper (Trimeresurus malabaricus) from the Western Ghats of India and evaluating its immunological cross-reactivity and in vitro cytotoxicity. *Int J Biol Macromol* 2020, 148:1029-1045.
[20] M. Vanuopadath, N. Sajeev, A.R. Murali, N. Sudish, N. Kangosseri, I.R. Sebastian, N.D. Jain, A. Pal, D. Raveendran, B.G. Nair, S.S. Nair: Mass spectrometry-assisted venom profiling of Hypnale hypnale found in the Western Ghats of India incorporating de novo sequencing approaches. *Int J Biol Macromol* 2018, 118:1736-1746.
[21] K. Shabbiri, C.H. Botting, A. Adnan, M. Fuszard: Charting the cellular and extracellular proteome analysis of Brevibacterium linens DSM 20158 with unsequenced genome by mass spectrometry-driven sequence similarity searches. *J Proteomics* 2013, 83:99-118.
[22] D.Z. Wang, C. Li, Z.X. Xie, H.P. Dong, L. Lin, H.S. Hong: Homology-Driven Proteomics of Dinoflagellates with Unsequenced Genomes Using MALDI-TOF/TOF and Automated De Novo Sequencing. *Evid Based Complement Alternat Med* 2011, 2011:471020.
[23] P. Waridel, A. Frank, H. Thomas, V. Surendranath, S. Sunyaev, P. Pevzner, A. Shevchenko: Sequence similarity-driven proteomics in organisms with unknown genomes by LC-MS/MS and automated de novo sequencing. *Proteomics* 2007, 7:2318-2329.
[24] D. Pla, J.M. Gutierrez, J.J. Calvete: Second generation snake antivenomics: comparing immunoaffinity and immunodepletion protocols. *Toxicon* 2012, 60:688-699.
[25] A. Silva, W.C. Hodgson, G.K. Isbister: Antivenom for Neuromuscular Paralysis Resulting From Snake Envenoming. *Toxins (Basel)* 2017, 9.
[26] M. Vanuopadath, D. Raveendran, B.G. Nair, S.S. Nair:Evaluating the Immunological cross-reactivity of Indian polyvalent antivenoms towards the venom of Hypnale hypnale (hump-nosed pit viper) from the Western Ghats bioRxiv. https://doi.org/10.1101/2020.08.01.232579
[27] P.K. Smith, R.I. Krohn, G.T. Hermanson, A.K. Mallia, F.H. Gartner, M.D. Provenzano, E.K. Fujimoto, N.M. Goeke, B.J. Olson, D.C. Klenk: Measurement of protein using bicinchoninic acid. *Anal Biochem* 1985, 150:76-85.
[28] A. Keller, A.I. Nesvizhskii, E. Kolker, R. Aebersold: Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 2002, 74:5383-5392.
[29] A.I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold: A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 2003, 75:4646-4658.
[30] T. Muth, L. Weilnbock, E. Rapp, C.G. Huber, L. Martens, M. Vaudel, H. Barsnes: DeNovoGUI: an open source graphical user interface for de novo sequencing of tandem mass spectra. *J Proteome Res* 2014, 13:1143-1146.
[31] A. Shevchenko, S. Sunyaev, A. Loboda, P. Bork, W. Ens, K.G. Standing: Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. *Anal Chem* 2001, 73:1917-1926.
[32] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman: Basic local alignment search tool. *J Mol Biol* 1990, 215:403-410.

[33] J.A. Vizcaino, E.W. Deutsch, R. Wang, A. Csordas, F. Reisinger, D. Rios, J.A. Dianes, Z. Sun, T. Farrah, N. Bandeira, P.A. Binz, I. Xenarios, M. Eisenacher, G. Mayer, L. Gatto, A. Campos, R.J. Chalkley, H.J. Kraus, J.P. Albar, S. Martinez-Bartolome, R. Apweiler, G.S. Omenn, L. Martens, A.R. Jones, H. Hermjakob: ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* 2014, 32:223-226.
[34] J.A. Vizcaino, A. Csordas, N. del-Toro, J.A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q.W. Xu, R. Wang, H. Hermjakob: 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 2016, 44:D447-456.
[35] R.M. Kini: Molecular moulds with multiple missions: functional sites in three-finger toxins. *Clin Exp Pharmacol Physiol* 2002, 29:815-822.
[36] Y.N. Utkin: Last decade update for three-finger toxins: Newly emerging structures and biological activities. *World J Biol Chem* 2019, 10:17-27.
[37] V. Tsetlin: Snake venom alpha-neurotoxins and other 'three-finger' proteins. *Eur J Biochem* 1999, 264:281-286.
[38] L.T. Potter: Snake toxins that bind specifically to individual subtypes of muscarinic receptors. *Life Sci* 2001, 68:2541-2547.
[39] J.M. Gutierrez, B. Lomonte: Phospholipases A2: unveiling the secrets of a functionally versatile group of snake venom toxins. *Toxicon* 2013, 62:27-39.
[40] M.K. Bhat, T.V. Gowda: Purification and characterization of a myotoxic phospholipase A2 from Indian cobra (Naja naja naja) venom. *Toxicon* 1989, 27:861-873.
[41] M.K. Bhat, B.N. Prasad, T.V. Gowda: Purification and characterization of a neurotoxic phospholipase A2 from Indian cobra (Naja naja naja) venom. *Toxicon* 1991, 29:1345-1349.
[42] L.M. Rudrammaji, T.V. Gowda: Purification and characterization of three acidic, cytotoxic phospholipases A2 from Indian cobra (Naja naja naja) venom. *Toxicon* 1998, 36:921-932.
[43] F.S. Markland, Jr.: Inventory of alpha- and beta-fibrinogenases from snake venoms. For the Subcommittee on Nomenclature of Exogenous Hemostatic Factors of the Scientific and Standardization Committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost* 1991, 65:438-443.
[44] J.W. Fox, S.M. Serrano: Structural considerations of the snake venom metalloproteinases, key members of the M12 reprolysin family of metalloproteinases. *Toxicon* 2005, 45:969-985.
[45] J. Hamako, T. Matsui, S. Nishida, S. Nomura, Y. Fujimura, M. Ito, Y. Ozeki, K. Titani: Purification and characterization of kaouthiagin, a von Willebrand factor-binding and -cleaving metalloproteinase from Naha kaouthia cobra venom. *Thromb Haemost* 1998, 80:499-505.
[46] M.S. Kumar, K.S. Girish, B.S. Vishwanath, K. Kemparaju: The metalloprotease, NN-PF3 from Naja naja venom inhibits platelet aggregation primarily by affecting alpha2beta1 integrin. *Ann Hematol* 2011, 90:569-577.
[47] A.S. Kamiguti: Platelets as targets of snake venom metalloproteinases. *Toxicon* 2005, 45:1041-1049.
[48] W.J. Wang, C.H. Shih, T.F. Huang: Primary structure and antiplatelet mechanism of a snake venom metalloproteinase, acurhagin, from Agkistrodon acutus venom. *Biochimie* 2005, 87:1065-1077.
[49] W.J. Wang: Purification and functional characterization of AAV1, a novel P-III metalloproteinase, from Formosan Agkistrodon acutus venom. *Biochimie* 2007, 89:105-115.
[50] F.S. Markland: Snake venoms and the hemostatic system. *Toxicon* 1998, 36:1749-1800.
[51] E.F. Sanchez, R.J. Flores-Ortiz, V.G. Alvarenga, J.A. Eble: Direct Fibrinolytic Snake Venom Metalloproteinases Affecting Hemostasis: Structural, Biochemical Features and Therapeutic Potential. *Toxins (Basel)* 2017, 9.
[52] K.K. Tan, B.H. Bay, P. Gopalakrishnakone: L-amino acid oxidase from snake venom and its anticancer potential. *Toxicon* 2018, 144:7-13.

[53] M. Pessatti, J.D. Fontana, M.F. Furtado, M.F. Guimaraes, L.R. Zanette, W.T. Costa, M. Baron: Screening of Bothrops snake venoms for L-amino acid oxidase activity. *Appl Biochem Biotechnol* 1995, 51-52:197-210.
[54] W.M. Chou, W.H. Liu, K.C. Chen, L.S. Chang: Structure-function studies on inhibitory activity of Bungarus multicinctus protease inhibitor-like protein on matrix metalloprotease-2, and invasion and migration of human neuroblastoma SK-N-SH cells. *Toxicon* 2010, 55:353-360.
[55] Y. Yamazaki, T. Morita: Snake venom components affecting blood coagulation and the vascular system: structural similarities and marked diversity. *Curr Pharm Des* 2007, 13:2872-2886.
[56] P. Lazarovici, C. Marcinkiewicz, P.I. Lelkes: From Snake Venom's Disintegrins and C-Type Lectins to Anti-Platelet Drugs. *Toxins (Basel)* 2019, 11.
[57] S.D. Aird: Taxonomic distribution and quantitative analysis of free purine and pyrimidine nucleosides in snake venoms. *Comp Biochem Physiol B Biochem Mol Biol* 2005, 140:109-126.
[58] S.D. Aird: Ophidian envenomation strategies and the role of purines. *Toxicon* 2002, 40:335-393.
[59] Y.F. Pung, P.T. Wong, P.P. Kumar, W.C. Hodgson, R.M. Kini: Ohanin, a novel protein from king cobra venom, induces hypolocomotion and hyperalgesia in mice. *J Biol Chem* 2005, 280:13137-13147.
[60] A. Chapeaurouge, M.A. Reza, S.P. Mackessy, P.C. Carvalho, R.H. Valente, A. Teixeira-Ferreira, J. Perales, Q. Lin, R.M. Kini: Interrogating the Venom of the Viperid Snake Sistrurus catenatus edwardsii by a Combined Approach of Electrospray and MALDI Mass Spectrometry. *PLoS One* 2015, 10:e0092091.
[61] S.D. Robinson, E.A.B. Undheim, B. Ueberheide, G.F. King: Venom peptides as therapeutics: advances, challenges and the future of venom-peptide discovery. *Expert Rev Proteomics* 2017, 14:931-939.
[62] G. de la Rosa, F. Olvera, I.G. Archundia, B. Lomonte, A. Alagon, G. Corzo: Horse immunization with short-chain consensus alpha-neurotoxin generates antibodies against broad spectrum of elapid venomous species. *Nat Commun* 2019, 10:3642.
[63] R. Teixeira-Araujo, P. Castanheira, L. Brazil-Mas, F. Pontes, M. Leitao de Araujo, M.L. Machado Alves, R.B. Zingali, C. Correa-Netto: Antivenomics as a tool to improve the neutralizing capacity of the crotalic antivenom: a study with crotamine. *J Venom Anim Toxins Incl Trop Dis* 2017, 23:28.
[64] World Health Organization. WHO Guidelines for the Production, Control and Regulation of Snake Antivenom
Immunoglobulins.2018,https://www.who.int/bloodproducts/AntivenomGLrevWHO_TRS_1004_web_Annex_5.pdf?ua=1

Table 1: Peptides representing various snake venom proteins identified through *de novo* analysis. *De novo* sequencing was performed using PEAKS and Novor softwares followed by MS-BLAST. The mutated amino acids observed are highlighted in bold and underlined. Proteins marked in bold were identified through both conventional proteomics and *de novo* methods.

Sl. No.	**Precursor ion mass (Da)**	**Observed m/z (Z)**	**ppm**	***De novo* sequence**	**Mutation observed**	**Search score**	**Query coverage (%)**	**% identity**	**E value**	**Protein name**	**Accession number**	**Snake species**	**Family**	**Methodology**
1	2141.924	1071.9618 (2)	2	V**R**TGVDIQCCSTDNCNPF	K49R	60	100	94.44	4.00E-12	Alpha-cobratoxin (LNTx)	1CTX_A	*Naja siamensis*	3FTX	RP-HPLC
2	1390.678	696.3388 (2)	3	**D**VLGCAATCPTVK	V58D	41	92	100	2.00E-05	Alpha-elapitoxin-Al2b (LNTx)	B2BRQ5	*Austrelaps labialis*	3FTX	
3	1449.786	484.2621(3)	1	LRCFLTP**N**LTSK	D8N	41	100	91.67	1.00E-05	Long neurotoxin 1	P25668	*Naja naja*	3FTX	
4	1357.564	679.7821(2)	6	**G**TWCDGFCSIR	K23G	39	90	100	3.00E-05	Long neurotoxin 1	P25668	*Naja naja*	3FTX	SDS-PAGE
5	1158.718	387.2393(3)	6	FPL**R**FPVKR	K54R	30	100	88.89	5.70E-02	Cardiotoxin V	1CVO_A	*Naja atra*	3FTX	RP-HPLC
6	1577.724	789.8619 (2)	10	NLCY**R**MYM**WNK**	K23R V27W A28N T29K	32	72	87.5	0.013	Cytotoxin 3a	P86539	*Naja naja*	3FTX	

7	1458.809	730.4047 (2)	6	LIPLAY**R**TCPAGK	K12R	41	100	92.31	4.00E-06					
8	1459.567	730.8283 (2)	5	**GL**CYQD**L**GDTYR	A103G M104L G109L	31	91	81.82	2.70E-02	Tissue-type plasminogen activator, partial	XP_026544671	*Notechis scutatus*	SVSP	RP-HPLC
9	1299.609	650.8045 (2)	2	L**F**PDSFCTSAR	Y449F	36	100	90.91	5.00E-04					
10	1107.224	554.6120 (2)	3	VA**DL**FQATYTYQNK	S422D V423L	40	100	85.71	3.00E-05	complement receptor type 1-like	XP_034271149	*Pantherophis guttatus*	CVF	
11	1140.577	571.2884 (2)	2	CP**L**PNLENGK	V96L	88	100	90	2.00E-03	C4b-binding protein alpha chain-like isoform X1	XP_026571379	*Pseudonaja textilis*	CVF	
12	1633.781	545.5937 (3)	2	LGD**K**VTFGCD**Y**GFR	Q115K F122Y	41	100	85.71	1.00E-05	C4b-binding protein alpha chain-like isoform X2	XP_007427549	*Python bivittatus*	CVF	
13	1862.129	932.0644(2)	8	LILNTPLN**A**QSLPLTVR	T409A	83	100	94.12	1.00E-09	**Venom factor**	**AAX86641**	***Austrelaps superbus***	CVF	SDS-PAGE
14	1838.799	920.3994(2)	10	DDNED**NM**IAD**MA**IISR	G738N F739M S743M D744A	36	100	75	1.00E-03	Chain B, Cobra Venom Factor	3PRX_B	*Naja kaouthia*	CVF	

15	1051.533	351.5109 (3)	2	SFAEWR**E**K	A281E	27	100	87.5	5.50E-01	**Metalloproteinase atrase A**	**ADF43026**	***Naja atra***	SVMP	RP-HPLC
16	1639.788	547.5960 (3)	3	TGCIV**T**VSP**H**DPDSR	P576T R580H	42	100	86.67	8.00E-06					
17	1179.604	590.8019 (2)	1	**DK**AVVQDYSR	S318D V319K	29	80	100	0.15					
18	1072.474	537.2370 (2)	1	**D**HQEYLLR	E381D	28	100	87.5	0.14					
19	1491.79	498.2632 (3)	3	VTLDLFGEW**K**ER	R43K	38	100	83.33	1.00E-04	Metalloproteinase, partial	ADI47661	*Echis coloratus*	SVMP	
20	1234.655	618.3275 (2)	2	LAY**L**GS**L**CNPK	V117L, M120L	32	100	81.82	0.016	Zinc metalloproteinase-disintegrin-like batroxstatin-1	C5H5D2	*Bothrops atrox*	SVMP	
21	1688.827	845.4135 (2)	7	VYEMLN**M**LR**D**SYR	T226M E229D	34	100	69.23	0.003	Acutusin 2	ADB43124	*Deinagkistrodon acutus*	SVMP	
22	1179.613	394.2043 (3)	2	T**A**AAVVQDYS**R**	S130A S139R	31	90	90	0.04	**Chain A, Atragin**	**3K7L_A**	***Naja atra***	SVMP	
23	1929.988	644.3293 (3)	2	CPIMTNQCIALRGP**RK**	G524R V525K	51	87	100	5.00E-09	Zinc metalloproteinase-disintegrin-like atragin	D3TTC2	*Naja atra*	SVMP	

24	1453.628	727.8141(2)	8	NS**F**ICNC**AE**SPR	M573F, S578A, I579E	30	100	75	0.078	Zinc metalloproteinase-disintegrin-like atragin	D3TTC2	*Naja atra*	SVMP	SDS-PAGE
25	1561.691	521.5637 (3)	1	LYCTEKNTMSC**R**	L367R	42	91	100	5.00E-06	**Hemorrhagic metalloproteinase-disintegrin-like kaouthiagin**	**P82942**	***Naja kaouthia***	SVMP	RP-HPLC
26	1645.836	549.6121(3)	1	AWVYE**F**VN**Y**LN**TK**	I229F T232Y E235T I236K	32	84	81.82	0.023	MAD-4b	QIV64942	*Crotalus atrox*	SVMP	
27	1989.047	664.0157 (3)	10	NDE**L**VVQ**Q**FTLGQLEEK	N241L R245Q	44	100	82.35	0.000001	Endonuclease domain-containing 1 protein-like	XP_026579205	*Pseudonaja textilis*	END	
28	1304.641	653.3204 (2)	3	Y**N**CPIDWLPK	T25N	36	100	90	0.0004	**C-type lectin BfL-1**	**Q90WI8**	***Bungarus fasciatus***	CTL	
29	1250.614	626.3071 (2)	3	NNEFCVELV**K**	N134K	34	90	100	0.002					
30	1215.593	406.1977 (3)	3	FWEADGLH**N**K	G362N G363K	32	80	100	0.016	**L-amino acid oxidase**	**A8QL51**	***Bungarus multicinctus***	LAAO	

31	1122.64	375.2134 (3)	3	IAE**T**VHLNAR	M273T	30	100	90	0.065	**L-amino acid oxidase**	**A8QL58**	***Naja atra***		
32	1141.591	571.7955 (2)	1	SITG**N**YIAFR	E78N	31	100	90	0.018	GPX3	AFS63888	*Thamnophis elegans*	Gprx	
33	1354.682	678.3412(2)	9	AATYFWPG**T**D**V**K	S26T I28V	38	100	83.3	1.00E-04	phosphodiesterase family member 3, partial	XP_007430581	*Python bivittatus*	PDE	SDS-PAGE
34	1853.827	927.9134 (2)	4	CNAYI**TA**FYYNP**D**S**R**	P43T S44A H50D H52R	39	93	78.57	7.00E-05	Kunitz inhibitor II	ABX82869	*Walterinnesia aegyptia*	KSPI	RP-HPLC
35	1187.558	594.7790 (2)	1	**N**SGTPVDDLDR	G25N	34	90	100	0.003	**Phospholipase A_2, partial**	**CAM34525**	***Naja atra***	PLA2	
36	1494.706	748.3528 (2)	8	FN**G**SPCVLGSPGFR	S101G	44	100	92.86	2.00E-06	**Vespryn**	**F8S122**	***Crotalus adamanteus***	OLP	
37	1048.544	525.272(2)	7	LTNCDSL**V**K	L196V	28	100	88.89	0.24	**Natrin**	**1XTA_A**	***Naja Atra***	CRISP	SDS-PAGE
38	1125.474	563.7368 (2)	9	N**G**DFNSQSTR	V20G	27	100	80	0.6	Natrin	AAP85301	*Naja atra*	CRISP	RP-HPLC
39	1292.563	647.2816 (2)	2	Q**N**SCQDDWLK	S217N	35	100	90	6.00E-04					

Table 3: Comparison of protein families identified in cobra venom collected from different geographical location of India.

Sl. No.	Venom family	*Naja naja*							
		West India [8]	**West India [14]**	**East India [9,10]**	**North India [11]**	**South India [5]**	**South India [13]**	**Kentucky Reptile Zoo [12]**	**Western Ghats #**
1	**PLA_2**	+	+	+	+	+	+	+	+
2	**PLB**	+	-	+	-	-	-	+	+
3	**SVMP**	+	+	+	+	+	+	+	+
4	**SVSP**	+	+	-	-	+	+	+	+
5	**CRISP**	+	+	+	+	+	+	+	+
6	**Cystatin**	+	-	+	-	-	+	+	-
7	**KSPI**	+	+	+	+	+	+	+	+
8	**3FTx**	+	+	+	+	+	+	+	+
9	**NGF**	+	-	+	+	+	+	+	+
10	**ACE**	+	-	+	-	-	+	+	+
11	**PDE**	+	-	+	-	+	-	-	+
12	**LAAO**	+	+	+	-	+	+	+	+
13	**CVF**	+	+	+	-	+	+	+	+
14	**5-NU**	+	+	+	-	+	+	+	+
15	**GPrx**	-	-	-	-	-	-	-	+
16	**Natriuretic peptide**	-	+	+	-		+	-	-
17	**AP**	+	-	-	-	+	-	-	-
18	**OLP**	+	+	+	-	+	+	+	+
19	**CTL**	-	-	-	-	-	-	+	+
20	**END**	-	-	-	-	-	-	-	+
21	**Hyal**	-	-	-	-	-	-	+	-

22	**VEGF**	**-**	**-**	**-**	**-**	**-**	**-**	**+**	**-**
	***- cobra venom obtained from Kentucky Reptile Zoo, #-current study**								

Table 4: Summary of cobra venom proteomic studies from India.

Sl. No.	Geographical location/place from venom was collected or procured	Number of snakes used for the study	Total number of proteins identified	Number of venom families identified	Reference
1	North-Western (Rajasthan, Gujarat)	17	26	6	[11]
2	Eastern India (Nadia District, West Bengal)-Calcutta Snake park	4	43	15	[9]
3	Eastern India-(Burdwan District, West Bengal)-Calcutta snake park, Kolkata	4	52	14	[10]
4	Southern India-Tamil Nadu-Irula snake catcher's society	Pooled venom	81	13	[5]
5	Southern India-Tamil Nadu-Irula snake catcher's society	4	45	14	[13]
6	Western India-Haffkine Institute, Mumbai	4	54	16	[8]
7	Maharashtra (West India)	5	56	11	[14]
8	Indian cobra venom obtained from Kentucky Reptile Zoo (exact geographical location not mentioned)	Pooled venom	48	17	[12]
9	Western Ghats region in Kerala (India)	2	115	17	Current study

Supplementary Table 1: Proteins identified from crude NNwg venom after SDS-PAGE and RP-HPLC fractionation. The proteins separated through SDS-PAGE and HPLC were subjected to in-gel and in-solution digestion using trypsin, chymotrypsin and V8 protease. The digests were then subjected to Q-TOF LC-MS/MS data collection and the obtained proteins were assigned to different venom protein families through Mascot and PEAKS analysis. #;peptides identified through chymotrypsin digestion, ^;peptides identified through V8 protease digestion, P#; peptides derived through PEAKS analysis after chymotrypsin digestion, P^; peptides derived through PEAKS analysis after V8 protease digestion. All the other peptides were obtained through trypsin digestion.

Sl. No.	**Protein name**	**Accession number**	**Snake species**	**Molecular mass (kDa)**	**No.of peptides**	**No.of unique peptides**	**MS/MS score**	**Sequence Coverage**	**Identified Peptides**	**Observed m/z (Z)**	**ppm**
Three finger toxins											
1	Short neurotoxin 1	P01427	*Naja oxiana*	7	2	2	97	37%	VKPGVNLNCCR	439.5582 (3)	3
									WWSDHRGTIIER	519.2664 (3)	5
2	Weak neurotoxin 6	P29180	*Naja naja*	8	4	3	112	52%	LTCLICPEK	567.2889 (2)	1
									YIRGCADTCPVR	489.9011 (3)	4
									EIVQCCSTDKCNH	551.2202 (3)	3
									LTCLICPE^	1005.4723 (1)	2
3	Chain A, Short neurotoxin	2MJ4_A	*Naja Oxiana*	7	1	1	59	39%	LECHNQQSSQPPTTKTCSGETNCY#	943.063 (3)	1

	1											
4	Chain A, Alpha-cobratoxin	1CTX_A	*Naja siamensis*	8	10	10	793	93%	CFITPDITSK	591.2994 (2)	3	
									VDLGCAATCPTVK	696.3388 (2)	3	
									TGVDIQCCSTDNCNPFPTR	747.9849 (3)	3	
									RVDLGCAATCPTVK	516.5957 (3)	4	
									IRCFITPDITSK	725.8918 (2)	3	
									IRCFITPDITSKDCPNGHVCYTK	696.582 (4)	4	
									CFITPDITSKDCPNGHVCYTK	838.7116 (3)	4	
									TWCDAFCSIR	658.2843 (2)	4	
									VDLGCAATCPTVKTGVDIQCCSTDNCNPFPTR	904.4042(4)	4	
									IRCFITPD [P^]	511.2638 (2)	7	
5	Long neurotoxin 1	P25668	*Naja naja*	8	8	8	758	69%	TWCDGFCSIR	651.2766 (2)	4	
									RVDLGCAATCPTVR	525.931 (3)	3	
									DCPNGHVCYTKTWCDGFCSIR	659.0287 (4)	5	
									VDLGCAATCPTVR	710.3422(2)	4	
									ITPDITSKDCPNGHVCY[#]	659.9621 (3)	2	

									IRCFITPDITSKDCPNGHVCY#	851.728 (3)	1	
									ITPDITSKDCPNGHVCYTKTW#	624.0448 (4)	1	
									GKRVDLGCAATCPTVR	587.6371 (3)	4	
6	Muscarinic toxin-like protein 2	P82463	*Naja kaouthia*	7	7	7	374	60%	SIFGVTTEDCPDGQNLCFKR	782.0311 (3)	4	
									EKSIFGVTTEDCPDGQNLCFKR	651.0602(4)	4	
									SIFGVTTEDCPDGQNLCFK	729.9981(3)	0	
									LTCVKEKSIFGVTTEDCPDGQNL#	871.0853 (3)	1	
									GVTTEDCPDGQNLCF#	856.8486 (2)	1	
									GCAATCPIAENR	660.2914 (2)	6	
									KSIFGVTTE^	491.2668 (2)	1	
7	Three-finger toxin precursor T.A	AKO63248	*Micrurus tener*	9	1	1	47	9%	AVVTIVCL#	874.5022 (1)	5	
8	Toxin B	764177A	*Naja naja*	8	7	3	460	83%	TGVDIQCCSTDDCPFPTRKRP	837.7108 (3)	5	
									IRCFITPDITSK	725.8918(2)	3	
									IRCFITPDITSKDCPNGHVCYTK	696.582 (4)	4	
									CFITPDITSK	591.2987(2)	2	

									CFITPDITSKDCPNGHVCYTK	838.7116(3)	4	
									RVDLGCAATCPTVR	525.931(3)	3	
									VDLGCAATCPTVR	473.8915(3)	9	
9	Weak toxin CM-9a	P25679	*Naja kaouthia*	7	2	1	106	42%	LTCLNCPEMF[#]	1300.5356 (1)	1	
									ITPDITSKDCPNGHVCY[#]	659.9624 (3)	2	
10	Cytotoxin 8	P86540	*Naja naja*	7	6	6	204	83%	LIPLAYKTCPAGK	477.9417 (3)	5	
									MYMVSDKTVPVKR	518.6127 (3)	4	
									GCIDVCPK	474.7189 (2)	3	
									NSLLVKYECCNTDRCN	682.6346 (3)	3	
									GCIDVCPKNSLL [P#]	688.3450 (2)	9	
									GCIDVCPKNSLLVKY [P#]	883.4595 (2)	8	
11	Neurotoxin-like protein NTL2	Q9W717	*Naja atra*	10	1	1	60	12%	GITRLPWVIR	605.8765 (2)	3	
12	Cytotoxin 9	A0A0U5AUY6	*Naja naja*	8	12	12	522	90%	CNKLVPLFYK	641.3564 (2)	2	
									LVPLFYKTCPAGK	498.613 (3)	4	
									LVPLFYKTCPAGKNLCYK	543.7929 (4)	7	
									NLCYKMYMVATPK	540.26	6	

										76 (3)		
									MYMVATPK	470.7368 (2)	3	
									MYMVATPKVPVKR	380.7189 (4)	3	
									GCIDVCPK	474.7181 (2)	1	
									GCIDVCPKSSLLVK	525.9475 (3)	3	
									SSLLVKYVCCNTDR	572.2832 (3)	6	
									SSLLVKYVCCNTDRCN	663.639 (3)	2	
									YVCCNTDRCN	681.2563 (2)	2	
									TCPAGKNLCYK	437.8775 (3)	1	
13	Cytotoxin 1	P01447	*Naja naja*	7	10	3	430	100 %	MYMVSNKTVPVKR	523.9463 (3)	7	
									NSLVLKYECCNTDRCN	682.6346 (3)	3	
									LIPLAYKTCPAGK	477.9417 (3)	5	
									GCIDVCPK	474.7189 (2)	3	
									NSLVLKYECCNTDR	591.2777 (3)	5	
									GCIDVCPKNSLVLK	535.281 (3)	6	
									LIPLAYKTCPAGKNLCYK	528.2877 (4)	5	
									LKCNKLIPLAYK	487.62	4	

										89 (3)		
									CNKLIPLAYK	610.3487 (2)	3	
									KTCPAGKNL#	987.5221(2)	5	
14	Cytotoxin 2	P01445	*Naja kaouthia*	7	10	7	460	100 %	MFMVSNKTVPVKR	523.95(3)	7	
									GCIDVCPK	474.72(2)	3	
									LIPLAYKTCPAGK	477.94(3)	5	
									LKCNKLIPLAYK	487.6289 (3)	4	
									CNKLIPLAYK	610.3487 (2)	3	
									LIPLAYKTCPAGK	477.941 (3)	4	
									LIPLAYKTCPAGKNLCYK	528.2877 (4)	5	
									GCIDVCPKNSLLVK	801.924 (2)	4	
									NSLLVKYVCCNTDRCN	672.6441 (3)	4	
									YVCCNTDRCN	681.2563 (2)	2	
15	Muscarinic toxin-like protein 3	P82464	*Naja kaouthia*	8	8	8	324	94%	TSETTEICPDSWYFCYK	1093.959 (2)	3	
									ISLADGNDVR	530.2762 (2)	2	
									TCPELRPTGIY#	653.827 (2)	1	
									TCPELRPTGIYVY#	784.89	3	

										49 (2)		
									GCTFTCPELRPTGIYVYCCR	837.3771 (3)	3	
									TRTSETTEICPDSWYF#	996.9373 (2)	1	
									CYKISLADGNDVRIKRGCTF#	594.2957 (4)	2	
									TICYNHLTR	393.1971 (3)	1	
16	Neurotoxin-like protein	CAB45156	*Naja naja*	10	6	6	294	65%	SIFSETIEICPDGHNF#	742.8895 (4)	2	
									SETIEICPDGHNF#	831.8872 (2)	1	
									SETIEICPDGHNFCF#	649.8186 (2)	1	
									VIRGCAATCPKAEARVY#	724.8637 (2)	1	
									ARVYVDCCARDKCNR^	643.3226 (2)	3	
									GITRLPWVIR	404.2532 (3)	3	
17	Chain A, Cardiotoxin V4II	1CDT_A	*Naja mossambica*	7	4	3	87	18%	LKCNKLIPIAY#	666.8889 (2)	0	
									KCNKLIPIAY#	610.3474 (2)	1	
									LKCNKLIPIAY P#	666.8954 (2)	10	
									IPIAY P#	576.3439 (1)	8	
18	Muscarinic toxin-like protein 1	P82462	*Naja kaouthia*	7	1	1	59	17%	NQAHLIYPGKY#	652.3433 (2)	1	

19	Chain A, Ringhalexin	4ZQY_A	*Hemachatus haemachatus*	7	2	2	92	34%	RLCLSDYSIFSE^	745.3546 (2)	2	
									GITRLPWVIR	404.254 (3)	5	
20	Chain A, Cardiotoxin V	1CVO_A	*Naja atra*	7	8	8	319	77%	LKCHNTQLPFIYK	554.6351 (3)	4	
									CHNTQLPFIYK	474.2421 (3)	5	
									NLCFKATLK	365.5405 (3)	4	
									KFPLKFPVK	368.5727 (3)	5	
									FPLKFPVK	488.3065 (2)	3	
									NSALLKYVCCSTDKCN	644.9608 (3)	3	
									FPLKFPVKR	377.9074 (3)	3	
									LKCHNTQLPFIYKTCPE^	717.0243(3)	1	
21	Cytotoxin 13, partial	BAU24670	*Naja naja*	8	3	2	134	33%	GCIDVCPK	474.721(2)	7	
									NSLLVKYVCCNTDR	581.286 (3)	4	
									NSLLVKYVCCNTDRCN	672.6443 (3)	4	
22	Cytotoxin 1	P0CH80	*Naja kaouthia*	7	9	3	417	100 %	GCIDVCPKNSLVLK	535.281 (3)	6	
									NSLVLKYVCCNTDRCN	672.6443 (3)	4	
									GCIDVCPK	474.721 (2)	7	

									MFMVSNKTVPVKR	384.9681 (4)	8	
									LKCNKLVPLFYK	761.9469 (2)	3	
									CNKLVPLFYK	641.3564 (2)	2	
									LVPLFYKTCPAGK	498.613 (3)	4	
									LVPLFYKTCPAGKNLCYK	543.7929 (4)	7	
									YVCCNTDRCN	681.2563 (2)	2	
23	Cytotoxin 14, partial	BAU24671	*Naja naja*	8	8	3	414	81%	NSLLVKYECCNTDR	591.2777 (3)	5	
									NSLLVKYECCNTDRCN	682.636 (3)	5	
									CNKLIPLAYK	610.3487 (2)	3	
									LIPLAYKTCPAGK	477.941 (3)	4	
									LIPLAYKTCPAGKNLCYK	528.2877(4)	5	
									GCIDVCPK	474.7181 (2)	1	
									GCIDVCPKNSLLVK	801.924 (2)	4	
									MYMVSNKTVPVKR	388.9647 (4)	2	
24	Cytotoxin D1	1007132A	*Naja atra*	7	2	1	105	40%	SSLLVKYVCCNTDRCN	663.6377 (3)	1	
									GCIDVCPK	474.721 (2)	7	
25	Cytotoxin 2	P01440	*Naja*	7	10	3	457	100	SSLVLKYVCCNTDRCN	663.63	2	

			naja					%		9 (3)		
									GCIDVCPK	474.7181 (2)	1	
									LKCNKLVPLFYK	761.9469 (2)	1	
									SSLVLKYVCCNTDR	572.2832 (3)	6	
									MYMVATPKVPVKR	507.2897 (3)	4	
									CNKLVPLFYK	641.3564 (2)	2	
									LVPLFYKTCPAGK	498.613 (3)	4	
									LVPLFYKTCPAGKNLCYK	543.7929 (4)	7	
									NLCYKMYMVATPK	540.2676 (3)	6	
									MYMVATPK	470.7368 (2)	3	
26	Cardiotoxin isoform 1, Cytotoxin isoform 1	AAB25732	*Naja naja*	7	4	2	188	40%	GCIDVCPKSNLLVK	801.924 (2)	4	
									SNLLVKYVCCNTDRCN	672.6441 (3)	4	
									GCIDVCPK	474.7181 (2)	1	
									YVCCNTDRCN	681.2563 (2)	2	
27	Cytotoxin SP15d	P60309	*Naja atra*	7	2	1	89	23%	GCINVCPKSSLLVK	788.4181 (2)	3	
									GCINVCPK	474.7181 (2)	1	
28	Toxin,cardio	0406231A	*Naja atra*	7	12	6	554	100%	MFMVATPKVPVKR	507.2897 (3)	4	

									LKCNKLVPLFYK	508.3007 (3)	4	
									CNKLVPLFYK	641.3564 (2)	2	
									LVPLFYKTCPAGK	498.613 (3)	4	
									LVPLFYKTCPAGKNLCYK	543.7929(4)	7	
									GCIDVCPK	474.7181 (2)	1	
									SSLLVKYVCCNTDR	572.2832 (3)	6	
									MFMVATPK	470.7366 (2)	3	
									SSLLVKYVCCNTDRCN	663.6368 (3)	1	
									KTCPAGKNL#	494.7678 (2)	4	
									KTCPAGKNLCY#	656.8085 (2)	6	
									VCCNTDRCN#	599.724 (2)	1	
29	Cytotoxin sagitoxin	P83345	*Naja sagittifera*	7	2	1	93	38%	SLLVKYECCNTDRCN	644.6205 (3)	3	
									GCIDVCPK	474.7181 (2)	1	
30	Chain A, Neurotoxin II	1NOR_A	*Naja Oxiana*	7	2	2	119	43%	LECHNQQSSQPPTTK	585.6074 (3)	3	
									VKPGVNLNCCR	439.5553 (3)	4	
31	Cytotoxin 7	P86382	*Naja naja*	7	3	3	144	48%	LIPLAYKTCPAGK	477.9402 (3)	2	
									NSLLVKYECCNTDR	591.27	2	

										62 (3)	
									NSLLVKYECCNTDRCN	682.6347 (3)	3
32	Cardiotoxin N	CAA90966	*Naja naja*	9	1	1	13	10%	VVATIVCL#	874.5021 (1)	5
33	Long chain neurotoxin precursor	AAT97250	*Ophiophagus hannah*	10	1	1	35	14%	CCSTDNCNPFPTR	543.5418 (3)	9
Kunitz-type serine protease Inhibitor (KSPI)											
34	Kunitz-type serine protease inhibitor	P20229	*Naja naja*	6	5	4	195	60%	RPGFCELPAAK	415.8872 (3)	5
									AHKPAFYYNK	413.5516 (3)	7
									FIYGGCGGNANR	643.2932 (2)	4
									IYGGCGGNANRF#	643.2911 (2)	1
									RPGFCELPAAKGL P#	643.2955 (2)	8
35	Kunitz-type serine protease inhibitor	Q5ZPJ7.1	*Naja atra*	9	3	3	94	42%	RPRFCELAPSAGSCF P#	585.6139 (3)	10
									FCELAPSAGSCF P#	1345.5602 (1)	4
									AFVPSYY P#	846.4078 (1)	5
Phospholipase A_2 (PLA_2)											
36	Chain A, Phospholipase A_2	1A3D_A	*Naja naja*	13	28	25	1448	100%	ISGCWPYFK	579.2771(2)	2
									NLYQFKNMIK	657.8491(2)	3

									NMIKCTVPSR	402.5434(3)	2	
									CTVPSRSWWDFADYGCYCGR	848.3523(3)	4	
									SWWDFADYGCYCGR	921.8565(2)	3	
									GGSGTPVDDLDR	594.7802(2)	3	
									CCQVHDNCYNEAEKISGCWPYFK	742.0635(4)	5	
									TYSYECSQGTLTCK	566.578 (3)	3	
									LAAICFAGAPYNDNNYNIDLK	1179.0696(2)	4	
									DFADYGCY#	1010.3547 (1)	1	
									CGRGGSGTPVDDLDRCCQVHDNCY#	700.2852 (4)	2	
									NEAEKISGCWPY#	727.324 (2)	1	
									NEAEKISGCWPYFKTY#	664.9779 (3)	2	
									SYECSQGTLTCKGDNNACAASVCDCDRL#	1071.4335 (3)	2	
									AEKISGCWPYFKTYSYE^	710.3292 (3)	2	
									NLYQFKNMIKCTVPSR	667.017 (3)	6	
									ISGCWPYFKTYSYECSQGTLTCK	946.0887 (3)	3	

									TYSYECSQGTLTCKGDNNACAASVCDCDR	841.5856 (4)	4	
									GDNNACAASVCDCDRLAAICFAGAPYNDNNYNIDLK	1341.5809 (3)	3	
									LAAICFAGAPYNDNNYNIDLKAR	646.8242 (4)	5	
									TCKGDNNACAASVCDCDRL[#]	729.6284 (3)	1	
									TCKGDNNACAASVCDCDRLAAICF[#]	1375.0691 (2)	2	
									AGAPYNDNNYNIDL[#]	777.3489 (2)	2	
									NDNNYNIDLKARCQ[#]	579.9395 (3)	4	
									GDNNACAASVCDCDR	842.8049 (2)	5	
									NDNNYNIDL[#]	1094.4761 (1)	1	
									KNMIKCTVPSRSWW [P#]	598.3008 (3)	6	
									DNCYNEAEKISGCWPY [P#]	1003.4136 (2)	7	
37	Phospholipase A_2 E3	0508173A	*Naja naja*	13	20	13	1083	100%	SWLDFANYGCYCGR	885.3587 (2)	3	
									LAAICFAGAPYNNDNYNINLK	1178.5765 (2)	3	
									LAAICFAGAPYNNDNYNINLKAR	861.77(3)	4	
									NLYQFKNMIK	657.85(2)	3	
									NLYQFKNMIKCTVPSR	500.52(	9	

										4)		
									NMIKCTVPSR	402.54(3)	2	
									GGSGTPVDDLDR	594.78(2)	3	
									DFANYGCY#	1010.3547 (1)	1	
									AGAPYNNDNYNINL#	776.8552 (2)	1	
									NNDNYNINL#	547.2497 (2)	1	
									NNDNYNINLKARCQ#	579.9383 (3)	2	
									CTVPSRSWLDFANYGCYCGR	824.02(3)	5	
									GGSGTPVDDLDRCCQIHDNCYNEAGKISGCWPYFK	816.15(5)	2	
									CCQIHDNCYNEAGKISGCWPYFK	727.81(4)	5	
									TYSYECSQGTLTCKGDNNSCAASVCDCDR	845.59(4)	10	
									ISGCWPYFK	579.278 (2)	3	
									ISGCWPYFKTYSYECSQGTLTCK	946.0887 (3)	3	
									TYSYECSQGTLTCK	566.5788 (3)	4	
									NEAGKISGCWPY#	691.8074 (2)	2	
									AGKISGCWPYFKTYSYE^	686.3242 (3)	5	
38	Chain A,	1LN8_A	*Naja*	13	3	2	185	39%	TYSYECTQGTLTCK	856.37	2	

	Phospholipase A_2		*sagittifera*							05 (2)		
									LAAICFAGAPYNDDNYNIDLK	786.7134 (3)	787	
									GGSGTPVDDLDR	594.7802 (2)	3	
39	Phospholipase A_2	BAA36404	*Naja kaouthia*	16	18	7	1002	94%	LAAICFAGAPYNNNNYNIDLK	1179.0696 (2)	4	
									LAAICFAGAPYNNNNYNIDLKAR	861.7676 (3)	4	
									SWWDFADYGCYCGR	921.8565 (2)	3	
									GGSGTPVDDLDR	594.7802 (2)	3	
									CCQVHDNCYNEAEKISGCWPYFK	742.0635 (4)	5	
									ISGCWPYFK	579.2774(2)	2	
									TYSYECSQGTLTCK	566.578 (3)	3	
									DFADYGCY#	1010.3547 (1)	1	
									CGRGGSGTPVDDLDRCCQVHDNCY#	700.2852 (4)	2	
									NEAEKISGCWPY#	727.324 (2)	1	
									NEAEKISGCWPYFKTY#	664.9779 (3)	2	
									GGSGTPVDDLDRCCQVHDNCYNEAEKISGCWPYFK	827.7593 (5)	5	
									CCQVHDNCYNEAEKISGCWPYFK	742.0653 (4)	8	

									ISGCWPYFKTYSYECSQGTLTCK	849.363 (2)	2	
									AGAPYNNNNYNIDL#	777.3489 (2)	2	
									NNNNYNIDLKARCQ#	579.9395 (3)	4	
									NNNNYNIDL#	1094.4763 (1)	1	
									NMIQCTVPSR	603.2881 (2)	6	
40	Chain A, Phospholipase A_2	1POA_A	*Naja atra*	13	8	1	482	91%	LAAICFAGAPYNDNDYNINLK	786.3834 (3)	5	
									SWWDFADYGCYCGR	921.8565 (2)	3	
									GGSGTPVDDLDR	594.7802 (2)	3	
									CCQVHDNCYNEAEKISGCWPYFK	742.0635 (4)	5	
									ISGCWPYFK	579.2774 (2)	2	
									TYSYECSQGTLTCK	566.578 (3)	3	
									GGSGTPVDDLDRCCQVHDNCYNEAEKISGCWPYFK	1034.4458 (4)	3	
									ISGCWPYFKTYSYECSQGTLTCK	946.0887(3)	3	
41	Chain A, Phospholipase A_2 Isoform 3	1T37_A	*Naja sagittifera*	13	9	4	544	67%	TYSYECTQGTLTCK	856.3708(2)	2	
									TYSYECTQGTLTCKGDNNACAASVCDCDR	1126.4497 (3)	3	
									LAAICFAGAPYNDDNYNIDLK	1179.5669(2)	8	

									LAAICFAGAPYNDDNYNIDLKAR	862.424(3)	5	
									GGSGTPVDDLDR	594.779(2)	1	
									TCKGDNNACAASVCDCDRL#	729.6284 (3)	1	
									TCKGDNNACAASVCDCDRLAAICF#	917.049 (3)	2	
									NLYQFKNMIQCTVPSR	667.0049 (3)	6	
									NMIQCTVPSR	603.2928 (2)	1	
42	Chain A, Phospholipase A_2, Acidic.	2WQ5_A	*Naja naja*	13	12	3	620	79%	TYSYKCSQGTLTCKGGNNACAASVCDCDR	1102.7773 (3)	2	
									NLYQFKNMIK	649.8522 (2)	3	
									NLYQFKNMIKCTVPSR	667.017 (3)	6	
									CTVPSRSWWDFADYGCYCGR	848.3534 (3)	5	
									SWWDFADYGCYCGR	921.8565 (2)	3	
									GGSGTPVDDLDR	594.779 (2)	1	
									LAAICFAGAPYNDNNYNIDLK	786.3829 (3)	5	
									LAAICFAGAPYNDNNYNIDLKAR	646.8242 (4)	5	
									TCKGGNNACAASVCDCDRLAAICF#	898.0412 (3)	2	
									DFADYGCY#	1010.3	1	

										569 (1)		
									NMIKCTVPSR	402.8665 (3)	10	
									TYSYKCSQGTLTCK	849.3706 (2)	10	
43	Phospholipase A_2	BAB72252	*Laticauda colubrina*	17	1	1	46	13%	AALCFAKSPYNNNNYNIDIK	778.0418 (3)	8	
44	Phospholipase A_2	PSNJ3K	*Naja kaouthia*	13	9	2	516	82%	NLYQFKNMIQCTVPSR	667.0049 (3)	6	
									NMIQCTVPSR	603.2881 (2)	6	
									SWWNFADYGCYCGR	614.9076 (3)	4	
									GGSGTPVDDLDR	594.7795 (2)	2	
									ISGCWPYFK	579.2781(2)	3	
									TYSYECSQGTLTCK	849.361 (2)	1	
									LAAICFAGAPYNNNNYNIDLK	786.7116 (3)	6	
									NFADYGCY#	1010.3559 (1)	1	
									NNNNYNIDLKARCQ#	579.9384 (3)	2	
45	Phospholipase A_2, partial	CAM34525	*Naja atra*	12	4	2	189	36%	IQCTVPSRSWW#	710.3474 (2)	3	
									DFADYGCY#	1010.3569(1)	1	
									NEAEKISGCWPY#	727.3268 (2)	3	

									SYECSQGTL	1044.4353(1)	5
46	Chain B, Phospholipase A_2 isoform 2	1S6B_B	*Naja sagittifera*	13	4	2	195	41%	TCTGRNNACAASVCDCDRLAAICF#	921.7215 (3)	1
									AGAPYNDNNYNIDL#	1010.356 (1)	1
									DFADYGCY#	777.3489 (2)	2
									NDNNYNIDL#	1094.4763 (1)	1
47	Phospholipase A_2	1MH7_A	*Naja sagittifera*	13	1	1	63	29%	TYTYQCTQGTLTCK	863.3789 (2)	3
Phospholipase B (PLB)											
48	Phospholipase B	JAC94989	*Opheodrys aestivus*	64	3	3	109	7%	KQVVPESLLAW#	635.3642 (2)	1
									AINGPPVEKGLPIFSW#	862.9733 (2)	1
									NFDFVTMKPVL#	655.8445 (2)	1
Vespryn /Ohanin family (Vesp/OLP)											
49	Vespryn	F8S122	*Crotalus adamanteus*	22	1	1	50	5%	EWAVGIAGK	465.7572 (2)	3
50	Vespryn	JAC96417	*Hypsiglena sp. JMG-2014*	20	2	2	129	16%	ADVTFDSNTAFKSLVVSPDKK	757.3839 (3)	8
									EWAVGLAGK	465.7572 (2)	3

51	Ohanin precursor	AAR07992	*Ophiophagus hannah*	21	1	1	37	11%	ADVTFDSNTAFESLVVSPDKK	757.3804 (3)	3	
52	Thaicobrin	P82885	*Naja kaouthia*	12	14	13	652	85%	ADVTFDSNTAFESLVVSPDKK	757.3839 (3)	8	
									TVENVGVSQVAPDNPER	905.953 (2)	3	
									TVENVGVSQVAPDNPERFDGSPCVLGSPGFR	1096.5327 (3)	3	
									FDGSPCVLGSPGFR	748.3528(2)	1	
									EWAVGLAGK	465.7572 (2)	3	
									SPPGNWQKADVTFDSNTAFESLVVSPDKK	791.8965 (4)	3	
									YGTQREWAVGLAGK	512.6065 (3)	5	
									SPPGNWQKADVTF#	723.8549 (2)	1	
									VVSPDKKTVENVGVSQVAPDNPERF#	678.6057 (4)	1	
									SPPGNWQKADVTFDSNTAFE^	737.6713 (3)	1	
									DSNTAFESLVVSPDKKTVENVGVSQVAPDNPERF#	919.7082 (4)	1	
									NVGVSQVAPDNPE^	663.3209 (2)	1	
									LRLVPEERIW P#	442.9255 (3)	7	
									RLVPEERIW P#	399.900 (3)	10	

53	Ophanin	AAO62996	*Ophiophagus hannah*	27	2	2	65	6%	KMQWYPEAASNAERW#	623.2855 (3)	1
									MQWYPEAASNAER	777.3405 (2)	3
54	Opharin precursor	ACN93671.1	*Ophiophagus hannah*	26	3	2	165	19%	NMLQMEWNSNAAQNAK	925.4233 (2)	4
									CSFAHSPPHLR	327.9121 (4)	1
									FSCGENLFMSSQPYAWSR	728.3113 (3)	6
Snake venom metalloprotease (SVMP)											
55	Metalloproteinase atrase A	ADF43026	*Naja atra*	68	9	9	456	12%	TRVYEMVNYLNTK	544.2818 (3)	3
									VYEMVNYLNTK	695.3404 (4)	1
									ERPQCILNKPSR	499.937 (3)	1
									CGTLYCTEIK	622.7862 (2)	1
									HIALIGLEIW#	582.8415 (2)	2
									GSVAVVQDY#	937.4582 (1)	5
									VYEMVNYLNTKYR	570.2841 (3)	1
									CGTLYCTEIKK	458.2257 (3)	2
									LGPNYTVGPAGCF#	676.8201 (2)	2
56	Hemorrhagic	P82942	*Naja kaouthia*	44	21	20	1090	51%	YYNYDKPAIK	425.5516 (3)	1

	metalloproteinase-disintegrin-like kaouthiagin								IRVYEMINAVNTK	517.6175 (3)	1	
									VYEMINAVNTK	641.3299 (2)	1	
									TAPAFQFSSCSIR	736.3567 (2)	4	
									DYQEYLLR	550.2765 (2)	3	
									AAKHDCDLPELCTGQSAECPTDSLQR	740.5821 (4)	3	
									HDCDLPELCTGQSAECPTDSLQR	897.0514 (3)	4	
									CPTLTNQCIALLGPHFTVSPK	785.4074 (3)	3	
									GCFDLNMR	506.7214 (2)	2	
									GCFDLNMRGDDGSFCR	641.591 (3)	1	
									MEDGTKIPCAAK	660.8176 (2)	1	
									GQCVDVQTAY	570.7528 (2)	1	
									EMINAVNTKF[#]	583.796 (2)	2	
									EMINAVNTKFRPL[#]	511.6128 (3)	1	
									NYDKPAIKIRVY[#]	493.9484 (3)	1	
									SSCSIRDYQEY[#]	704.2942 (2)	2	
									TNTPEQDRYLQAEK	564.946 (3)	3	

									YIEFYVIVDNR	715.8756 (2)	9	
									TAPAFQFSSCSIRDYQEYLLR	851.4133 (3)	1	
									STDIVSPAICGNY#	698.8252 (2)	1	
									NGKCPTLTNQCIAL#	795.8866 (2)	3	
57	K-like metalloprotease precursor, partial	ACN50005	*Naja atra*	66	9	2	531	17%	VAKDDCDLPELCTGQSAECPTDSLQR	989.1071 (3)	3	
									SFGEWRETVLLPR	530.6213 (3)	4	
									IRVYEMINAVNTK	517.6175 (3)	1	
									VYEMINAVNTK	641.3299 (2)	1	
									TAPAFQFSSCSIR	736.3567 (2)	4	
									GCFDLNMR	506.7214 (2)	2	
									GCFDLNMRGDDGSFCR	641.591 (3)	1	
									MEDGTKIPCAAK	660.8176 (2)	1	
									GQCVDVQTAY	570.7528 (2)	1	
58	Chain A, Atragin	3K7L_A	*Naja atra*	48	23	20	1427	62%	TNTPEQDRYLQAK	521.93 (3)	1	
									ATLNLFGEWR	604.3111 (2)	4	
									ATLNLFGEWREK	488.9225 (3)	3	

									TKPAYQFSSCSVR	510.919 (3)	4	
									AAKDDCDLPELCTGQSAECPTDVFQR	995.1032 (3)	3	
									CPIMTNQCIALR	738.8614 (2)	2	
									VSRDSCFTLNQR	494.9097 (3)	2	
									DSCFTLNQR	570.7601 (2)	2	
									DPNYGMVEPGTK	654.3011 (2)	1	
									KRNDNAQLLTGIDFNGTPVGLAYIGSICNPK	837.4404 (4)	5	
									RNDNAQLLTGIDFNGTPVGLAYIGSICNPK	1073.55 (3)	3	
									NDNAQLLTGIDFNGTPVGLAYIGSICNPK	1531.7691 (2)	2	
									RTKPAYQFSSCSVR	562.9542 (3)	6	
									TKPAYQFSSCSVR	765.8754 (2)	4	
									DDCDLPELCTGQSAECPTDVFQR	905.0475 (3)	4	
									NGLPCQNNQGYCYNGK	943.9026 (2)	4	
									CPIMTNQCIALRGPGVK	638.9951 (3)	0	
									NSMICNCSISPR	727.8159 (2)	4	
									QCVDVKTAY	542.2624 (2)	3	

									QRNGLPCQNNQGY#	775.3461 (2)	2	
									TGIDFNGTPVGLAY#	712.8567 (2)	1	
									TGIDFNGTPVGL#	595.806 (2)	1	
									LQHEAQCDSEECCEK	641.5876 (3)	9	
59	Metalloproteinase	ADI47610	*Echis coloratus*	69	1	1	64	2%	VTLNLFGEWR	618.3277 (2)	5	
60	Mocarhagin 1	AAM51550	*Naja mossambica*	68	6	5	227	10%	DRPQCILNKPSR	495.2666 (3)	4	
									AAKNDCDFPELCTGR	585.5891 (3)	3	
									RVYEMVNALNTMYR	587.2954 (3)	6	
									SVAVVQDHSK	535.2900 (2)	8	
									FKGAGAECR	498.2449 (2)	10	
61	Cobrin precursor	AAF00693	*Naja naja*	68	32	18	1901	59%	DPSYGMVEPGTK	648.7924 (2)	1	
									ATLDLFGEWR	604.3115 (2)	4	
									TSAAVVQDYSK	584.7976 (2)	3	
									DSCFTLNQR	570.7616 (2)	5	
									TSAAVVQDYSKSTR	504.9276 (3)	2	
									ATLDLFGEWREK	604.3111 (2)	3	

									TKPAYQFSSCSVR	510.919 (3)	4	
									VSRDSCFTLNQR	494.9097 (3)	2	
									YIEFYMVVDNIMYR	936.4425 (2)	4	
									KRNDNAQLLTGIDFNGTPVGLAYIGSICNPK	837.4404 (4)	5	
									RNDNAQLLTGIDFNGTPVGLAYIGSICNPK	1073.55 (3)	3	
									NDNAQLLTGIDFNGTPVGLAYIGSICNPK	1021.517 (3)	4	
									RTKPAYQFSSCSVR	562.9542 (3)	6	
									TKPAYQFSSCSVR	765.8754 (2)	4	
									AAKDDCDLPELCTGQSAECPTDVFQR	746.5811 (4)	5	
									DDCDLPELCTGQSAECPTDVFQR	905.0475 (3)	4	
									CPIMTNQCIALR	738.8618 (2)	3	
									CPIMTNQCIALRGPGVK	638.9951 (3)	1	
									NSMICNCSISPR	727.8159 (2)	4	
									QCVDVKTAY	542.2624 (2)	3	
									QAEKYIEFY[#]	595.7939 (2)	6	
									TGIDFNGTPVGL[#]	595.806 (2)	1	

									TGIDFNGTPVGLAY#	712.8567 (2)	1	
									LQHEAQCDSEECCEK	641.5876 (3)	9	
									FKGAGAECR	994.4753 (2)	10	
									YMVVDNIMY#	574.2642 (2)	4	
									SNINEINVQSDVKATL#	582.3073 (3)	1	
									HIALIGLEIW#	582.8419 (2)	1	
									IGSICNPKTSAAVVQDY#	608.3064 (3)	2	
									AYIGSICNPKTSAAVVQDY#	686.3396 (3)	2	
									NGKCPIMTNQCIAL#	810.3895 (2)	1	
									KGAGAECRAAKDDCDLPELCTGQSAECPTDVF#	882.8868 (4)	3	
62	Metalloprotease	XP_015671647	*Protobothrops mucrosquamatus*	133	1	1	2	1%	QKPIASQACNINF#	746.8542 (2)	6	
63	P-III snake venom metalloprotease, partial	AHZ08819.1	*Micropechis ikaheca*	68	1	1	57	1%	CPTDSFQR	505.7263 (2)	9	
64	Scutellatease-1	ABQ01136	*Oxyuranus scutellatu*	69	2	1	88	4%	LTNTPEQER	544.2776 (2)	9	
									LQHEAQCDSEECCEK	641.58	9	

			s							76 (3)		
Nerve growth factor (NGF)												
65	Nerve growth factor beta chain precursor	A59218	*Naja kaouthia*	28	8	8	477	48%	EDHPVHNLGEHSVCDSVSAWVTK	651.558 (4)	6	
									TTATDIKGNTVTVMENVNLDNKVYK	692.8625 (4)	5	
									GIDSSHWNSYCTETDTFIK	754.3351 (3)	4	
									ALTMEGNQASWR	682.3266 (2)	3	
									TTATDIKGNTVTVMENVNLDNK	793.4036 (3)	4	
									HSVCDSVSAWVTKTTATDIKGNTVTVME^	760.1237 (4)	4	
									NVNLDNKVYKQYFFE^	960.9794 (2)	1	
									TDTFIKALTME^	635.3242 (2)	1	
66	Chain A, Venom growth factor	4EC7_A	*Naja atra*	13	10	10	422	83%	VTKTTATDIKGNTVTVMENVNLDNKVY#	742.8895 (4)	2	
									NSYCTETDTFIKAL#	831.8872 (2)	1	
									CTETDTFIKAL#	649.8186 (2)	1	
									IKALTMEGNQASW#	724.8637 (2)	1	
									TDTFIKALTME^	643.3226 (2)	3	
									TTATDIKGNTVTVMENVNLDNKVYK	692.8623 (4)	5	

									GIDSSHWNSYCTETDTFIK	754.3369 (3)	6	
									ALTMEGNQASWR	682.3261 (2)	2	
									TTATDIKGNTVTVMENVNLDNK	793.4070 (3)	8	
									EDHPVHNLGEHSVCDSVSAWP#	759.0060 (3)		
67	Nerve growth factor II	AAS94269	*Naja sputatrix*	27	14	13	800	48%	EDHPVHNLGEHSVCDSVSAWVTK	651.5576 (4)	5	
									TTATDIKGNTVTVMENVNLDNK	793.404 (3)	4	
									TTATDIKGNTVTVMENVNLDNKVYK	692.8624 (4)	5	
									GNTVTVMENVNLDNKVYK	680.0197 (3)	8	
									GIDSSHWNSYCTETDTFIK	754.3358 (3)	5	
									ALTMEGNQASWR	682.3267 (2)	3	
									FIRIDTACVCVITK	565.9709 (3)	3	
									FIRIDTACVCVITKK	456.7553 (4)	7	
									IDTACVCVITK	640.3291 (2)	10	
									GNTVTVMENVNLDNK	824.405 (2)	1	
									VYKQYFFETK	676.8492 (2)	2	

									GIDSSHWNSYCTETDTFIKALTMEGNQASWR	906.1592 (4)	4	
									IDTACVCVITKK	469.9174 (3)	3	
									TDTFIKALTME^	643.3222 (2)	2	
68	Venom nerve growth factor precursor	ABA60123	*Pseudechis porphyriacus*	27	2	1	147	13%	GIDSSHWNSYCTTTQTFVRALTMEGNQASWR	902.1619 (2)	2	
									ALTMEGNQASWR	682.3267 (2)	3	
69	Venom nerve growth factor	P61899	*Naja kaouthia*	13	3	3	145	35%	VTKTTATDIKGNTVTVMENVNLDNKVY#	742.8912 (4)	4	
									NSYCTETDTFIKAL#	831.8878 (2)	1	
									CTETDTFIKAL#	649.8201 (2)	2	
Cysteine-rich secretory proteins (CRISP)												
70	Kaouthin-2 precursor	ACH73168.1	*Naja kaouthia*	26	17	17	1089	77%	ENQKQIVDKHNALR	423.9856 (4)	4	
									NMLQMEWNSNAAQNAK	627.9481 (3)	4	
									NMLQMEWNSNAAQNAKR	669.3174 (3)	2	
									CSFAHSPPHLR	436.8813 (3)	2	
									TVGKIGCGENLFMSSQPYAWSR	830.0661 (3)	4	
									IGCGENLFMSSQPYAWSR	706.9863 (3)	3	
									VIQSWYDENK	641.30	1	

										98 (2)		
									VIQSWYDENKK	470.5747 (3)	2	
									YLYVCQYCPAGNIIGSIATPYK	851.0855 (3)	2	
									SGPPCGDCPSACVNGLCTNPCK	803.3262 (3)	3	
									HHNVFSNCQSLAK	514.5816 (3)	2	
									QNACQTEWMK	648.2808 (2)	3	
									CAASCFCRTEII	744.3279 (2)	3	
									VYGVGANPPGSVIGHY#	793.9009 (2)	1	
									GVGANPPGSVIGHY#	662.8339 (2)	3	
									CPAGNIIGSIATPY#	717.3552 (2)	4	
									SRVIQSWYDENKKF P#	600.6443 (3)	6	
71	Natrin	1XTA_A	*Naja atra*	25	27	27	1500	86%	NVDFNSESTR	584.7662 (2)	2	
									VSPTASNMLK	524.2794 (2)	1	
									MEWYPEAASNAER	785.3376 (2)	3	
									VLEGIQCGESIYMSSNAR	1015.4722 (2)	3	
									VSPTASNMLKMEWYPEAASNAER	872.071 (3)	3	
									SNCPASCFCR	629.74	4	

										43 (2)		
									NVDFNSESTRR	442.2138 (3)	2	
									KMEWYPEAASNAERW#	623.2863 (3)	1	
									TEIIHLWHDEY#	485.9007 (3)	2	
									KQKEIVDLHNSLR	527.3047 (3)	5	
									QKEIVDLHNSLR	484.6051 (3)	3	
									QKEIVDLHNSLRR	402.7301 (4)	1	
									RVSPTASNMLK	602.3298 (2)	1	
									RVSPTASNMLKMEWYPEAASNAER	689.333 (4)	5	
									MEWYPEAASNAERWANTCSLNHSPDNLR	830.6245 (4)	4	
									TATPYKLGPPCGDCPSACDNGLCTNPCTIYNK	901.6417 (4)	4	
									WANTCSLNHSPDNLR	595.9396 (3)	4	
									TWTEIIHLWHDEYK	468.4854 (4)	4	
									LGPPCGDCPSACDNGLCTNPCTIYNK	1471.1084 (2)	2	
									ANTCSLNHSPDNLRVL#	604.2994 (3)	2	
									RVLEGIQCGESIY#	762.376 (2)	5	
									EIVDLHNSLR	598.32	1	

										45 (2)		
									VSPTASNMLK	532.2771 (2)	2	
									ANTCSLNHSPDNLRVL#	604.3004 (3)	1	
									NHSPDNLRVLEGIQCGESIY#	767.7023 (3)	1	
									AASNAERWANTCSLNHSPDNLRVLE^	707.0907 (4)	1	
									RWANTCSLNHSPDNLRVLE^	571.2838 (4)	2	
Phosphodiesterase (PDE)												
72	Phosphodiesterase 1	JAB52814	*Micrurus fulvius*	97	34	34	1720	40%	MANVLCSCSEDCLTK	902.3754 (2)	3	
									MANVLCSCSEDCLTKK	644.618 (3)	3	
									DCCTDYKSICKR	402.1783 (4)	6	
									AEYLETWDTLMPNINK	969.472 (2)	3	
									AEYLETWDTLMPNINKLK	732.3765 (3)	4	
									AATYFWPGSEVK	678.3368 (2)	2	
									AKRPDFSTLYIEEPDTTGHK	577.045 (4)	5	
									RPDFSTLYIEEPDTTGHK	527.2597 (4)	1	
									TLGMLMEGLK	546.7952 (2)	3	
									TLGMLMEGLKQR	464.91	4	

										82 (3)		
									NLHNCVNLILLADHGMEAISCNR	666.8294 (4)	6	
									SKNVPKDFYTFDSEAIVK	522.7749 (4)	6	
									NVPKDFYTFDSEAIVK	624.9855 (3)	2	
									DFYTFDSEAIVK	717.8457 (2)	4	
									IDKVNLMVDR	401.5591 (3)	5	
									IDKVNLMVDRQWLAVR	489.7796 (4)	6	
									VNLMVDRQWLAVR	533.966 (3)	3	
									YCSGGTHGYDNEFK	545.5581 (3)	3	
									NPFYNPSPAK	567.7837 (2)	2	
									CSSITDLEAVNQR	746.8596 (2)	4	
									YISAYSQDILMPLWNSYTISK	837.0864 (3)	5	
									LWNYFHSTLLPK	506.945 (3)	6	
									VLSFILPHRPDNSESCADKSPNNLWVEER	682.9384 (5)	6	
									SPNNLWVEER	622.7999 (2)	2	
									QPLSETLR	472.2655 (2)	3	
									CSCSEDCLTKKDCCTDY#	734.60	2	

										32 (3)		
									VNHYTIVTGLY#	640.3384 (2)	1	
									GPVSGQVIKSL#	542.8232 (2)	3	
									CLLHQTKY#	531.7742 (2)	1	
									DALITSNIVPMYKEF#	878.9474 (2)	3	
									NCPPGSLKVL#	542.7957 (2)	2	
									NCPPGSLKVLSF#	659.8455 (2)	1	
									KQPLSETL#	458.2616 (2)	1	
									EARVMEVL#	481.7536 (2)	3	
73	Phosphodiesterase, partial	ALA20853	*Macropisthodon rudis*	91	7	2	390	11%	TEVTSFENIEVYNLMCDLLK	1209.5852 (2)	3	
									EKTEVTSFENIEVYNLMCDLLK	892.4386 (3)	2	
									LWNYFHSTLLPK	506.945 (3)	6	
									NLHNCVNLILLADHGMEAISCNR	666.8294 (4)	6	
									NPFYNPSPAK	567.7837 (2)	2	
									QPLSETLR	472.2655 (2)	3	
									SPDNLWVEER	622.7999 (2)	2	
74	Phosphodie	AHJ8088	*Macrovip*	96	5	1	273	6%	SMQAIFLAHGPGFK	507.59	5	

	sterase	5	*era lebetina*							3 (3)		
									AATYFWPGSEVK	678.3368 (2)	2	
									TLGMLMEGLK	554.7938 (2)	5	
									TLGMLMEGLKQR	459.5866 (3)	4	
									NPFYNPSPAK	567.7837 (2)	2	
75	Phosphodiesterase family member 3, partial	XP_007430581	*Python bivittatus*	94	9	4	399	9%	AATYFWPGSDIK	678.3368 (2)	2	
									TLGMLMEGLK	554.7938 (2)	5	
									TLGMLMEGLKQR	459.5866 (3)	4	
									IDKVNLMVDR	401.5591 (3)	5	
									YIEEPDTTGHKF[#]	479.563 (3)	4	
									IEEPDTTGHKF[#]	425.2086 (3)	4	
									CLLHQTKY[#]	531.7742 (2)	1	
									KQPLSETL[#]	458.2616 (2)	1	
									DYQDTCVLPTQSW[#]	806.8536 (2)	3	
76	Phosphodiesterase family member 3	JAG67101	*Boiga irregularis*	101	12	3	542	11%	AEYLETWDTLMPNINK	969.472 (2)	3	
									AEYLETWDTLMPNINKLK	732.3765 (3)	4	
									AATYFWPGSEVK	678.3368 (2)	2	

									NPFYNPSPAK	567.7837 (2)	2	
									LWNYFHSTLLPK	506.945 (3)	6	
									QPLSETLR	472.2655 (2)	3	
									TFDSEAIVKKL#	625.854 (2)	1	
									DSEAIVKKL#	501.7958 (2)	1	
									ANNIRIDKVNL#	423.9153 (3)	3	
									EARVMEVL#	481.7536 (2)	3	
									KQPLSETL#	458.2616 (2)	1	
									DTLMPNINKL#	579.8135 (2)	1	
77	Phosphodiesterase	JAS05142	*Micrurus tener*	97	9	9	486	13%	RPDFSTLYIEEPDTTGHK	527.2624 (4)	6	
									NPFYNPSPAK	567.7848 (2)	4	
									CSSITDLEAVNQR	746.858 (2)	2	
									SPNNLWVEER	622.8051 (2)	10	
									DTLMPNINKL#	579.8135 (2)	1	
									VNHYTIVTGLY#	640.3389 (2)	2	
									GPVSGQVIKSL#	542.8226 (2)	2	

									GPVPSPDLSGCKCSSITDLEAVNQRL[#]	934.1275 (3)	3	
									DALITSNIVPMYKEF[#]	870.949 (2)	2	
78	Chain A, Snake Venom Phosphodiesterase [P#^]	pdb|5GZ4|A	*Naja atra atra*	94	9	8	620	15%	TIVTGLYAETHGIIDNNMY	1071.0222 (2)	7	
									LYPPDFSSSGPEQY	793.8597 (2)	7	
									YNPSPAKEQSPPLY	795.8998 (2)	8	
									GPVSGQVIKSL	542.8268 (2)	10	
									NLIDQAKMQSEADNLPY	975.4763 (2)	9	
									WGGQPIWHTASY	701.8361 (2)	8	
									SLSGSNMRNAAW	647.3091 (2)	8	
									RAVYPTKTF	541.8075 (2)	8	
									LLTGLDFYSALKQPLSE	948.0171 (2)	6	
5' nucleotidase (5-NU)												
79	Ecto-5'-nucleotidase	JAS05143	*Micrurus tener*	63	30	30	1585	57%	LTILHTNDVHAR	463.9257 (3)	2	
									YDAMALGNHEFDNGLNGLLDPLLK	882.7688 (3)	2	
									NVKFPILSANIRPK	399.9963 (4)	3	
									FPILSANIRPK	419.257 (3)	3	

									IINVGSEKVGIIGYTTK	598.0133 (3)	2	
									VGIIGYTTK	476.28 (2)	2	
									ETPVLSNPGPYLEFRDEVEELQK	897.1196 (3)	3	
									QVPVVQAYAFGK	653.8614 (2)	1	
									ASGNPILLNK	513.8015 (2)	1	
									FHECNLGNLICDAVVYNNLR	807.7188 (3)	1	
									ANNGIITLEELTAVLPFGGTFDLLQIK	963.5335 (3)	7	
									HGQGTGELLQVSGIK	508.612 (3)	2	
									VVSLNVLCTECR	725.3642 (2)	2	
									VPTYVPLQMEK	653.3439 (2)	3	
									VLLPSFLAAGGDGYYMLK	966.0096 (2)	7	
									MGKVFPAMEGR	408.2085 (3)	3	
									VFPAMEGR	453.7298 (2)	2	
									IINVGSEK	430.2492 (2)	3	
									SIQEDPAVK	493.7645 (2)	6	
									QAFEHSVHR	555.7784 (2)	5	

									KIINVGSEKVGIIGY#	795.469 (2)	3	
									TTKETPVLSNPGPY#	752.3899 (2)	2	
									LEFRDEVEEL#	639.8155 (2)	2	
									GVNKIIAL#	827.5274 (1)	9	
									GVNKIIALGHSGF#	656.8721 (2)	1	
									MQLSDDGRQVPVVQAY#	903.4476 (2)	1	
									SDDGRQVPVVQAY#	717.3557 (2)	1	
									HECNLGNLICDAVVY#	888.9075 (2)	2	
									CTECRVPTYVPLQMEKTY#	759.3554 (3)	5	
									KVLLPSFL#	916.5877 (1)	1	
80	Snake venom 5'-nucleotidase	JAG67188	*Boiga irregularis*	65	15	3	796	30%	GREVVQFMNSLR	479.2564 (3)	5	
									EVVQFMNSLR	611.8174 (2)	2	
									QVPVVQAYAFGK	653.8614 (2)	1	
									VGIIGYTTK	476.28 (2)	2	
									ETPVLSNPGPYLEFRDEVEELQK	897.1196 (3)	3	
									HGQGTGELLQVSGIK	508.612 (3)	2	

									VPTYVPLEMEK	661.3424 (2)	5	
									FHECNLGNLICDAVVYNNLR	807.7188 (3)	1	
									TTKETPVLSNPGPY#	752.3899 (2)	2	
									LEFRDEVEEL#	639.8155 (2)	2	
									GVNKIIALGHSGF#	656.8721 (2)	1	
									SDDGRQVPVVQAY#	717.3557 (2)	1	
									HECNLGNLICDAVVY#	888.9075 (2)	2	
									KVLLPSFL#	916.5877 (1)	1	
									QVSGIKVVY#	992.579 (1)	2	
81	5' nucleotidase, partial	BAN82018	*Protobothrops flavoviridis*	57	9	3	422	23%	GREVVKFMNSLR	479.2695 (3)	7	
									FHECNLGNLICDAVIYNNVR	807.7188 (2)	1	
									GDSSNHSSGNLDISIVGDYIKR	584.292 (4)	6	
									IINVGSEKVGIIGYTTK	598.0133 (3)	2	
									VGIIGYTTK	476.28 (2)	2	
									ETPVLSNPGPYLEFR	573.9617 (3)	7	
									QVPVVQAYAFGK	653.8614 (2)	1	
									ASGNPILLNK	513.80	1	

										15 (2)		
									VVSLNVLCTECR	725.3642 (2)	2	
82	5'-nucleotidase, partial	ETE67245	*Ophiophagus hannah*	28	6	1	348	30%	HADKLTTLGVNK	432.9155 (3)	4	
									HGQGTGELLQVSGIK	508.612 (3)	2	
									VVSLNVLCTECR	725.3642 (2)	2	
									VPTYVPLEMEK	653.3439 (2)	3	
									VLLPSFLAAGGDGYYMLK	966.0096 (2)	7	
									VFPAMEGR	453.7298 (2)	2	
83	Snake venom 5'-nucleotidase	JAS04365	*Agkistrodon contortrix*	65	7	5	302	13%	LLDAGDQY#	447.7135 (2)	1	
									KIINVGSEKVGIIGY#	530.6476 (3)	1	
									TTKETPVLSNPGPY#	752.3889 (2)	1	
									DLLQIKGSTL#	544.3216 (2)	1	
									QVSGIKVVY#	496.7928 (2)	1	
									LEFRDEVEEL#	639.8155 (2)	2	
									GVNKIIALGHSGF#	656.8721 (2)	1	
84	Ecto-5' nucleotidase, partial	AEJ31979.1	*Crotalus adamanteus*	58	108	3	2	6%	VGIIGYTTK	476.2776 (2)	3	
									QVPVVQAYAFGK	653.8606(2)	1	

									VLSLNVLCTECR	732.3704 (2)	1
Snake venom serine proteinase (SVSP)											
85	Thrombin-like serine protease	AAL68708	*Gloydius ussuriensis*	29	1	1	51	3%	TKVFDYTEW[#]	594.7799 (2)	4
86	Thrombin-like enzyme TLP	P86545.1	*Naja naja*	4	1	1	37	28%	CAGTLLNHEW[#]	600.7764 (2)	1
87	Tissue-type plasminogen activator	JAI11009	*Crotalus adamanteus*	60	4	4	234	10%	AVCYQDIGDTYR	730.829 (2)	2
									GTWSTTESGAECVNWK	906.897(2)	2
									TLDNDIALLK	558.3208 (2)	3
									TVTENMLCAGDTR	734.3349 (2)	4
88	Tissue-type plasminogen activator, partial	ETE66683	*Ophiophagus hannah*	32	4	3	152	15%	GGAYSYIESHPWQAGIFVK	704.0202 (3)	3
									LYPDSFCTSAR	658.8021 (2)	2
									TVTENMLCAGDTR	742.3355 (2)	8
									SYIESHPWQAGIF [P#]	767.8721 (2)	3
89	Tissue-type plasminogen activator, partial	XP_007443169	*Python bivittatus*	38	2	2	70	4%	LIGIISWGVGCGK	680.3778 (2)	2
									IGIISW [P#]	704.3993 (1)	2
90	Tissue-type	XP_0139	*Thamnop*	51	3	2	190	6%	LPDWTECEVSGYGK	820.86	3

	plasminogen activator	24331	*his sirtalis*							95 (2)		
									TVTENMLCAGDTR	742.3355 (2)	8	
									SYIESHPW#	509.7357 (2)	1	
									KLPDWTECEVSGY#	792.3562 (2)	1	
91	Snake venom serine protease NaSP	A8QL53.1	*Naja atra*	31	1	1	68	5%	FPCAQLLEPGVYTK	811.9148 (2)	2	
Cobra Venom Factor (CVF)												
92	Complement decay-accelerating factor transmembrane isoform, partial	ETE63384	*Ophiophagus hannah*	79	1	1	55	1%	IVSSNDLR	452.7419 (2)	3	
93	Pseutarin C	AAO38805	*Pseudonaja textilis*	166	1	1	10	1%	KQEEPRDALSGL#	671.8464 (2)	8	
94	Chain B, Cobra Venom Factor	3PRX_B	*Naja kaouthia*	185	53	53	2739	37%	QLDIFVHDFPR	462.9135 (3)	7	
									QNQYVVVQVTGPQVR	857.9704 (2)	6	
									GIYTPGSPVLYR	661.8634 (2)	7	
									YVLPSFEVR	555.3068 (2)	7	

									IPIIDGDGK	464.2642 (2)	7	
									IPIIDGDGKATLKR	374.9726 (4)	8	
									LILNIPLNAQSLPITVR	938.0822 (2)	8	
									VGLVAVDK	400.7512 (2)	9	
									VGLVAVDKAVYVLNDKYK	499.293 (4)	9	
									DDNEDGFIADSDIISR	891.399 (2)	4	
									YRINYENALLAR	499.2721 (3)	1	
									INYENALLAR	588.8223 (2)	1	
									LNQDITVTASGDGK	709.8604 (2)	1	
									ATMTILTFYNAQLQEK	936.4821 (2)	1	
									VAVIIYLNK	516.8269 (2)	1	
									HFEVGFIQPGSVK	482.2579 (3)	1	
									VYSYYNLDEK	647.3094 (2)	9	
									IDVPLQIEK	527.8109 (2)	1	
									ACETNVDYVYK	681.3118 (2)	9	
									IEEQDGNDIYVMDVLEVIK	741.3712 (3)	4	
									KCQEALNLK	552.29	1	

										82 (2)		
									KCQEALNLKVNDDYLIWGSR	808.0767 (3)	1	
									VNDDYLIWGSR	669.3328 (2)	7	
									SDLLPTKDKISYIITK	612.3611 (3)	9	
									QTRVDMNPAGGML#	695.3375 (2)	4	
									VVVQVTGPQVRL#	647.8958 (2)	1	
									VVVQVTGPQVRLEKVVLL#	988.6198 (2)	3	
									LFIQTDKGIY#	599.3298 (2)	1	
									IQTDKGIYTPGSPVLY#	876.4661 (2)	2	
									NLPDLVSLGTW#	1214.6428 (1)	1	
									LYGEEVEGVAF#	606.7939 (2)	2	
									TRIPIIDGDGKATL#	735.4209 (2)	1	
									YQVGNNEIVADSVW#	797.3836 (2)	3	
									VDVKDTCMGTL#	619.7919 (2)	1	
									KISQAKIWDTIEKSDF#	637.0084 (3)	2	
									TPTKGICVAEPY#	668.3351 (2)	1	
									LILNRQQPDGAF#	686.3739 (2)	2	

									ALAAADQLNDDRVL#	742.8890 (2)	1	
									DQTGPIVRW#	536.2831 (2)	1	
									EIQMPTHKDLNLDITIELPDREVPIRY#	812.9323 (4)	3	
									YIDGNENFHVSITARY#	633.6415 (3)	4	
									NLDITIELPDREVPIRY#	686.0439 (3)	5	
									RINYENALL#	553.3039 (2)	1	
									NAQLQEKANVCNKF#	555.2797 (3)	2	
									TGFLPDAEDLTRL#	724.3776 (2)	3	
									EVDNNMAQKVAVIIY#	853.9431 (2)	1	
									EVGFIQPGSVKVY#	711.8854 (4)	1	
									HQERIDVPLQIEKACETNVDY#	668.5794 (4)	7	
									QIEKACETNVDYVY#	866.3968 (2)	2	
									NHQERIDVPLQIEKACETNVDYVY#	734.1098 (4)	2	
									TKIIIQGDPVAQIIE^	819.4784 (2)	1	
									NALLARTVE^	493.7859 (2)	2	
									FQKLCDDFAQFSYTLTE^	1056.9837 (2)	1	

95	Chain A, Cobra Venom Factor Alpha Chain	3FRP_A	*Naja kaouthia*	70	9	2	471	17%	ALYTLITPAVLR	665.91 (2)	2	
									YVLPSFEVR	555.3068 (2)	7	
									IPIIDGDGK	464.2642 (2)	7	
									QLDIFVHDFPR	462.9135 (3)	7	
									GIYTPGSPVLYR	661.8634 (2)	7	
									LILNIPLNAQSLPITVR	938.0822 (2)	8	
									VGLVAVDK	400.7512 (2)	9	
									VGLVAVDKAVYVLNDKYK	499.293 (4)	9	
									ALYTLITPAVLRTDTEE^	953.51(2)	2	
96	OVF precursor protein	AFJ59923	*Ophiophagus hannah*	184	20	4	982	14%	TDTEEQILVEAHGDNTPK	666.3263 (3)	9	
									QLDIFVHDFPR	462.9135 (3)	7	
									GIYTPGSPVLYR	661.8634 (2)	7	
									HFEVGFIQPGSVK	482.2579 (3)	1	
									RTDTEEQILVEAHGDNTPKQL#	599.3055 (4)	4	
									RTDTEEQILVEAHGDNTPKQLDIF#	693.1002 (4)	3	
									ALAAADRLNDDRVL#	504.9446 (3)	4	

									VVVQVTGPQVRL#	647.8958 (2)	1	
									IQTDKGIYTPGSPVLY#	876.4661 (2)	2	
									LYGEEVEGVAF#	606.7939 (2)	2	
									YQVGNNEIVADSVW#	797.3836 (2)	3	
									VDVKDTCMGTL#	619.7919 (2)	1	
									KISQAKIWDTIEKSDF#	637.0084 (3)	2	
									TPTKGICVAEPY#	668.3351 (2)	1	
									LILNRQQPDGAF#	686.3739 (2)	2	
									DQTGPIVRW#	536.2831 (2)	1	
									NAQLQEKANVCNKF#	555.2797 (3)	2	
									TGFLPDAEDLTRL#	724.3776 (2)	3	
									EVDNNMAQKVAVIIY#	853.9431 (2)	1	
									EVGFIQPGSVKVY#	711.8854 (2)	1	
97	Complement-depleting factor, partial	ABN72542	*Bungarus fasciatus*	143	8	3	319	5%	IEEKDGNDIYVMDVLEVIK	741.3753 (3)	7	
									GNANSLNQIK	529.7876 (2)	8	
									VGLVAVDK	400.751 (2)	9	
									VGLVAVDKAVYVLNDKYK	499.29	9	

										3 (4)		
									HFEVGFIQPGSVK	482.2579 (3)	1	
									KCQEALNLK	552.2982 (2)	1	
									AVPFVIVPLQQGLHDIEVK	701.7338 (3)	1	
									SDFGCTAGSGQNNLGVFEDAGLALTTSTNLNTK	1120.8595 (3)	1	
98	Venom factor	AAX86641	*Austrelaps superbus*	185	6	2	291	4%	YFTYLILTK	581.3316 (2)	1	
									AVYVLNDKYK	606.8333 (2)	3	
									VGLVAVDK	400.7464 (2)	3	
									HFEVGFIQPGSVK	482.2569 (3)	2	
									DTCMGTLVVK	562.2756 (2)	4	
									VELLYNPAFCSASTEGQR	681.328 (3)	2	
C-type lectin (CTL)												
99	C-type lectin BfL-1	Q90WI8	*Bungarus fasciatus*	19	5	5	96	23%	YIWEWTDR	584.7768 (2)	3	
									ISDYLTGQGHVW[#]	688.3337 (2)	3	
									SEYISDYLTGQGHVW [P#]	877.8997 (2)	6	
									HSNADAVEFSEY [P#]	692.7882 (2)	3	

									ASLHSNADAVEFSEY [P#]	828.3643 (2)	2
Endonucleases (END)											
100	Endonuclease domain-containing 1 protein	ETE56723	*Ophiophagus hannah*	18	9	6	525	53%	GHLNPNGHQPDYSAK	545.5948 (3)	2
									SSTFTLTNIVPQFIK	565.9838 (3)	3
									GCQQTFAVVGAVPGDTYIAR	704.0215 (3)	5
									LAQLYNVNHVSLFHSDCPR	568.2862 (4)	4
									LAQLYNVNHVSLFHSDCPRQ	600.3018 (4)	6
									AVVGAVPGDTY[#]	524.7672 (2)	4
									TLTNIVPQFIKL[#]	693.9181 (2)	5
									SAACCLIDNNHL[#]	694.3016 (2)	10
									SAACCLIDNNHLRSW[#]	606.2776 (3)	2
101	Endonuclease domain-containing 1 protein	JAB54304	*Micrurus fulvius*	32	4	1	176	14%	TLTNIVPQFIKL[#]	693.9181 (2)	5
									SAACCLIDNNHL[#]	694.3016 (2)	10
									SAACCLIDNNHLRSW[#]	606.2776 (3)	2
									ARNDQNVVQKF[#]	440.233 (3)	1
L- aminoacid oxidases (LAAO)											

102	L-amino-acid oxidase	A8QL58	*Naja atra*	51	47	47	2445	88%	VTLLEASER	509.2818 (2)	1	
									VIKIQYDAEK	402.8965 (3)	2	
									VIKIQYDAEKVR	487.954 (3)	3	
									TFVTADYVIVCSTSR	573.6209 (3)	4	
									IYFEPPLPPK	600.8379 (2)	2	
									TCADIVINDLSLIHDLPK	679.696 (3)	2	
									TCADIVINDLSLIHDLPKR	549.0504 (4)	4	
									EIQALCYPSIK	661.3424 (2)	4	
									EIQALCYPSIKK	483.9317 (3)	2	
									VLADDSDFF[#]	1028.4577 (1)	1	
									IQYDAEKVR	561.3017 (2)	1	
									TSNPKHVVVVGAGMAGLSAAYVLAGAGHK	553.3035 (5)	4	
									HVVVVGAGMAGLSAAYVLAGAGHK	563.5583 (4)	4	
									HVVVVGAGMAGLSAAYVLAGAGHKVTLLEASER	809.1965 (4)	6	
									EGWYVNMGPMR	670.3011 (2)	2	
									LNEFFQENENAWYYINNIR	1239.0828 (2)	3	

									LNEFFQENENAWYYINNIRK	869.0918 (3)	5	
									VWEVKKDPSLLK	361.2174 (4)	4	
									KDPSLLKYPVKPSEEGK	479.5207 (4)	4	
									KVIEELKR	338.8827 (3)	3	
									YDSYSTKEYLIK	503.9226 (3)	4	
									SDALFSYEK	530.2547 (2)	3	
									SDALFSYEKR	405.8724 (3)	2	
									VRVTYQTPAK	388.2247 (3)	1	
									VTYQTPAKTFVTADYVIVCSTSR	869.7785 (3)	3	
									RIYFEPPLPPK	452.9286 (3)	3	
									RIYFEPPLPPKK	371.9726 (4)	4	
									IYFEPPLPPKK	443.5934 (3)	3	
									IFLTCSKK	332.8576 (3)	3	
									KFWEADGIHGGK	448.8964 (3)	4	
									FWEADGIHGGK	406.1977 (3)	3	
									STTDLPSR	438.7257 (2)	2	

									REIQALCYPSIK	493.2672 (3)	2	
									REIQALCYPSIKK	402.227 (4)	5	
									DAEKVRVTY#	540.7887 (2)	2	
									WEADGIHGGKSTTDLPSRF#	519.2571 (4)	4	
									YPVKPSEEGK	567.297 (2)	3	
									VTYQTPAK	454.2491 (2)	3	
									VLAGAGHKVTLL#	589.8667 (2)	1	
									AGAGHKVTLL#	483.7905 (2)	1	
									SRGAVDMIGDLLNEDSSY#	979.4456 (2)	2	
									EKRFDEIVGGFDQLPISMY#	748.7078 (3)	4	
									EADGIHGGKSTTDLPSRF#	629.9805 (3)	4	
									DTKTCADIVINDLSLIHDLPKREIQAL#	773.6705 (4)	5	
									NLDKYTMGSITSF#	738.8577 (2)	3	
									GWYVNMGPMRLPE^	775.3707 (2)	3	
									VKKDPSLLKYPVKPSEE^	490.0318 (4)	2	
103	L-amino-acid	A8QL51	*Bungarus multicinc*	59	3	1	142	2%	TMGALTSFTPY#	1204.5557 (1)	1	

	oxidase		*tus*						VLADDSDFF#	1028.4578 (1)	1	
									EADGIHGGKSTTDLPSRF#	629.9805 (3)	4	
104	L-amino-acid oxidase	P0DI91	*Naja oxiana*	11	12	3	688	82%	RSPLEECFQQNDYEEFLEIAR	891.75 (3)	4	
									SPLEECFQQNDYEEFLEIAR	839.7163 (3)	4	
									SDALFSYEK	530.2547 (2)	3	
									KFWEADGIHGGK	448.8964 (3)	4	
									FWEADGIHGGK	406.1977 (3)	3	
									REIQALCYPSIK	493.2672 (3)	2	
									REIQALCYPSIKK	402.227 (4)	5	
									EIQALCYPSIK	661.3424 (2)	4	
									EIQALCYPSIKK	483.9317 (3)	2	
									HLSFIESL#	473.26 (2)	9	
									DDRRSPLEECFQQNDYEEFLEIAR	765.5999 (4)	1	
									CFQQNDYEEFLE^	811.3284 (2)	1	
105	L-amino acid oxidase precursor	AAY89682	*Pseudechis australis*	59	3	1	110	6%	SASQLYRESLQK	1408.7372 (2)	4	
									FWEADGIHGGK	1215.5713 (3)	3	

									STTDLPSR	438.7257 (2)	2
106	L-amino acid oxidase, partial	JAC94991	*Opheodrys aestivus*	40	3	1	123	8%	QAIAETVHL#	491.2726 (2)	2
									WEADGIHGGKSTTDLPSRF#	519.2571 (4)	4
									EADGIHGGKSTTDLPSRF#	472.737 (4)	3
107	L-Amino Acid Oxidase	3KVE_A	*Vipera ammodytes ammodytes*	55	2	2	105	7%	HVVVVGAGMSGLSAAYVLAGAGHK	563.5583 (4)	4
									STTDLPSR	438.7257 (2)	2
108	Secreted L-amino acid oxidase precursor	ACF70483	*Daboia russelii*	57	1	1	61	3%	NLLLETADYVIVCTTSR	656.6753 (3)	1
109	L-amino acid oxidase precursor	AAY89680.1	*Oxyuranus scutellatus*	59	5	1	239	15%	NPLEECFREADYEEFLEIAR	844.3866 (3)	2
									HVVVVGAGMAGLSAAYVLAGAGHK	745.7428 (3)	3
									VTLLEASER	509.2811 (2)	3
									IYFEPPLPPK	600.8363 (2)	1
									KFWEADGIHGGK	448.8939 (2)	2
Gluthathione peroxidase (GPrx)											
110	Glutathione peroxidase 3, partial	ETE68810	*Ophiophagus hannah*	30	20	20	943	77%	LVILGFPCNQFGK	746.9027 (2)	1
									QEPGQNSEILQGIK	770.90	2

										36 (2)		
									GDVNGENEQKIYTFLK	618.9817 (3)	4	
									NSCPPVVETFGDTAK	811.3824 (2)	3	
									AKVDCYDSVK	592.7864 (2)	3	
									TDRLVILGFPCNQFGK	622.3319 (3)	4	
									LVILGFPCNQFGKQEPGQNSEILQGIK	1005.532 (3)	1	
									FLVNPQGKPVMR	462.5971 (3)	3	
									TNVSTVKNDIIR	453.9265 (3)	4	
									GKQEPGQNSEIL#	650.3313 (2)	1	
									QKGDVNGENEQKIY#	541.5946 (3)	2	
									TFLKNSCPPVVETF#	819.9152 (2)	2	
									LKNSCPPVVETF#	695.8555 (2)	1	
									KNSCPPVVETF#	639.3131 (2)	1	
									KNSCPPVVETFGDTAKL#	621.6494 (3)	1	
									KNSCPPVVETFGDTAKLF#	670.6727 (3)	2	
									VDCYDSVK	493.2219 (2)	7	
									LVNPQGKPVMRW#	475.60	2	

										01 (3)		
									QKIYTFLKNSCPPVVE^	641.6744 (3)	2	
									TFGDTAKLFWSPLKIHDIKWNFE^	559.4953 (5)	3	
111	Glutathione peroxidase 3	XP_015679695	*Protobothrops mucrosquamatus*	28	3	2	143	9%	ILFVNVATY#	1039.5821 (1)	1	
									GKQEPGQNSEIL#	650.3313 (2)	1	
									LVNPQGKPVMRW#	475.6001 (3)	2	
Acetylcholinesterase (ACE)												
112	ACN-Sut-1, partial	JAA74986	*Suta fasciata*	63	3	3	145	6%	SLPVLDGHVSAF#	621.3313 (2)	2	
									GESAGAASVGMHLL#	650.3216 (2)	1	
									TASRPQYVQL#	581.8148 (2)	1	
113	Acetylcholinesterase	XP_007444166	*Python bivittatus*	67	2	1	106	5%	GVPHANDIATEAVVL#	753.407 (2)	7	
									GESAGAASVGMHLL#	650.3216 (2)	1	
114	Acetylcholinesterase	XP_015668657.1	*Protobothrops mucrosquamatus*	65	2	2	94	3%	VATQTGWVR	509.2783 (2)	2	
									VYAYLFDHR	395.2006 (3)	2	
115	Acetylcholinesterase	4QWW_A	*Bungarus fasciatus*	60	2	2	171	8%	VGAFGFLGLPGSPEAPGNMGLLDQR	834.0948 (3)	2	
									LLNATVDPPR	548.3103 (2)	2	