Journal of Proteomics 133 (2016) 20–32

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Deep venomics of the *Pseudonaja* genus reveals inter- and intra-specific variation

Timothy Reeks [a], Vincent Lavergne [a], Kartik Sunagar [b], Alun Jones [a], Eivind Undheim [a], Nathan Dunstan [c], Bryan Fry [a,d], Paul F. Alewood [a,*]

[a] Institute for Molecular Bioscience, The University of Queensland, QLD 4072, Australia
[b] Department of Ecology, Evolution and Behavior The Alexander Silberman Institute for Life Sciences, Hebrew University of Jerusalem, Israel
[c] Venom Supplies Pty Ltd., Australia
[d] Venom Evolution Lab, School of Biological Sciences, The University of Queensland, QLD 4072, Australia

ARTICLE INFO

Article history:
Received 7 July 2015
Received in revised form 14 November 2015
Accepted 24 November 2015
Available online 26 November 2015

Keywords:
Venomics
Topic:
Venom gland transcriptome
Proteome
Snake toxins
Glycopeptides
Pseudonaja genus

ABSTRACT

Australian elapid venom remains an under-investigated resource of novel bioactive peptides. In this study, the venom gland transcriptomes and proteomes of the Australian western brown snakes, *Pseudonaja aspidorhyncha* and *Pseudonaja nuchalis*, were compared to *Pseudonaja textilis*. A deep venomics strategy incorporating high throughput 454 pyrosequencing gave a total of 200,911 raw reads for the three venoms. Subsequent annotation identified 5716 transcripts from 20 different toxin families with inter-specific variation between species observed in eight of the less abundant families. Integration of each venom proteome with the corresponding annotated reads identified 65 isoforms from six toxin families; high sequence coverage highlighted subtle differences between sequences and intra and inter-specific variation between species. High quality MS/MS data identified unusual glycoforms with natriuretic peptides from *P. aspidorhyncha* and *P. nuchalis*containing *O*-linked trisaccharides with high homology to the glycosylated region of TNPc. Molecular evolutionary assessments indicated the accelerated evolution of all toxin families with the exception of both natriuretic peptides and *P. aspidorhyncha* PLA2s that were found to be evolutionarily constrained under purifying selection pressures. This study has revealed a wide range of novel peptide sequences from six bioactive peptide families and highlights the subtle differences between toxins in these closely related species.

Biological significance: Mining Australia's vastly untapped source of toxins from its venomous creatures has been significantly advanced by employing deep venomics methodology. Technological advances in transcriptome analysis using next generation sequencing platforms and proteome analysis by highly sensitive tandem mass spectrometry allowed a more comprehensive interrogation of three underinvestigated brown snake (*Pseudonaja*) venoms uncovering many novel peptide sequences that are unique to these closely related species. This generic strategy will provide invaluable information when applied to other venomous snakes for a deeper understanding of venom composition, envenomation, venom evolution, as well as identifying research tools and drug leads.

1. Introduction

Snake venoms have evolved over millions of years to rapidly induce profound physiological changes during prey envenomation that results in death and preliminary digestion of the target animal. Elapid venoms contain numerous protein and peptide families that act as enzymes or ligands of membrane bound receptors with agonist or inhibitory mechanisms of action [1]. A number of families, such as the phospholipase A_2 (PLA_2) enzymes, and three finger toxins (3FTxs) exhibit conserved structural frameworks that present a broad range of pharmacological effects [2]. Diversity of function within each protein family is a key attribute of snake venom [3]. The proteins normally contain a characteristic, preserved, disulfide rich scaffold that provides high stability, as well as resistance against enzymatic degradation within the venom gland and the prey tissue. These frameworks present amino acid residues in a precise orientation at the surface of the protein, promote the effect of mutations of key residues and provide subtle differences in sequence and conformation. Overall, this translates into a myriad of toxins with a diverse range of biological effects [2]. Moreover, the formation of polymeric (mostly covalently linked) toxin complexes often enhances the binding efficiency toward specific molecular targets compared to their monomeric forms and leads to new pharmacologies [4]. Not only is this variation a result of toxin gene duplication followed by neofunctionalization, but can also be a function of phylogeny, ecology

* Corresponding author.
E-mail address: p.alewood@imb.uq.edu.au (P.F. Alewood).

http://dx.doi.org/10.1016/j.jprot.2015.11.019
1874-3919/

and/or geography [5]. Such diversity often leads to both inter- and intra-specific snake venom variation [6].

The common brown snake (*Pseudonaja textilis*) has earned the reputation of being one of the most dangerous snakes in Australia not only because of its lethal venom, but because of its aggressive nature and presence in both urban and country areas of the eastern and central parts of the country [7]. The main component of *P. textilis* venom is a protein complex (Pseutarin C) comprised of an enzymatic (mutated blood factor Xa) and non-enzymatic subunits (blood factor Va) [8]. Upon envenomation, this toxin causes venom-induced consumptive coagulopathy (VICC) resulting in hemorrhage, potentially leading to stroke and cardiac arrest [9,10]. *P. textilis* venom is also characterized by neurotoxic proteins and peptides, such as the Type II and Type III 3FTxs as well as the PLA_2 family [11]. Additionally, complexes of PLA_2 toxins, such as textilotoxin, are capable of producing highly neurotoxic effects [4]. The complexity of *P. textilis* venom is further enhanced by the presence of post-translational modifications, such as glycosylation, γ-carboxylation and phosphorylation [11]. Apart from its neurotoxic components, the venom contains additional toxin families, including nerve growth factors (NGFs), cysteine-rich secretory proteins (CRVPs), kunitz type protease inhibitors and natriuretic peptides [11,12,13].

Compared to Asian and American snakes, Australian terrestrial elapid venom remains largely unexplored. Research has predominantly focused on the 'medically significant' genera (*Acanthophis*; *Notechis*; *Pseudechis*; *Pseudonaja*; and *Oxyuranus*), which are considered to be the most dangerous to humans. However, only one or two species of these taxa (e.g. *Pseudonaja*) have been studied in any depth [14]. Furthermore, investigation of Australian elapid venom has mainly focused on the isolation and characterization of the most toxic or abundant protein components that are readily observed by 2D gel chromatography [1].

One clade, the *Pseudonaja nuchalis* species (generically referred as the western brown snake), was recently split into three lineages based on their karyotypes, mitochondrial DNA sequences, and multivariate morphometric analysis [15,16]. Skinner et al. showed that these lineages comprise *Pseudonaja aspidorhyncha* (distributed in southern central Australia), *Pseudonaja mengdeni* (western and central Australia), and *P. nuchalis* (tropical Northern Territory and northern Queensland) [16]. All three species are potentially as dangerous as *P. textilis* with at least three fatalities connected to the 'western brown' in recent years [17]. To our knowledge, based on the locality data provided in most of the research conducted to date, *P. aspidorhyncha* has predominantly been investigated amongst the '*P. nuchalis*'-complex.

Unraveling the complexity of snake venoms has largely been dependent upon advances in separation technologies combined with modern proteomic analysis [2]. Nevertheless, mass fingerprinting protein identification still remains limited by reliance on universal protein databases that only provide known annotated protein sequences. This bottleneck is being removed by utilization of next generation transcriptome sequencing platforms (NGS), which allow in-depth analysis of messenger RNA (mRNA) transcripts encoding polypeptide toxins in the venom-producing cells of the venom gland. Integrated venomics techniques are now matching high quality proteome mass spectrometry data to toxin-encoding transcripts isolated from the same tissue thus increasing the likelihood of discovery, identification and sequencing of novel toxins and their isoforms [18]. In the present study, we report the analysis of the venom gland transcriptomes of *P. aspidorhyncha*, *P. nuchalis* and *P. textilis* and integrate high quality MS/MS proteomic data from their crude venoms to delineate full length translated peptides and proteins (Fig. 1). This has led to the discovery of numerous novel PLA_2 sequences, three finger toxins, natriuretic peptides and kunitz type serine protease inhibitors that are secreted as the major toxin families within the three brown snake species. The high degree of fidelity between isoforms enabled the identification of intra- and inter-specific variation between these species for the first time. Although this strategy focused on peptide and small proteins less than 16 kDa, fragments from larger toxins such as prothrombin activators were also identified.

2. Material & methods

2.1. Venom and venom gland samples

Venom glands and associated milked venoms from individual adult specimens of *P. textilis* (Queensland population), *P. aspidorhyncha* and *P. nuchalis* were obtained from Venom Supplies Pty. Ltd. Snakes were euthanized 4 days after their last milking and the venom gland dissected and stored in RNAlater™ solution (Qiagen) at −20 °C.

2.2. mRNA isolation and Roche 454 sequencing

Each tissue was first processed using a rotor-stator homogenizer, and total RNA was extracted using TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Messenger RNA was isolated using the RNeasy mRNA mini kit (Qiagen) protocol and quantified using a nanodrop spectrophotometer. Samples of 200 ng of mRNA were submitted to the Australian Genomic research Facility (AGRF) for cDNA library construction and sequencing using a Roche GS FLX Titanium platform.

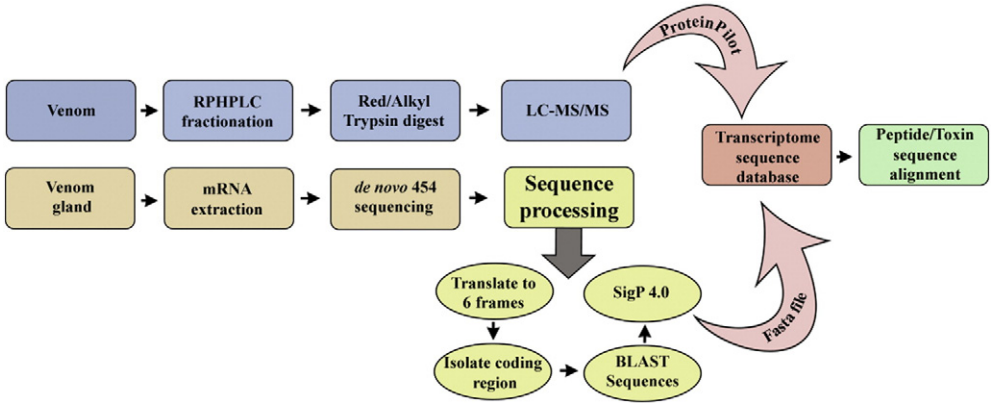

Fig. 1. Flow diagram of the transcriptome and proteome analysis procedure. Venom gland of each *Pseudonaja* species was homogenized and enriched for mRNA. The cDNA library was constructed using 454 sequencing on a Roche GS FLX Titanium platform with a ¼ plate. An in-house stepwise algorithm was used to identify Blastp and SignalP 4.0 hits from the raw reads. Crude venom of each species was fractionated using RPHPLC then reduced/alkylated and trypsin digested for analysis by ESI LC-MS/MS. The high quality spectral data was matched to the transcriptomes sequences using ABSCIEX *ProteinPilot 4.0* software.

2.3. RP-HPLC fractionation

Crude venom from each species (1.0 mg) was dissolved in 0.5% trifluoroacetic acid and fractionated using a Phenomenex C18 (4.6 mm × 250 mm 5 μm 300 Å) column. The column was fitted to a Shimadzu LC-20 AT HPLC system and UV absorbance monitored at 214 nm and 280 nm. A linear gradient of 0.5% min^{-1} at flow rate 1.0 mL min^{-1} was utilized with 0.043% trifluoroacetic acid/90% acetonitrile as elution buffer B and 0.05% trifluoroacetic acid as buffer A. Fractions were collected across the entire elution range based on eluent absorption at 214 and 280 nm; freeze dried and stored at −20 °C.

2.4. Reduction/alkylation and tryptic digestion

Lyophilized fractions were resuspended in 100 μL ultrapure water. A volume of 45 μL was removed from each HPLC fraction and 5 μL of 1.0 M ammonium carbonate (pH 11) was added. A cocktail of 50 μL solution containing 97.5% acetonitrile, 2% iodoethanol, and 0.5% triethylphosphine (*v*/*v*/v) was used as the reduction/alkylation buffer. Samples were incubated for 120 min at 37 °C and subsequently dried in a vacuum concentrator. A 100 μL formic acid solution (0.1%) was added to each fraction, from which 50 μL were removed, and freeze-dried prior to trypsin digestion. Each fraction was reconstituted in 100 mM ammonium carbonate (pH 8.5), before addition of trypsin (1.0 mg/mL in 1.0 mM HCl) in a protease:peptide ratio of 1:20. Samples were then incubated overnight at 37 °C.

2.5. uHPLC and nano electrospray ionization HPLC–ESI–MS–MS/MS

P. textilis and *P. aspidorhyncha* fractions were reconstituted in 0.1% formic acid solution. The extracts were analyzed by uHPLC–MS/MS on a Shimadzu Nexera uHPLC (Japan) coupled to a Triple TOF 5600 mass spectrometer (ABSCIEX, Canada) equipped with a TurboSpray ion source. Aliquots (2 μL) of extracts were injected onto a 2.1 mm × 100 mm Zorbax SB 300 Å C18 1.8 μm column (Agilent) at 250 μL/min. For uHPLC analyses the ion-spray voltage was set to 5300 V, declustering potential at 100 V, curtain gas flow for both GS1 and GS2 set at 50, nebulizer gas 1 at 12 and GS2 heater at 500 °C.

P. nuchalis fractions were analyzed by nano HPLC–MS–MS/MS on a Shimadzu Prominance Nano HPLC (Japan) coupled to a Triple TOF 5600 (Hybrid Quadrupole-TOF) mass spectrometer (Framingham, MA, USA) equipped with a nano electrospray ion source. 2 μL of each sample was injected onto a 50 mm × 300 μm C18 trap column (Agilent) at 60 μL/min. The trap column was then placed in-line with the analytical Agilent nano HPLC column (150 mm × 100 μm (ID) C18, 3.5 μm 300 Å) for mass spectrometry analysis at 500 nl/min. For nano HPLC analyses the ion-spray voltage was set to 2000 V, declustering potential at 100 V, curtain gas flow at 25, nebulizer gas 1 at 12 and interface heater at 150 °C.

For both nano and uHPLC, solvent A consisted of 0.1% (*v*/*v*) formic acid and solvent B contained 90% (*v*/*v*) acetonitrile, 0.1% (*v*/*v*) formic acid in water. Linear gradients of 1–40% solvent B over 25 min were used for protein elution.

For both nano and uHPLC analyses the mass spectrometer acquired 250 ms full scan TOF-MS data followed by twenty 100 ms full scan product ion acquisitions in an Information Dependent Acquisition (IDA) mode. Full scan TOF-MS data was acquired over the mass range of 300–2000 Da and for product ion ms/ms 80–2000 Da. Ions observed in the TOF-MS scan exceeding a threshold of 150 counts and a charge state of +2 to +5 were set to trigger the acquisition of product ion, ms/ms spectra of the resultant 20 most intense ions. The data was acquired and processed using Analyst TF 1.5.1 software (Framingham, MA, USA). Interpretation of spectral data for *de novo* sequencing was performed manually, aided by both Analyst TF 1.5.1 and *ProteinPilot 4.0* software (Framingham, MA, USA).

2.6. Proteomic data analysis

Reconstruction of native LC–ESI–MS data was achieved by utilizing Analyst TF 1.5.1 LCMS and Bayesian peptide reconstruct BioTools (Framingham, MA, USA). The mass range was set between 1000 and 16,000 Da with mass tolerance set to 0.2 Da and S/N threshold set to 10. Manual inspection of reduced/alkylated and tryptic digest LC–ESI–MS and MS/MS data was carried out using the same parameters.

2.7. Bioinformatics

Each data set of reads was treated independently by a step-wise algorithm developed in-house. Initially, cDNA sequences were translated into 6 reading frames using the universal genetic code, and the coding regions were isolated in order to retrieve full-length precursor protein sequences. To annotate the transcriptome reference databases, these sequences were submitted to *BLASTp* (NCBI-BLAST-2.2.28+ package) for similarity searches, using the BLOSUM62 scoring matrix, against the last update of the reviewed UniProtKB-SwissProt database [19]. BLAST hits showing e-value similarity score < 10 have been sorted into a group named "blast+" (query sequences with score above this threshold or for which no similarity was found have been tagged as "blast−"). Each group of sequences was submitted to *SignalP 4.1c* program (using the default D value cutoff) to detect the presence of a potential N-terminal signal region, based on the recognition of consensus eukaryotic cleavage sites [20]. The 4 groups thus created ("blast + _signal+", "blast + _signal−", "blast−_signal+", and "blast−_signal−") have been matched to LC–ESI–MS/MS data of the respective snake species (as described above) using ABSCIEX *ProteinPilot 4.0*. Only LC–ESI–MS/MS spectral data with a matching confidence value ≥99% were used for subsequent analysis. Biological modifications, amino acid substitutions and post-translational modifications including amidation, deamidation, hydroxylation of proline and valine, oxidation of methionine, carboxylation of glutamic acid, cyclization of N-terminal glutamine (pyroglutamate), bromination of tryptophan, and sulfation of tyrosine and *N* linked and *O* linked glycosylation were predicted by ABSCIEX *ProteinPilot 4.0* software. Glycosylation (*O*- and *N*-linked) parameters in the paragon algorithm were manually set to 1% in anticipation of glycosylated proteins that have previously been discovered in *Pseudonaja* venom [11,1]. Finally, precursor protein sequences containing a signal region have been sorted according to the molecular function of their corresponding BLAST hit, then aligned with CLUSTALW (BLOSUM cost matrix, gap open and extension penalties of 10 and 0.1 respectively). In this study, only sequences showing a similarity score (e-value) $\leq 10^{-4}$ when compared to their respective BLAST hit were reported.

2.8. Selection analyses

Maximum-likelihood models [21] implemented in Codeml of the PAML package [22] were utilized to assess the influence of natural selection on various *Pseudonaja* toxin families examined in this study. Site-specific models, which estimate positive selection statistically as a non-synonymous (dN) to synonymous (dS) nucleotide-substitution rate ratio (ω) significantly greater than 1, were employed [23]. As no a priori expectation exists, we compared likelihood values for a pair of models with different assumed ω distributions: M7 (β) versus M8 (β and ω). The results are considered significant only if the alternative model (M8) that allow sites with $\omega > 1$ shows a better fit in Likelihood Ratio Test (LRT) relative to its null model (M7), which does not allow sites with $\omega > 1$. LRT is estimated as twice the difference in maximum likelihood values between nested models and compared with the χ^2 distribution with the appropriate degree of freedom — the difference in the number of parameters between the two models. The Bayes empirical Bayes (BEB) approach [24] was used for identifying sites experiencing positive selection by computing the posterior probabilities that a

particular site belongs to a given selection class (neutral, conserved or highly variable). Sites with greater posterior probability (PP ≥ 95%) of belonging to the 'ω > 1 class' were inferred as positively selected. FUBAR [25] implemented in HyPhy [26] was used to identify sites evolving under the pervasive influence of diversifying and purifying selection pressures. MEME [27] was used for detecting sites that experience episodic influence of diversifying selection. Sequence data sets were screened for recombination events using the Single Breakpoint Algorithm [28] of the HyPhy package [26]. When recombination breakpoints were detected, each recombinant was allowed to have independent phylogenies, prior to conducting selection analyses [28].

3. Results

3.1. Venom gland transcriptome analysis

The three venom gland transcriptomes of *Pseudonaja* species were sequenced using the Roche GS-FLX XLR70 titanium system. We obtained 60,674 (381.61 bp on average; N50 = 457), 83,944 (379.75 bp; N50 = 459) and 56,293 (381.84 bp; N50 = 455) reads for *P. textilis*, *P. aspidorhyncha* and *P. nuchalis*, respectively (Table 1). A high-throughput algorithm developed in-house was designed to isolate the coding region(s) of the raw read and to remove the duplicates (nucleic acid sequences lacking start and/or stop codon(s) were discarded in order to keep only full-length precursor protein sequences). The program identified BLASTp and SignalP hits from these coding sequences. This pipeline reduced 55.7% (33,769 — *P. textilis*), 38.6% (32,419 — *P. aspidorhyncha*) and 40.5% (22,783 — *P. nuchalis*) sequences of the initial number of cDNA reads, leading to a total of 5716 reads identified as toxins (2231–19% of the total reads for *P. textilis*; 3051 — 21% of the total reads for *P. aspidorhyncha*; 3115 — 29% of the total reads for *P. nuchalis*) (Fig. 2), with an average sequence length of 251.66 bp (252.03 bp/N50 = 258 for *P. textilis*; 249.04 bp/N50 = 255 for *P. aspidorhyncha*; 253.91 bp/N50 = 261 for *P. nuchalis*) (Table 1). We also successfully identified housekeeping protein transcripts that made up 56% (*P. textilis*), 52% (*P. aspidorhyncha*) and 47% (*P. nuchalis*) of the total processed reads (Fig. 2). Only 25% (*P. textilis*), 27% (*P. aspidorhyncha*) and 24% (*P. nuchalis*) of the reads could not be annotated (Fig. 2). The N50 and N90 values for the raw reads and processed toxins of the three transcriptomes (Table 1) as well as average Phred + 33 quality scores of >30 [29], indicated that the processed transcripts could be utilized for small peptide and protein identification from each venom proteome. For cysteine-rich secreted proteins (CRVPs), reads were manually assembled to the complete precursor with both reads containing methionine as the first residue.

A closer inspection of the annotated toxin transcripts isolated from these three *Pseudonaja* species allowed us to classify them into 20 known families of snake venom proteins with some precursors markedly more abundant than others (Fig. 2). As expected, *P. textilis* typically showed high expression of toxin transcripts involved in the biosynthesis of coagulation factors, PLA_2s, 3FTxs, natriuretic peptides, kunitz type serine protease inhibitors and CRVPs [11]. Both *P. aspidorhyncha* and *P. nuchalis* showed similar higher levels of toxin expression for these families compared to *P. textilis*, especially coagulation factors, PLA_2s and venom nerve growth factors whereas low levels of CRVPs were observed for *P. nuchalis*; and high proportions of Type II and III α-neurotoxic 3FTxs were observed for *P. aspidorhyncha* (Fig. 2). However, inter-specific variation was evident in the low abundant transcriptome families of vespryn, CVF-like, calglandulin, calmodulin, acetylcholinesterase, cystatin and waprins as well as the previously mentioned CRVPs (Fig. 2). For example, transcripts from vespryn, CVF-like, cystatin and waprins were observed in *P. textilis* only (Fig. 2). Interestingly, there was no evidence of transcripts for PLA_2 inhibitors previously reported for *P. nuchalis* [11].

Transcript sequences not identified by the BLASTp search ("unidentified sequences") were also compared between species. Of the 9508 cysteine rich and linear sequences, 4.82% contained a signal region. Identical sequences present in all three transcriptomes constituted 1.92% of these transcripts, while sequences with an exact match within 2 venoms made up 4.96% of total unidentified sequences. A number of non-annotated sequences were represented by a high number of reads and the exact cysteine rich and linear sequences were identified in either two or three venom transcriptomes (Table S1). While the *SignalP 4.1c* program did not identify a signal peptide in these sequences, they may suggest a new type of housekeeping protein or toxin.

3.2. Proteome analysis of Pseudonaja venoms

Venoms used for mass spectrometric analysis were milked from the same specimens that the venom gland transcriptome data was obtained. A comprehensive analysis of the crude venom was undertaken commencing with RP-HPLC fractionation. Each of the fractions was reduced, alkylated, trypsin digested and subsequently analyzed using LC–MS/MS. The chromatography profiles were similar between venoms with multiple peaks dominated by 7–8 large fractions (Fig. S1). High quality MS/MS spectra were submitted to *ProteinPilot 4.0* software to identify proteins and polypeptides using the corresponding venom transcriptomes as the reference database. Only peptides showing ≥99% matching confidence from the *ProteinPilot* analysis were used. The sequences were then aligned with known proteins using ClustalW. All matched MS/MS data were manually checked (Table S3).

Proteomic data was matched to the corresponding transcriptomes and placed in toxin superfamilies before being clustered and aligned with each other to highlight sequence similarities. We identified 65 isoforms of PLA_2s, 3FTxs, Kunitz type protease inhibitors, nerve growth factors, cysteine-rich secretory proteins and natriuretic peptides in all three venoms (Table S3). An additional eight toxin sequences were identical to those reported in the literature. From *P. textilis* venom, 22 (1.0%) toxin isoforms (including those previously identified in the literature) were matched to the 2231 toxin transcripts; from *P. aspidorhyncha* venom, 27 (0.9%) complete isoforms were matched to the 3051 toxin transcripts and from *P. nuchalis* venom, 23 (0.7%) complete isoforms were matched to the 3115 toxin transcripts. From the four toxin groups identified, sequence coverage ranged from 12.8% in acidic PLA_2 to 100% coverage in 3FTxs, kunitz type protease inhibitors and natriuretic peptides (Table 2). Moreover, apart from basic PLA_2 sequences, a total of 48 isoforms showed greater than 80% sequence coverage are described (Table 2). In addition, coagulation factors Xa and Va, as well as CRVPs, were identified in all three venoms, while venom nerve growth factor was only identified in the proteome of *P. nuchalis*.

3.3. Three finger toxins

Members belonging to this toxin superfamily are typically structured with three distinct β-stranded loops extending from a disulfide rich hydrophobic core [30]. α-Neurotoxic 3FTxs may be of Types I and III (short-chain) composed of 60–62 residues with 4 disulfides or Type

Table 1
Comparison of the raw reads and processed toxins for each *Pseudonaja* species.

Species	# Reads	Length interval	Average length	N50	N90	%GC	%N
Raw reads							
P. textilis	60,674	25–736	381.61	457	519	47.3744	0.0623
P. aspidorhyncha	83,944	26–778	379.75	459	520	47.0343	0.0555
P. nuchalis	56,293	29–849	381.84	455	511	46.2961	0.0522
Processed toxins							
P. textilis	2231	153–477	252.03	258	390	45.1228	0
P. aspidorhyncha	3051	153–471	249.04	255	387	46.5514	0
P. nuchalis	3115	153–474	253.91	261	393	46.0003	0

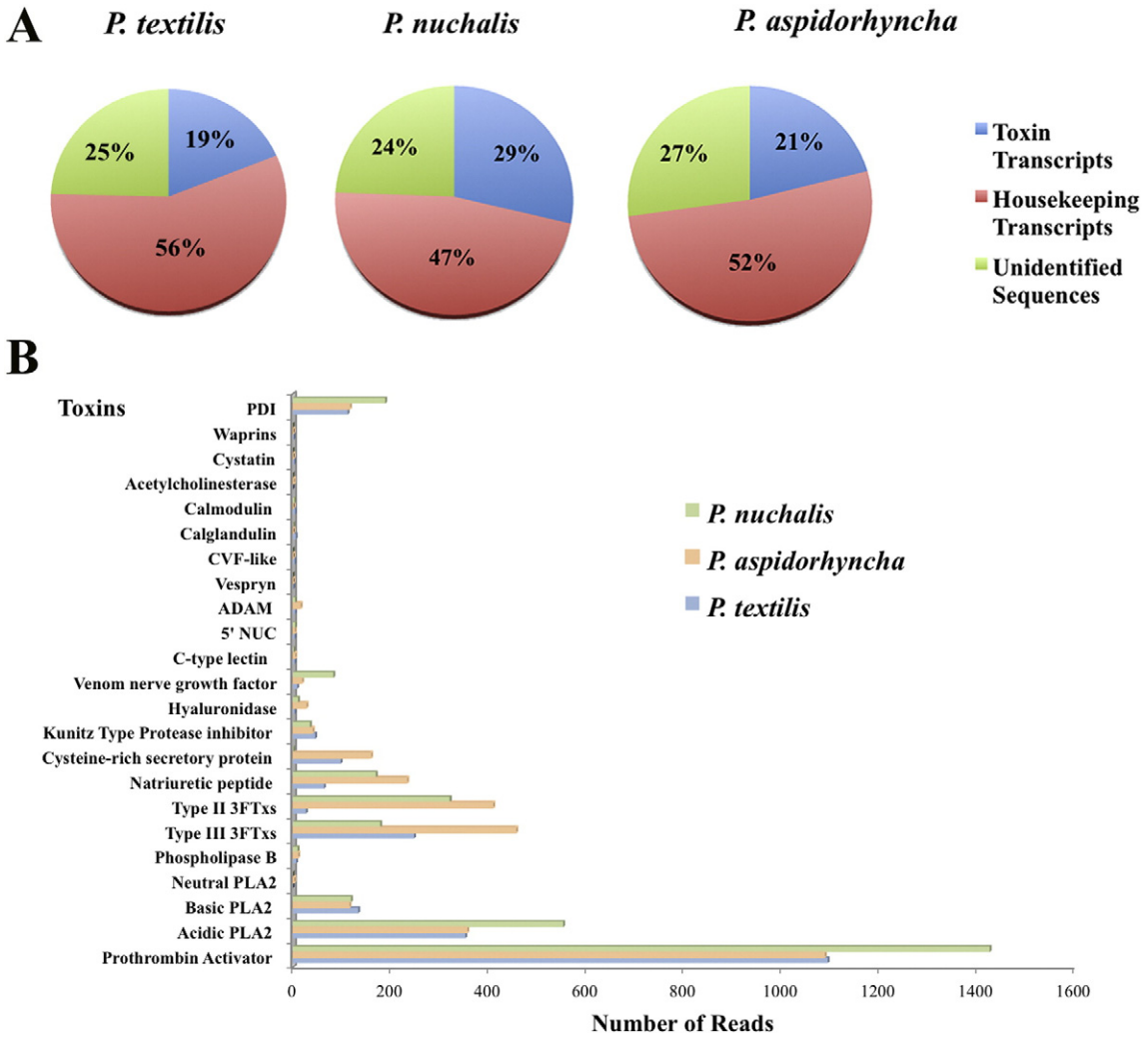

Fig. 2. Transcriptome profiles of *P. nuchalis*, *P. aspidorhyncha* and *P. textilis*. A) Processed transcript reads were submitted to *Blastp* from the NCBI-BLAST-2.2.28+ package for similarity search of toxin and non-toxin proteins (house-keeping proteins) against the UniProtKB-SwissProt database then analyzed for the presence of a signal peptide using *SignalP 4.1c*. 'Unidentified sequences' could not be confidently aligned by *Blastp*. B) Toxin transcriptome profiles of *P. nuchalis*, *P. aspidorhyncha* and *P. textilis*. Transcripts identified as toxins were split into 20 families and relative abundances expressed as number of reads. PLA2, phospholipase A2 enzyme; 3FTx, three finger toxin; 5'NUC, 5′-nucleotidase; ADAM, a disintegrin and metalloproteinase; CVF-like, cobra venom factor-like; PDI, protein disulfide isomerase.

II (long-chain) α-neurotoxin 3FTx composed of 66–74 residues and 5 disulfide bonds. While Type III 3FTx are poorly characterized, both Type I and Type II 3FTx are well-characterized and antagonize neuromuscular nicotinic acetylcholine receptors (α1 nAChRs) whereas only Type II block neuronal nicotinic acetylcholine receptors (α7 nAChRs) [31].

Proteomic analysis allowed us to identify Type III isoforms containing 4 disulfide bonds. The alignment of the identified Type III 3FTxs, compared with known homologs, showed variation in the exposed loops between the three species, but also intra-specific variation in *P. textilis* (Fig. S2). Generally, the isoforms could be segregated into two clusters; one cluster was closely aligned to 3FTx-Oxy6 (UniProt accession number A7X4T2) and the other aligned to *P. textilis* Type III SNTX8 (UniProt accession number A8HDK1) (Fig. S2). The isoforms aligned to 3FTx-Oxy6 contained 81 residues with at least 85% similarity between sequences (Fig. S2). However, greater variation was observed between the isoforms aligned to *P. textilis* Type III SNTX8 (UniProt accession number A8HDK1) with similarity between sequences as low as 55%. All sequences contained the customary 8 cysteine residues that form the 4 disulfide bonds apart from three sequences (one from each species) that contained nine cysteine residues (Fig. S2). Interestingly, the additional cysteine residue is located at the same position of the first cysteine of the loop 2 disulfide observed in Type II 3FTxs. 3FTxs with odd number of cysteines have traditionally been associated with complex formation (e.g. irditoxin [32]) and not linked to receptor binding through a covalent bond. A number of type III postsynaptic neurotoxins from *P. textilis* were partially characterized showing they antagonize muscle nAChRs in mice though with reduced binding constants compared to other 3FTxs [33]. The functional residues from Type III neurotoxins are yet to be identified.

The Type II 3FTx toxins identified in the three venom proteomes generally showed sequences that contain five disulfide bonds (Fig. S3). However, three sequences displayed nine cysteine residues,

Table 2
Sequence coverage of toxin families.

Toxin group	Min % protein sequence coverage	Max % protein sequence coverage	No. of protein sequences	No. of seq >80% coverage
Type III 3FTx	82.7	100	15	15
Type II 3FTx	35.8	100	13	11
Kunitz type	40.7	100	13	10
Natriuretic peptide	30.0	100	5	3
Acidic PLA_2	12.8	96.6	15	9
Basic PLA_2	61.0	72.8	7	0
CRiSP	51.4	72.9	2	0

including LNTX-2 from *P. textilis* with a truncated C-terminus. Sequences from *P. textilis* and *P. nuchalis* are homologous to *P. textilis* LNTX 1 (UniProt accession number A8HDK6), thereby lacking the characteristic disulfide in the 2nd loop, a hallmark of some *Pseudonaja* Type II 3FTxs [34]. The fifth disulfide in the 2nd loop, that sustains α7 nAChR binding activity, is present in the remaining sequences. Moreover, functional residues of Australian elapid Type II neurotoxins (UniProt accession number A8HDK6 (*P. textilis*) and Q45Z11 (*O.s. scutellatus*)) have been identified between residues Trp47 and Arg58 (Trp47, Asp49, Phe51, Arg55, Arg58) of loop 2 of the precursor protein [34]. Trp47 and Arg55 are common to all Type II isoforms in this study with variation observed at residue positions 49, 51 and 58 (Fig. S3).

Selection analyses of three-finger toxins highlighted the rapid evolution of this toxin type within *Pseudonaja* spp. Computed ω values were in the range of 1.56 to 5.81 for Type II and Type III toxins from various *Pseudonaja* spp. (Table S2), exemplifying the rapid evolution of these toxins [30]. Further, several positively selected sites (ranged between 4 and 32) were identified in these toxins types (Table S2). A significant influence of episodic diversifying selection was also identified (Table S2).

3.4. Kunitz peptides

Members of this ubiquitous protein family are between 57 and 60 amino acids long and characterized by the presence of three disulfide bonds in a specific cysteine rich motif [35]. Kunitz inhibitors bind to the active sites of serine proteases in a substrate-like manner, through an exposed binding loop that mimics the enzyme substrate [36]. Generally, a positively charged P1 residue (arginine or lysine) denotes trypsin inhibition [37]. Where P1 is a tyrosine, leucine or phenylalanine residue (i.e. large hydrophobic) chymotrypsin inhibition is observed, whereas valine or isoleucine in position P1 (i.e. small hydrophobic residues) enables elastase inhibition [37]. Seven kunitz type serine protease inhibitor isoforms have previously been identified in *P. textilis* venom that show variation in the P3–P3′ motif [38]. Textilinin-1 and -2 showed plasmin inhibition with both peptides containing an arginine residue at P1 while textilinin-3 (P1-asparagine) showed no inhibition [39]. Overall, the protease inhibitor isoforms identified here from the *Pseudonaja* venoms show a high degree of similarity to previously discovered textilinins from *P. textilis*, as well as a high conservation rate of their cysteine residues [40]. The sequences identified in this study share at least 64% identity. Comparison of the signal peptide regions showed significant variations between isoforms with two isoforms exhibiting an extended sequence that were not identified as signal peptides by SigP (Fig. 3). Inter-specific variation is evident in the P3–P3′ domain as well as the hairpin turn between beta sheets (residues 66–70) of all isoforms (Fig. 3). Textilinin-1,-2,-3 and -4 from a previous report [38] were identified in this study, as well as two more (textilinin -8 -9; both P1 is aspartate) isoforms highlighting intra-specific variation within the species. Crystal structures of textilinin-1 with microplasmin showed that the unique P3–P3′ motif (Pro15–Cys16–Arg17–Val18–Arg19–Phe20) was involved in the inhibition activity with the side chain of Arg17 fitting in the S1 pocket from the S3–S3′ sites of microplasmin [41]. An alignment of the kunitz peptides isolated in this study with textilinin-1 demonstrates that aspidorynchin-1 (*P. aspidorhyncha*) has the same P3–P3′ motif but shows variation in the primary structure (Fig. 3). Interestingly, nuchalinin-1 (*P. nuchalis*) has a similar P3–P3′ motif as textilinin-1 except for lysine replacing Arg-19 (Fig. 3). As textilinin-1 is an experimental drug (Q8008, Venomics Pty Ltd.) being trialed as an aprotinin replacement, these isoforms may provide drug leads as novel anti-bleeding agent for patients undergoing surgical procedures [42].

Kunitz peptides were also found to be rapidly evolving with the computed ω in the range of 1.54–6.81 (Table S2). While the BEB approach identified 9–16 sites as positively selected within this toxin type, MEME identified several toxin types as experiencing episodically adaptive selection (Table S2).

3.5. Natriuretic peptides

Natriuretic peptides found in snake venom are structurally similar to those found in mammals (ANP, BNP, CNP) and contain a 17 membered ring stabilized by one disulfide bond [43]. They are synthesized as pre-pro-hormones and undergo proteolytic cleavage of the propeptide by serine proteases to produce the mature natriuretic peptide [44,45]. The effects of these peptides are widespread and include regulation of blood volume, blood pressure, ventricular hypertrophy, pulmonary hypertension, fat metabolism and long bone growth in response to a number of pathological conditions [46]. In this study, the natriuretic peptide transcripts identified from the *Pseudonaja* proteomes also exhibit the typical 17 membered ring (Fig. 4). These isoforms can be segregated into two clusters that differ by a 3-residue (Lys-Val-Ala) insertion at the N-terminus of the mature peptide (Fig. 4). Residues Phe8, Arg14 and the C-terminus of ANP are important for binding and inducing natriuresis and diuresis effects [47]. One of the identified clusters contains the functionally important arginine residue while the other cluster has this residue replaced by histidine. Similar sequences have shown that the substitution of arginine by histidine was a contributing factor to the decrease of cGMP concentration in HEK293 cells overexpressing rat NPR-A [48]. Here we observe that the conserved functional Phe residue, Phe10/13, is present in all sequences. Additionally, fragments of the propeptide region of the *P. nuchalis* natriuretic peptide, PnNPa, were expressed in the proteome (Fig. 4). Natriuretic peptide precursors from viper venoms are proteolytically cleaved into natriuretic peptides and proline rich bradykinin potentiating peptides that have been shown to inhibit angiotensin converting enzyme (ACE) [49]. Likewise, propeptides isolated from *Psammophis mossambicus* snake venom metalloproteases (SVMP) were shown to inhibit mammalian α7 neuronal nicotinic acetylcholine receptors [50]. Hence, while the natriuretic peptide propeptides isolated in this study do not show homology to the propeptides from viper natriuretic peptide and SVMP precursors described, they may represent potential novel pharmacology.

From the MS/MS data, the natriuretic peptide from *P. aspidorhyncha* and *P. nuchalis* contains glycosylated residues. Marker ions corresponding to hexose-N-acetyl hexosamine (204 Da; 366 Da) and sialic acid (292 Da; dehydrated 274 Da) were observed on the native and tryptic MS/MS spectra containing the proline rich C-terminal sequence (Fig. S6). The marker ions were observed for *P. aspidorhyncha* and *P. nuchalis* (PasNPa, PasNPb, PnNPa) natriuretic peptides containing the fragment sequence LDHIGTL/SSGMGCNWPV/AQNRPK/NPTPAGS (Fig. 4). As the only potentially *N*-linked motif (Asn-X-Ser/Thr, X ≠ Pro) contains a proline residue, it is likely that glycosylation is *O*-linked as observed with TcNPa. TcNPa is a similar novel natriuretic peptide isolated from the venom of *Tropidechis carinatus* but with a disaccharide (galactosyl-β(1–3)-N-acetylgalactosamine (Gal-GalNAc)) *O*-linked to the C-terminal threonine residue. The synthetic glycosylated and non-glycosylated (TcNP-ng) variants both displayed bifunctionality to NPR-A and NPR-B receptors with similar EC_{50} values suggesting the *O*-linked glycan is not directly involved in receptor binding. Likewise, NMR structural analysis and stability testing did not infer a structural role for the glycan [51]. Interestingly, the glycosylated region in the C-terminus of TcNPa (−PKPTPGG) is highly similar to the corresponding regions of PasNPa, PasNPb and PnNPa (−PK/NPTPAGS) indicating the threonine residue of these isoforms may be *O*-linked glycosylated.

Unlike any other toxin type examined in this study, the evolutionary assessment of natriuretic peptides revealed a strong influence of negative selection, similar to what has been previously observed for lizard venom forms [52]. The computed ω was 0.85, while the BEB approach failed to identify positively selected sites (Table S2). Similarly, MEME failed to identify sites experiencing episodic influence of adaptive

Fig. 3. Multiple sequence alignment of Kunitz type protease inhibitor isoforms. Underlined sequence in italics depicts signal peptide. Underlined sequence NOT in italics means the sequence was not identified as a signal peptide by *SignalP 4.1c* (# and ^ depict full sequences for Aspidorhynchin-2 and Nuchalinin-4 respectively). The red bar indicates the hairpin turn between beta sheets and residues highlighted in gray represent P1. Sequences in bold represent ≥99% confidence MS/MS data matched to transcript sequences. Similarity scores (E-values) between newly discovered precursor toxins and their respective BLAST hit (UniProt accession number Q90WA1) are listed.

selection (Table S2). Thus, natriuretic peptides seem to have evolved under the constraints of negative selection.

3.6. PLA_2 toxins

Elapid venom is known as a rich supply of PLA_2 enzymes that exhibit a broad range of pharmacological effects [53]. They are subgrouped into group IA or group IB depending on the presence of a characteristic elapid or pancreatic loop respectively; the elapid loop form contains a five amino acid deletion [54]. Generally, most elapid PLA_2 belong to group IA though some PLA_2 toxins from *Oxyuranus microlepidotus*, *Pseudonaja textilis* and *Notechis scutatus* have been classified as group IB PLA_2 [55]. It must be noted that the IA/IB division is artificial as the pancreatic loop has been lost on multiple occasions [14].

In this investigation, numerous group 1 PLA_2 isoforms were identified in all venom proteomes. These were grouped into acidic and basic forms with all forms containing the functional residue Asp49, which is characteristic of group 1 PLA_2s. The acidic forms are subdivided into three groups (textilotoxin chain C, textilotoxin chain D, acidic PLA_2 1 (UniProt accession number Q9W7J4), which show inter-specific variation (47.5% similarity) across all sequences (Fig. 5). They are dominated

Toxin ID	Toxin Sequence	Length	E-value
TNP-b (P83228) *O. scutellatus*	*MVGLSRLAGGGLLLLLLLALLPLALD*gkpaplpqalpealaggttalrrdvteeqqqqlv	60	
PnNPb *P. nuchalis*	*MVGLSRLAGGGLLLLLL-ALLPLALD*gkpvplpqalpeapeggmmallqelteeqqqpp-	58	1.0E-21
PtNP-b *P. textilis*	*MVGLSRLAGGGLLLL---ALLPLALD*gkpvplpqalpealeggmmasrqelteeqqqpp-	56	6.0E-26
PasNPa *P. aspidorhyncha*	*MVGLSRLAGGGLLLLLL-ALLPLALD*gkpaplpqalpeapaggvmas---vteeqr**rlpa**	56	1.0E-28
PasNPb *P. aspidorhyncha*	*MVGLSRLAGGGLLLLLLLALLPLALD*gkpaplpqalpeapaggvmas---vteeqr**rlpa**	57	2.0E-30
PnNPa# *P. nuchalis*	MV/GLSRLAGGGLLLLLALLPLALDgkpaplp----**eapaggvmas---vteeqrrlpa**	81	4.0E-22
TcNPa *T. carinatus*	----------	0	
TNP-b (P83228) *O. scutellatus*	aeessgpaagrSDPK---IGDGCFGLPLDHIGSVSGLGCNRPVQNRPKQIPGGS	111	
PnNPb *P. nuchalis*	aeessgpaagrSG**SK---IGNGCFGLPLDRISNTSGMGCR**SPVQKRPKSTPGGS	112	
PtNP-b *P. textilis*	aeessghaagrSGSK---**IGNGCFGLPLDR**ISNTSGMGCRNPIQNSPKSTPGGS	110	
PasNPa *P. aspidorhyncha*	**aeespdpaagr**SGSKVAKLGSGCFGIR**LDHIGTSSGMGCNWPVQNRPKPTPAGS**	110	
PasNPb *P. aspidorhyncha*	**aeespdpaagr**SGSKVAKLGSGCFGIR**LDHIGTLSGMGCNWPVQNRPKPTPAGS**	111	
PnNPa *P. nuchalis*	**aeespdpaagr**SSSKVAK**LGSGCFGTRLDHIGTSSGMGCNWPAQNRPNPTPAGS**	135	
TcNPa *T. carinatus*	----------SGSETAKIGDGCFGLPIDRIGSASGMGCGS-V---PKPTPGGS	31	

MVHPSPSPLLSLTPRHLTFSLSLSSTRGKMVGLSRLAGGGLLLLLLALLPLALD

Fig. 4. Multiple sequence alignment of natriuretic peptide isoforms. Underlined sequence in italics depicts signal peptide and residues in lower case depict propeptides. Underlined sequence NOT in italics means the sequence was not identified as a signal peptide by *SignalP 4.1c* (# depicts full sequence for PnNPa). Sequences in bold represent ≥99% confidence MS/MS data matched to transcript sequences. Similarity scores (E-values) between newly discovered precursor toxins and their respective BLAST hit (UniProt accession number P83228) are listed.

by PLA_2s that contain the pancreatic loop though we also observed PLA_2s similar to textilotoxin chain C (Uniprot accession number P30811) containing the elapid loop (Fig. 5). Intra-specific variation was observed for *P. textilis*: two PLA_2 1 (Uniprot accession number Q9W7J4) homologs with variation at the C-terminus and isoform PLA_2 3 were observed. Generally, all sequences contained seven disulfide bonds though four sequences contained 13 cysteine residues due to variation at the C-terminus. All isoforms of textilotoxin chain D contain the extra cysteine residue at the N-terminus responsible for the covalently linked dimer that contributes to the quarternary structure of the complex [56]. Likewise, the signal peptide and propeptide regions showed a high degree of similarity (Fig. 5).

Basic PLA_2s were also identified in all venoms. The sequences were homologous to the A and B chains (P23026 and P23027 UniProtKB accession numbers respectively) of textilotoxin. As well, their signal peptides were very similar with the exception of textilotoxin chain A homolog from *P. aspidorhyncha* that has a 6 residue insertion (Fig. S4). The two textilotoxin chain B homologs from *P. aspidorhyncha* could not be separated by unique peptide evidence (Fig. S4). We also observed that the proteins contain either 13 or 14 cysteine residues. Textilotoxin chain A homolog from *P. aspidorhyncha* C-terminus is truncated resulting in 13 cysteine residues (Fig. S4). While PLA_2s with odd number of cysteines have been discovered in snake venom before, including the D chain of the complex textilotoxin, most have been associated with forming covalent complexes with other molecules [4]. All four chains of the highly toxic textilotoxin (A, B, C and D) were identified in *P. textilis*, *P. aspidorhyncha* and *P. nuchalis* venoms.

For both acidic and basic PLA2s from this study, the Ca^{2+} binding loop (Y25-G-C-Y/F-C-G/S-S/K-G/R-G33) is relatively conserved, and with Asp49, form the bipyramidal cage for Ca^{2+}; an important cofactor for catalysis [57]. The PLA_2 enzymatic active motif (CCxxHDxCx; Position 44–52) that includes the active site dyad His48 and Asp99 is also relatively conserved. However, as previously reported, a number of sequences aligned to PLA2 1 from *P. textilis* (UniProt accession number Q9W7J4) have Tyr28 and Tyr52 substituted by phenylalanine residues that may contribute to reduced enzymatic activity [55] (Fig. 5).

Molecular evolutionary assessment provided fascinating insights into the evolution of PLA_2s in *Pseudonaja* spp. While PLA_2s in *P. nuchalis* and *P. textilis* experienced a significant influence of positive selection (ω of 2.04 and 2.20, respectively), PLA_2s in *P. aspidorhyncha* evolved under the constraints of negative selection ($\omega = 0.37$) (Table S2). While BEB approach identified as many as 34 and 45 positively selected sites in *P. nuchalis* and *P. textilis*, respectively, it failed to identify such sites in PLA_2s of *P. aspidorhyncha* (Table S2). However, MEME identified eight sites in PLA_2s of this species as experiencing short bursts of adaptive selection (Table S2).

3.7. Other venom toxins — factor X, factor V, venom nerve growth factor and CRVP

Coagulation factors have been discovered in many snake species and are functionally classified into four groups based on cofactor requirements. *P. textilis* is known to contain the group C prothrombin activators consisting of the serine proteinase factor Xa, non-covalently linked to the non-enzymatic cofactor Va, while combining with Ca^{2+} and phospholipids to make up the prothrombinase complex [58]. This large complex (>250 kDa) is involved in the cleavage of prothrombin to thrombin and causes severe hemostatic unbalance upon envenomation [9]. Here we noticed that mRNA-encoding coagulation factors were the most highly expressed transcripts in all three *Pseudonaja* transcriptomes, counting 1094 (49.0% of toxin transcripts), 1087 (35.6% of toxin transcripts) and 1425 (45.7% of toxin transcripts) for *P. textilis*, *P. aspidorhyncha* and *P. nuchalis* respectively. Peptides were characterized from both prothrombin activator sub-units in all three venoms with the exception of *P. nuchalis* showing no peptide evidence of factor Xa chain (Table S3). Peptides from the coagulation factor X isoform 2 (UniProt accession number Q1L658); known to be expressed in the liver tissue but also expressed in the venom gland of Queensland *Pseudonaja* populations were identified in *P. textilis* and *P. aspidorhyncha* [59,60].

Transcript evidence for venom nerve growth factor fragments was observed in all venoms. However, only 9 transcripts were identified in *P. textilis*, whereas 18 were found in *P. aspidorhyncha* and 83 in *P. nuchalis*. Protein evidence from the propeptide and mature protein was observed from the *P. nuchalis* nerve growth factor (Table S3).

Similarly, while *P. textilis* and *P. aspidorhyncha* contained 98 and 159 transcripts of CRVPs respectively, only one was observed in *P. nuchalis*. Evidence of CRVP protein expression was retrieved in *P. textilis* (51.9% mature protein sequence coverage) and *P. aspidorhyncha* (75.2% mature protein sequence coverage) (Fig. S5). The *P. textilis* CRVP is identical to the reported sequence while the *P. aspidorhyncha* isoform shows two substitutions within this sequence. Both CRVPs expressed the short propeptide sequence.

4. Discussion

Using next generation sequencing in combination with sensitive and powerful mass spectrometry we compared the protein toxin repertoires of two western brown species, *P. aspidorhyncha* and *P. nuchalis* and the eastern brown, *P. textilis*, with an emphasis on polypeptides < 16 kDa. In the past, Australian snake venoms have been under-investigated with a focus on the medically important snake venoms and the more abundant

Toxin ID	Toxin Sequence	Length	E-value
Textilotoxin Chain C (P30811) *P. textilis*	*MHPAHLLVLLGVYVSLLGA*aripplplNLIQFSNMIKCTIPGSQPLLDYANYGCYCGPGN	60	
Textilotoxin Chain C homolog *P. nuchalis*	*MHPAHLLVLLAVYVSLLGA*aripplpl**NLVQFSNMIKCTIPGSQPLLDYADYGCYCGSGG**	60	2.0E-98
Textilotoxin Chain C homolog *P. aspidorhyncha*	*MHPAHLLVLLGVYVSLLGA*aripplpl**NLIQFSNMIKCTIPGSQPLLDYANYGCYCGPGN**	60	3.0E-102
Textilotoxin Chain C* *P. textilis*	*MHPAHLLVLLGVYVSLLGA*ar**ipplplNLIQFSNMIKCTIPGSQPLLDYANYGCYCGPGN**	60	3.0E-101
Textilotoxin Chain D (P23028) *P. textilis*	*MHPAHLLVLLGVCVSLLGA*asiprpslNIMLFGNMIQCTIPCEQSWLGYLDYGCYCGSGS	60	
Textilotoxin Chain D* *P. textilis*	*MHPAHLLVLLGVCVSLLGA***asiprpslNIMLFGNMIQCTIPCEQSWLGYLDYGCYCGSGS**	60	6.0E-110
Textilotoxin Chain D *P. aspidorhyncha*	*MHPAHLLVLLGVCVSLLGA*asiprpslNIMLFGNMIQCTIPCEQSWLGYLDYGCYCGSGS	60	9.0E-109
Textilotoxin Chain D *P. nuchalis*	*MHPAHLLVLLGVCVSLLGA*asiprpslNIMLFGNMIQCTIPCEQSWLGYLDYGCYCGLGS	60	8.0E-56
Acidic PLA_2 1 (Q9W7J4) *P. textilis*	*MHPAHLLVLLGVCVSLLGA*aripplplSLDDFSNLITCANRGSRSWLDYAHYGCFCGSGG	60	
PLA_2 1 homolog *P. textilis*	*MHPAHLLVLLGVCVSLLGA*aripplpl**SLDDFSNLITCANR**GSR**SWLDYAHYGCFCGSGG**	60	4.0E-108
PLA_2 1 homolog 2 *P. textilis*	*MHPAHLLVLLGVCVSLLGA*aripplpl**SLDDFSNLITCANR**GSRSWLDYAHYGCFCGSGG	60	7.0E-96
PLA_2 1 *P. nuchalis*	*MHPAHLLVLLGVCVSLLGA*aripllpl**SLDNFSSLITCANRGSRNWWDYAHYGCFCGSGG**	60	2.0E-93
PLA_2 1 *P. aspidorhyncha*	*MHPAHLLVLLGVCVSLLGA*aripplpl**SLDNFSNLITCANR**GSR**NWWDYAHYGCFCGSGG**	60	8.0E-97
PLA_2 2 *P. aspidorhyncha*	*MYPAHLLVLLGVCVSLLGA*aripplpl**SLDNFSNLITCANR**GSR**NWWDYAHYGCFCGSGG**	60	6.0E-96
PLA_2 3 *P. aspidorhyncha*	*MHPAHLLVPLAVCVSLLGA*aripplplSLLQFRILIKCANHNSR**DVLDYANYGCYCGKGG**	60	2.0E-94
PLA_2 4 *P. aspidorhyncha*	*MHPAHLLVLLGVCVSLLGA*tripplplSLVQFRILIKCANHSSR**NVLDYADYGCYCGKGG**	60	1.0E-96
PLA_2 2 *P. nuchalis*	*MHPAHLLVPLAVCVSLLGA*ar**ipplplSLLQFRILIKCANHNSRDVLDYANYGCYCGKGG**	60	3.0E-105
PLA_2 3 *P. textilis*	*MHPAHLLVLLGVCVSLLGA*aripplpl**SLVQFRILIKCANHNSRNVLDYADYGCYCGKGG**	60	4.0E-109
Textilotoxin Chain C (P30811) *P. textilis*	NGTPVDDVDRCCQAHDECYDEASNH-GC-----YPELTLYDYYCDTGVPYCK-ARTQCQV	113	
Textilotoxin Chain C homolog *P. nuchalis*	**SGTPVDDVDRCCQAHDECYDEASNH-GC-----YPELTLYDYYCDTGVPYCK**-ARTQCQV	113	
Textilotoxin Chain C homolog *P. aspidorhyncha*	**NGTPVDDVDRCCQAHDECYDEASNH-GC-----YPELTLYDYYCDTGVPYCK-ARTQCQV**	113	
Textilotoxin Chain C* *P. textilis*	**NGTPVDDVDRCCQAHDECYDEASNH-GC-----YPELTLYDYYCDTGVPYCK-ARTQCQV**	113	
Textilotoxin Chain D (P23028) *P. textilis*	SGIPVDDVDKCCKTHDECYYKAGQIPGCSVQPNEVFNVDYSYECNEGQLTCNESNNECEM	120	
Textilotoxin Chain D* *P. textilis*	**SGIPVDDVDKCCK**THDECYYKAGQIPGCSVQPNEVFNVDYSYECNEGQLTCNESNNECEM	120	
Textilotoxin Chain D *P. aspidorhyncha*	SGTPVDDVDRCCKTHDECYYKAGQIPGCSVQPNEVFNIDYSYECNEGQLTCNESNNECEM	120	
Textilotoxin Chain D *P. nuchalis*	SGIPVDDVDRCCR**THDECYYK**AGQIPGCSVQPNEVFNVDYSYECNEGQLTCNESNNECEM	120	
Acidic PLA_2 1 (Q9W7J4) *P. textilis*	SGTPVDDLDRCCQVHDNCFGDAEKLPACNYLFSGPYWNPYSYKCNEGEITCTDDNDECAA	120	
PLA_2 1 homolog *P. textilis*	**SGTPVDDLDRCCQVHDNCFGDAEKLPACNYLFSGPYWNPYSYKCNEGEITCTDDNDECAA**	120	
PLA_2 1 homolog 2 *P. textilis*	**SGTPVDDLDRCCQVHDNCFGDAEKLPACNYLFSGPYWNPYSYKCNEGEITCTDDNDECKA**	120	
PLA_2 1 *P. nuchalis*	**SGTPVDELDRCCQEHDNCFGDAEKLPACNYRFSGPYWNPYSYECNEGEITCTDDKDECKA**	120	
PLA_2 1 *P. aspidorhyncha*	**SGTPVDELDRCCQAHDNCFGDAEKLPACNYRFSGPYWNPYSYKCNEGEITCTDDNDECKA**	120	
PLA_2 2 *P. aspidorhyncha*	**SGTPVDELDRCCQAHDNCFGDAEKLPACNYRFSGPYWNPYSYKCNEGEITCTDDNDECKA**	120	

Fig. 5. Multiple sequence alignment of acidic PLA_2 isoforms. Underlined sequence in italics depicts signal peptide and residues in lower case depict propeptides. Sequences in bold represent ≥99% confidence MS/MS data matched to transcript sequences. Similarity scores (E-values) between newly discovered precursor toxins and their respective BLAST hit (UniProt accession number P30811, P23028 and Q9W7J4) are listed.

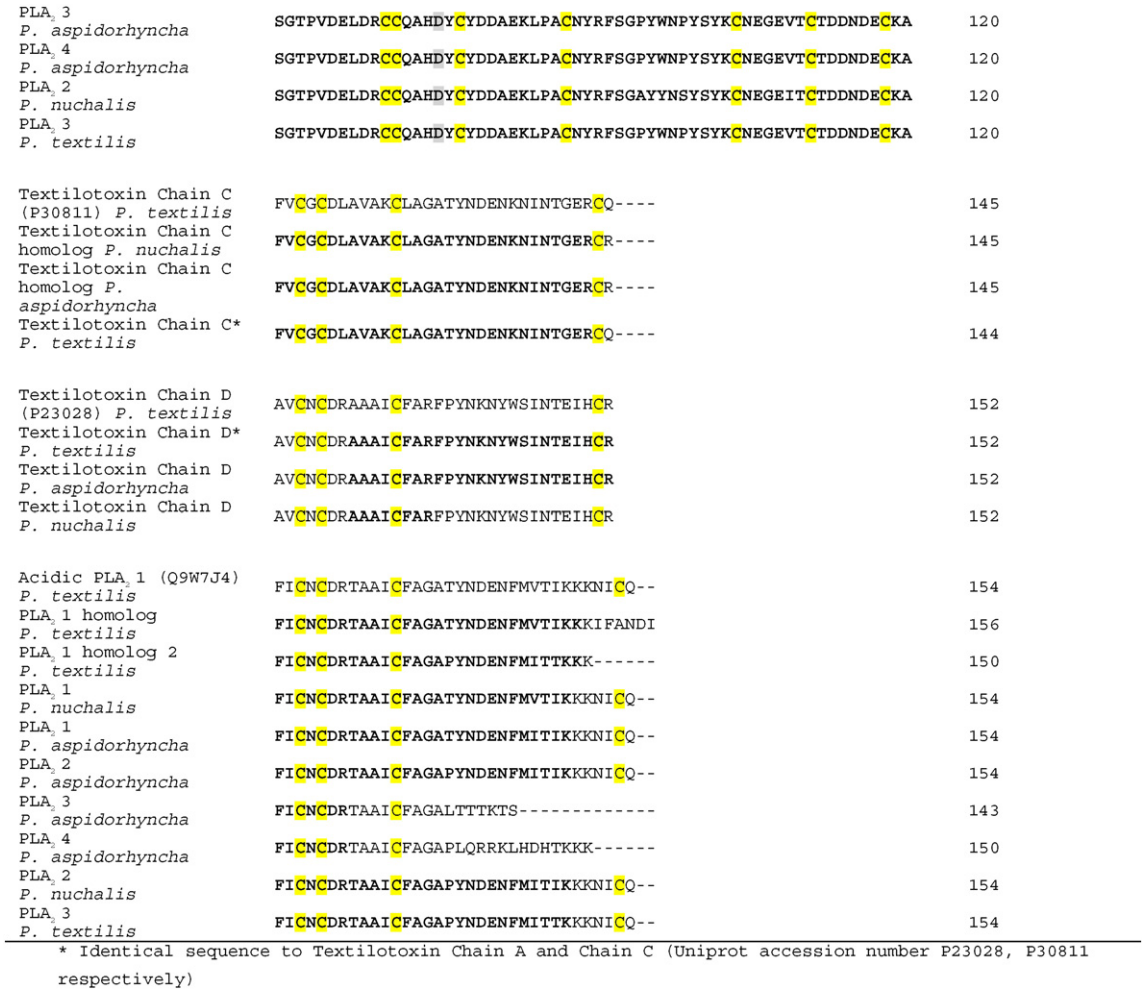

Fig. 5 (*continued*).

toxins identified through low throughput technologies employed at the time. The majority of this research has centered on antivenom investigation, but as the last 20 years of research has shown, snake venom also provides insights into evolutionary and ecological adaptations and is a largely untapped source of compounds with potential as research tools or therapeutics [2]. The advent of high throughput NGS platforms (transcriptomic data) and highly sensitive and accurate mass spectrometry (proteomic data) has driven the growth of analytical techniques to a depth and refinement that is both rapid and relatively cost effective [61]. This has led to an increased number of 'proteotranscriptomic' studies of venoms of different genera utilizing different NGS systems and mass spectrometry platforms; the latter being pivotal in providing an accurate picture of venom composition. Nevertheless, *de novo* sequencing of cDNA can suffer from erroneous assembly, especially if there is no reference genome as a template, which is the case for nearly all snake venoms. Likewise, studies utilizing public databases and/or venom milked from multiple specimens risk overestimating the number of isoforms identified because they are unable to distinguish between different toxin family members due to protein similarity [61]. As *Pseudonaja* venom is known to comprise an abundance of components <16 kDa, we chose to utilize only the reads from the 454 pyrosequencing of the three venoms to avoid the vagaries associated with assembly processing of the data [62]. Using the whole venom and venom glands from the same specimen with integration of the annotated transcripts and high quality MS/MS data from the venom proteome led to the identification of 65 isoforms from six snake toxin families. By using a specimen-specific proteotranscriptomic strategy, we were able to attain up to 100% sequence coverage in all toxin families (<16 kDa) except for PLA_2 (96.6%). A study using Illumina NGS and the highly sensitive LTQ orbitrap MS/MS data on *Protobothrops flavoviridis* and *Ovophis okinavensis* from the tissue of the same specimens achieved similar sequence coverage to this study (PLA_2 1–98.3%) [63]. To illustrate the difference between species specific and specimen specific sequences, an Illumina assembly of *Crotalus adamanteus* produced 123 unique full length toxin transcripts with the toxin sequences accounting for 35.4% of total reads [64]. This Illumina transcriptome was used as the database in a subsequent study to match the venom proteome of the same species resulting in identifying 36 toxins based on unique peptide evidence [65]. However, 19 different proteins from the same toxin family could not be differentiated due to protein similarity and mass spectrometry limitations [65]. Hence, the contrasting results between species specific and specimen specific studies are not only dependent on the quality of MS/MS data utilized, but more so, correlating the transcriptome and proteome from the same specimen allows for accurate and extensive integration between the two 'omics'. This approach can be valuable when analyzing venom for drug mining, research tools, or anti-venom and evolutionary research.

The high level of proteomic sequencing of the transcript isoforms highlighted intra- (*P. textilis*) and inter-specific variation within the venoms. Three populations of *P. textilis* are recognized (Queensland, South Australia and southern New Guinea) [16] with variation in venom procoagulation and neurotoxin components and potency reported previously between Queensland and South Australian populations [59,66]. LC-MS analysis of the South Australian venom sample revealed postsynaptic α-neurotoxins, including type II (UniProt

accession number Q9W7J5, P13495 and A8HDK6) and type III (UniProt accession number Q9W7K2, Q9W7K1, Q9W7J7, Q9W7J6) 3FTx. In a previous study, no type II 3FTxs and only a single postsynaptic neurotoxin (UniProt accession number Q9W7J6) from the α-neurotoxin type III group was found in Queensland venom sample [59]. The current study identified five new type III and two new type II (plus one previously reported in the literature) postsynaptic 3FTxs from the Queensland *P. textilis* specimen. However, it should be noted that the number of type II neurotoxin transcripts identified for *P. textilis* was only 27 (3 type II isoforms; ratio peptide:transcript 1:9) compared to 409 (5 type II isoforms; ratio peptide:transcript 1:82) and 327 (5 type II isoforms; ratio peptide:transcript 1:65) for *P. aspidorhyncha* and *P. nuchalis* respectively, which may be an indication why type II neurotoxins have not been identified in previous studies. Such a difference in peptide/transcript ratio highlights the importance of proteome-based strategies in deconstructing venom composition. Furthermore, recent research is beginning to identify microRNA (miRNA) as an integral factor in modulating the phenotypic profile of snake venom [67,68]. Comparing the miRNA profiles of the three *Pseudonaja* venom populations may provide further clues to the molecular mechanisms underlying transcriptional and translational activity of the venom gland.

The differences in primary structure amongst the 3FTxs, PLA_2s, kunitz type protease inhibitors and natriuretic peptide toxin families from the three *Pseudonaja* venoms indicate a degree of intra-specific and inter-specific variation. Different venom composition can be a factor of geographical disposition along with individual, gender-specific, regional, dietary and seasonal variations [69]. Overall, the diet of these three species are similar (frogs, birds, reptiles and mammals) that may differ on a microgeographic scale [70]. The obvious difference between these species is geographical and may be a factor in the subtle variation identified between isoforms. Further research to determine any correlation between venom composition and ecological factors would require larger population studies and also include the species *Pseudonaja modesta* and *Pseudonaja inframacula* due to the lack of mammals in their diet [70].

This study identified six 3FTx isoforms and five PLA_2 isoforms from *P. textilis*, *P. aspidorhyncha* and *P. nuchalis* containing an odd number of cysteine residues that may be involved in toxin complex formation via interchain disulfide bridging between toxins. For example, textilotoxin is a potent hexameric presynaptic neurotoxic PLA_2 complex consisting of six PLA_2 subunits whereby subunit D consists of two covalently linked polypeptide chains; each consisting of 133 residues [56,71]. Irditoxin is a covalently linked heterodimeric 3FTx isolated from *Boiga irregularis* with each peptide containing 11 cysteine residues [32]. Additionally, a number of Type II 3FTxs containing nine cysteine residues were identified from the transcriptome of *Pseudonaja modesta* recently [14]. Further, these toxins containing an odd-number of cysteine residues may also represent a new toxin class. A recent study identified the conotoxin, μO§-conotoxin GVIIJ, which contains seven cysteine residues where the extra cysteine was post-translationally modified and was involved in inhibiting Na_V1-channels [72].

Novel glycosylated natriuretic peptides containing a hexose-N-acetyl hexosamine and sialic acid were identified in *P. aspidorhyncha* and *P. nuchalis* venom. An *O*-linked glycosylated natriuretic peptide isolated from *Tropidechis carinatus* showed drug lead potential by demonstrating bi-functionality to both NPR-A and NPR-B natriuretic peptide receptors [51]. Furthermore, the peptide showed α-helix secondary structure and relative stability in plasma and trypsin solution. Interestingly, *P. aspidorhyncha* and *P. nuchalis* natriuretic peptides contain a C-terminal threonine residue surrounded by proline residues (PK/NPTPAGS) similar to the *T. carinatus* natriuretic peptide that may indicate the trisaccharide is localized at the C-terminus (Fig. 4). Further investigations are warranted to determine whether the glycan plays a structural or functional role for this toxin class.

Comparing the various isoform sequences of all toxin families from *P. textilis*, *P. aspidorhyncha* and *P. nuchalis* identified inter-specific variation within the *Pseudonaja* clade and exemplified the influence of adaptive selection on the evolution of these toxin types. While most families evolved under the influence of positive selection, natriuretic peptides were surprisingly found to have evolved under the constraints of negative selection. Similarly, PLA_2s in *P. aspidorhyncha* experienced a significant influence of negative selection, while the same toxin type evolved rapidly in *P. textilis* and *P. nuchalis*. However, several episodically adaptive sites were found in PLA_2s of *P. aspidorhyncha*, suggesting that certain sites in PLA_2s of this species undergo episodic molecular fine-tuning. Thus, the molecular evolutionary assessments in this study unraveled several fascinating insights into the interspecific variation in venom.

Importantly, such rich interspecific chemical diversity highlighted from the isoforms of this study needs to be screened and characterized against molecular targets. Not only is there a potential for novel bioactivity discovery but the high degree of isoform diversity may represent fine tuning of efficacy in receptor binding. Functional studies done on Australian elapid type II and type III neurotoxins identified they bind to skeletal nAChR with various EC50 values [33,36]. The primary structure diversity identified in 3FTxs in previous studies has been attributed to contributing to prey flexibility for the snake [14]. While this may be true for the diverse diet range for *Pseudonaja* species; the accelerated evolution observed in 3FTxs in this study may also facilitate increased efficacy in receptor binding. However, more functional activity must be done for the various toxin families of Australian elapids, especially the type III neurotoxins.

The *Pseudonaja* group has traditionally been problematic concerning their taxonomic classification at the species level [15]. Recently, the *P. nuchalis* species was split into 3 species (*P. aspidorhyncha*, *P mengdeni* and *P. nuchalis*) based on karyotypes, mitochondrial DNA sequences, and a multivariate morphometric analysis [15,16]. Overall, the transcriptome and proteome profiles of *P. aspidorhyncha* and *P. nuchalis* were similar to *P. textilis*, that is, prothrombin activators, PLA_2s, kunitz type peptides, natriuretic peptides and 3FTxs were highly expressed. Hence, in the absence of LD_{50} data for the western brown species, we may expect the venom of *P. aspidorhyncha* and *P. nuchalis* to be similar in action upon envenomation when compared to *P. textilis* by inducing cardiovascular, hemostatic and neuromuscular disruptions. Considering *P. textilis* is the second most venomous snake worldwide and well regarded as the most dangerous snake in Australia; both *P. aspidorhyncha* and *P. nuchalis* should be treated with equal respect.

Funding

This study was supported by grants from the National Health and Medical Research Council (PFA; 569927) and the Australian Research Council (PFA; dP130103275). BGF was funded by the Australian Research Council and the University of Queensland. KS was supported by the Marie Skłodowska-Curie Individual Fellowship (654294).

Conflict of interest

None declared.

Transparency document

The Transparency document associated with this article can be found, in the online version.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2015.11.019.

References

[1] G.W. Birrell, S.T.H. Earl, T.P. Wallis, P.P. Masci, J. de Jersey, J.J. Gorman, et al., The diversity of bioactive proteins in Australian snake venoms, Mol. Cell. Proteomics 6 (2007) 973–986, http://dx.doi.org/10.1074/mcp.M600419-MCP200.

[2] T.A. Reeks, B.G. Fry, P.F. Alewood, Privileged frameworks from snake venom, Cell. Mol. Life Sci. (2015) 1–20, http://dx.doi.org/10.1007/s00018-015-1844-z.

[3] B.G. Fry, N. Vidal, L. van der Weerd, E. Kochva, C. Renjifo, Evolution and diversification of the Toxicofera reptile venom system, J. Proteome 72 (2009) 127–136, http://dx.doi.org/10.1016/j.jprot.2009.01.009.

[4] R. Doley, R.M. Kini, Protein complexes in snake venom, Cell. Mol. Life Sci. 66 (2009) 2851–2871, http://dx.doi.org/10.1007/s00018-009-0050-2.

[5] S.D. Aird, S. Aggarwal, A. Villar-Briones, M.M. Tin, K. Terada, A.S. Mikheyev, Snake venoms are integrated systems, but abundant venom proteins evolve more rapidly, BMC Genomics 16 (2015) 647, http://dx.doi.org/10.1186/s12864-015-1832-6.

[6] S. Eichberg, L. Sanz, J.J. Calvete, D. Pla, Constructing comprehensive venom proteome reference maps for integrative venomics, Expert Rev. Proteomics 12 (2015) 557–573, http://dx.doi.org/10.1586/14789450.2015.1073590.

[7] C.J. Stewart, Snake bite in Australia: first aid and envenomation management, Accid. Emerg. Nurs. 11 (2003) 106–111, http://dx.doi.org/10.1016/S0965-2302(02)00189-3.

[8] V.S. Rao, S. Swarup, R.M. Kini, The nonenzymatic subunit of pseutarin C, a prothrombin activator from eastern brown snake (Pseudonaja textilis) venom, shows structural similarity to mammalian coagulation factor V, Blood 102 (2003) 1347–1354, http://dx.doi.org/10.1182/blood-2002-12-3839.

[9] R.M. Kini, The intriguing world of prothrombin activators from snake venom, Toxicon 45 (2005) 1133–1145, http://dx.doi.org/10.1016/j.toxicon.2005.02.019.

[10] T.N. Minh Le, M.A. Reza, S. Swarup, R.M. Kini, Gene duplication of coagulation factor V and origin of venom prothrombin activator in Pseudonaja textilis snake, Thromb. Haemost. 93 (2005) 420–429, http://dx.doi.org/10.1267/THRO05030420.

[11] G.W. Birrell, S. Earl, P.P. Masci, J. de Jersey, T.P. Wallis, J.J. Gorman, et al., Molecular diversity in venom from the Australian brown snake, Pseudonaja textilis, Mol. Cell. Proteomics. 5 (2006) 379–389, http://dx.doi.org/10.1074/mcp.M500270-MCP200.

[12] S.T.H. Earl, G.W. Birrell, T.P. Wallis, L.D.S. Pierre, P.P. Masci, J. de Jersey, et al., Post-translational modification accounts for the presence of varied forms of nerve growth factor in Australian elapid snake venoms, Proteomics 6 (2006) 6554–6565, http://dx.doi.org/10.1002/pmic.200600263.

[13] P.P. Masci, A.N. Whitaker, L.G. Sparrow, J. de Jersey, D.J. Winzor, D.J. Watters, et al., Textilinins from *Pseudonaja textilis* textilis. Characterization of two plasmin inhibitors that reduce bleeding in an animal model, Blood Coagul. Fibrinolysis 11 (2000) 385–393.

[14] T.N.W. Jackson, K. Sunagar, E.A.B. Undheim, I. Koludarov, A.H.C. Chan, K. Sanders, et al., Venom down under: dynamic evolution of Australian elapid snake toxins, Toxins 5 (2013) 2621–2655, http://dx.doi.org/10.3390/toxins5122621.

[15] A. Skinner, S.C. Donnellan, M.N. Hutchinson, R.G. Hutchinson, A phylogenetic analysis of Pseudonaja (Hydrophiinae, Elapidae, Serpentes) based on mitochondrial DNA sequences, Mol. Phylogenet. Evol. 37 (2005) 558–571, http://dx.doi.org/10.1016/j.ympev.2005.06.020.

[16] A. Skinner, A multivariate morphometric analysis and systematic review of Pseudonaja (Serpentes, Elapidae, Hydrophiinae), Zool. J. Linnean Soc. 155 (2009) 171–197, http://dx.doi.org/10.1111/j.1096-3642.2008.00436.x.

[17] R.K. Judge, P.J. Henry, P. Mirtschin, G. Jelinek, J.A. Wilce, Toxins not neutralized by brown snake antivenom, Toxicol. Appl. Pharmacol. 213 (2006) 117–125, http://dx.doi.org/10.1016/j.taap.2005.09.010.

[18] I. Vetter, J.L. Davis, L.D. Rash, R. Anangi, M. Mobli, P.F. Alewood, et al., Venomics: a new paradigm for natural products-based drug discovery, Amino Acids 40 (2011) 15–28, http://dx.doi.org/10.1007/s00726-010-0516-4.

[19] UniProt Consortium, Update on activities at the Universal Protein Resource (UniProt) in 2013, Nucleic Acids Res. 41 (2013) D43–D47, http://dx.doi.org/10.1093/nar/gks1068.

[20] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods 8 (2011) 785–786, http://dx.doi.org/10.1038/nmeth.1701.

[21] Z. Yang, Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution, Mol. Biol. Evol. 15 (1998) 568–573.

[22] Z. Yang, PAML 4: phylogenetic analysis by maximum likelihood, Mol. Biol. Evol. 24 (2007) 1586–1591, http://dx.doi.org/10.1093/molbev/msm088.

[23] K. Sunagar, W.E. Johnson, S.J. O'Brien, V. Vasconcelos, A. Antunes, Evolution of CRISPs associated with toxicoferan-reptilian venom and mammalian reproduction, Mol. Biol. Evol. 29 (2012) 1807–1822, http://dx.doi.org/10.1093/molbev/mss058.

[24] Z. Yang, W.S.W. Wong, R. Nielsen, Bayes empirical Bayes inference of amino acid sites under positive selection, Mol. Biol. Evol. 22 (2005) 1107–1118, http://dx.doi.org/10.1093/molbev/msi097.

[25] B. Murrell, S. Moola, A. Mabona, T. Weighill, D. Sheward, S.L. Kosakovsky Pond, et al., FUBAR: a fast, unconstrained Bayesian approximation for inferring selection, Mol. Biol. Evol. 30 (2013) 1196–1205, http://dx.doi.org/10.1093/molbev/mst030.

[26] S.L.K. Pond, S.D.W. Frost, S.V. Muse, HyPhy: hypothesis testing using phylogenies, Bioinformatics (Oxford, England) 21 (2005) 676–679, http://dx.doi.org/10.1093/bioinformatics/bti079.

[27] B. Murrell, J.O. Wertheim, S. Moola, T. Weighill, K. Scheffler, S.L. Kosakovsky Pond, Detecting individual sites subject to episodic diversifying selection, PLoS Genet. 8 (2012), e1002764http://dx.doi.org/10.1371/journal.pgen.1002764.

[28] S.L. Kosakovsky Pond, D. Posada, M.B. Gravenor, C.H. Woelk, S.D.W. Frost, Automated phylogenetic detection of recombination using a genetic algorithm, Mol. Biol. Evol. 23 (2006) 1891–1901, http://dx.doi.org/10.1093/molbev/msl051.

[29] V. Lavergne, I. Harliwong, A. Jones, D. Miller, R.J. Taft, P.F. Alewood, Optimized deep-targeted proteotranscriptomic profiling reveals unexplored Conus toxin diversity and novel cysteine frameworks, Proc. Natl. Acad. Sci. 112 (2015) E3782–E3791, http://dx.doi.org/10.1073/pnas.1501334112.

[30] K. Sunagar, T. Jackson, E. Undheim, S. Ali, A. Antunes, B. Fry, Three-fingered RAVERs: rapid accumulation of variations in exposed residues of snake venom toxins, Toxins 5 (2013) 2172–2208, http://dx.doi.org/10.3390/toxins5112172.

[31] B.G. Fry, W. Wüster, R.M. Kini, V. Brusic, A. Khan, D. Venkataraman, et al., Molecular evolution and phylogeny of elapid snake venom three-finger toxins, J. Mol. Evol. 57 (2003) 110–129, http://dx.doi.org/10.1007/s00239-003-2461-2.

[32] J. Pawlak, S.P. Mackessy, N.M. Sixberry, E.A. Stura, M.H. Le Du, R. Ménez, et al., Irditoxin, a novel covalently linked heterodimeric three-finger toxin with high taxon-specific neurotoxicity, FASEB J. 23 (2009) 534–545, http://dx.doi.org/10.1096/fj.08-113555.

[33] N. Gong, A. Armugam, K. Jeyaseelan, Postsynaptic short-chain neurotoxins from *Pseudonaja textilis*. cDNA cloning, expression and protein characterization, Eur. J. Biochem. 265 (1999) 982–989.

[34] L.S. Pierre, H. Fischer, D.J. Adams, M. Schenning, N. Lavidis, J. de Jersey, et al., Distinct activities of novel neurotoxins from Australian venomous snakes for nicotinic acetylcholine receptors, Cell. Mol. Life Sci. 64 (2007) 2829–2840, http://dx.doi.org/10.1007/s00018-007-7352-z.

[35] V. Lavergne, R.J. Taft, P.F. Alewood, Cysteine-rich mini-proteins in human biology, Curr. Top. Med. Chem. 12 (2012) 1514–1533.

[36] M. Laskowski Jr., M.A. Qasim, What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes? Biochim. Biophys. Acta Protein Struct. Mol. Enzymol. 1477 (2000) 324–337, http://dx.doi.org/10.1016/S0167-4838(99)00284-8.

[37] M. Laskowski Jr., I. Kato, Protein inhibitors of proteinases, Annu. Rev. Biochem. 49 (1980) 593–626, http://dx.doi.org/10.1146/annurev.bi.49.070180.003113.

[38] L.S. Pierre, S.T. Earl, I. Filippovich, N. Sorokina, P.P. Masci, J. De Jersey, et al., Common evolution of waprin and kunitz-like toxin families in Australian venomous snakes, Cell. Mol. Life Sci. 65 (2008) 4039–4054, http://dx.doi.org/10.1007/s00018-008-8573-5.

[39] I. Filippovich, N. Sorokina, P.P. Masci, J. De Jersey, A.N. Whitaker, D.J. Winzor, et al., A family of textilinin genes, two of which encode proteins with antihaemorrhagic properties, Br. J. Haematol. 119 (2002) 376–384, http://dx.doi.org/10.1046/j.1365-2141.2002.03878.x.

[40] E.K.I. Millers, M. Trabi, P.P. Masci, M.F. Lavin, J. de Jersey, L.W. Guddat, Crystal structure of textilinin-1, a Kunitz-type serine protease inhibitor from the venom of the Australian common brown snake (*Pseudonaja textilis*), FEBS J. 276 (2009) 3163–3175, http://dx.doi.org/10.1111/j.1742-4658.2009.07034.x.

[41] E.K.I. Millers, L.A. Johnson, G.W. Birrell, P.P. Masci, M.F. Lavin, J. de Jersey, et al., The structure of human microplasmin in complex with textilinin-1, an aprotinin-like inhibitor from the Australian brown snake, PLoS ONE 8 (2013)http://dx.doi.org/10.1371/journal.pone.0054104.

[42] S.T.H. Earl, P.P. Masci, J. de Jersey, M.F. Lavin, J. Dixon, Drug development from australian elapid snake venoms and the venomics pipeline of candidates for haemostasis: textilinin-1 (Q8008), Haempatch™ (Q8009) and CoVase™ (V0801), Toxicon (2010)http://dx.doi.org/10.1016/j.toxicon.2010.12.010.

[43] S. Vink, A.H. Jin, K.J. Poth, G.A. Head, P.F. Alewood, Natriuretic peptide drug leads from snake venom, Toxicon (2010)http://dx.doi.org/10.1016/j.toxicon.2010.12.001.

[44] W. Yan, F. Wu, J. Morser, Q. Wu, Corin, a transmembrane cardiac serine protease, acts as a pro-atrial natriuretic peptide-converting enzyme, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 8525–8529, http://dx.doi.org/10.1073/pnas.150149097.

[45] C. Wu, F. Wu, J. Pan, J. Morser, Q. Wu, Furin-mediated processing of pro-C-type natriuretic peptide, J. Biol. Chem. 278 (2003) 25847–25852, http://dx.doi.org/10.1074/jbc.M301223200.

[46] N.E. Zois, E.D. Bartels, I. Hunter, B.S. Kousholt, L.H. Olsen, J.P. Goetze, Natriuretic peptides in cardiometabolic regulation and disease, Nat. Rev. Cardiol. 11 (2014) 403–412, http://dx.doi.org/10.1038/nrcardio.2014.64.

[47] H. Ogawa, Y. Qiu, C.M. Ogata, K.S. Misono, Crystal structure of hormone-bound atrial natriuretic peptide receptor extracellular domain rotation mechanism for transmembrane signal transduction, J. Biol. Chem. 279 (2004) 28625–28631, http://dx.doi.org/10.1074/jbc.M313222200.

[48] B.G. Fry, J.C. Wickramaratana, S. Lemme, A. Beuve, D. Garbers, W.C. Hodgson, et al., Novel natriuretic peptides from the venom of the inland taipan (*Oxyuranus microlepidotus*): isolation, chemical and biological characterisation, Biochem. Biophys. Res. Commun. 327 (2005) 1011–1015, http://dx.doi.org/10.1016/j.bbrc.2004.11.171.

[49] A.C.M. Camargo, D. Ianzer, J.R. Guerreiro, S.M.T. Serrano, Bradykinin-potentiating peptides: beyond captopril, Toxicon (2011)http://dx.doi.org/10.1016/j.toxicon.2011.07.013.

[50] A. Brust, K. Sunagar, E.A.B. Undheim, I. Vetter, D.C. Yang, N.R. Casewell, et al., Differential evolution and neofunctionalization of snake venom metalloprotease domains, Mol. Cell. Proteomics 12 (2013) 651–663, http://dx.doi.org/10.1074/mcp.M112.023135.

[51] T. Reeks, A. Jones, A. Brust, S. Sridharan, L. Corcilius, B.L. Wilkinson, et al., A defined α-helix in the bifunctional *O*-glycosylated natriuretic peptide TcNPa from the venom of *Tropidechis carinatus*, Angew. Chem. Int. Ed. (2015)http://dx.doi.org/10.1002/anie.201411914 n/a.

[52] I. Koludarov, K. Sunagar, E.A.B. Undheim, T.N.W. Jackson, T. Ruder, D. Whitehead, et al., Structural and molecular diversification of the Anguimorpha lizard mandibular venom gland system in the arboreal species *Abronia graminea*, J. Mol. Evol. 75 (2012) 168–183, http://dx.doi.org/10.1007/s00239-012-9529-9.

[53] L. Cendron, I. Mičetić, P. Polverino de Laureto, M. Paoli, Structural analysis of trimeric phospholipase A2 neurotoxin from the Australian taipan snake venom, FEBS J. 279 (2012) 3121–3135, http://dx.doi.org/10.1111/j.1742-4658.2012.08691.x.

[54] R. Doley, X. Zhou, R.M. Kini, Snake venom phospholipase A2 enzymes, Handb. Venoms Toxins Reptil. 1st ed.CRC Press, Boca Raton, FL 2010, pp. 173–205.
[55] A. Armugam, N. Gong, X. Li, P.Y. Siew, S.C. Chai, R. Nair, et al., Group IB phospholipase A2 from *Pseudonaja textilis*, Arch. Biochem. Biophys. 421 (2004) 10–20, http://dx.doi.org/10.1016/j.abb.2003.09.045.
[56] J.A. Aquilina, The major toxin from the Australian common brown snake is a hexamer with unusual gas-phase dissociation properties, Proteins Struct. Funct. Bioinf. 75 (2009) 478–485, http://dx.doi.org/10.1002/prot.22259.
[57] S. Banumathi, K.R. Rajashankar, C. Nötzel, B. Aleksiev, T.P. Singh, N. Genov, et al., Structure of the neurotoxic complex vipoxin at 1.4 Å resolution, Acta Crystallogr. D Biol. Crystallogr. 57 (2001) 1552–1559, http://dx.doi.org/10.1107/S0907444901013543.
[58] P.P. Masci, A.N. Whitaker, J. de Jersey, Purification and characterization of a prothrombin activator from the venom of the Australian brown snake, *Pseudonaja textilis* textilis, Biochem. Int. 17 (1988) 825–835.
[59] J. Skejić, W.C. Hodgson, Population divergence in venom bioactivities of elapid snake *Pseudonaja textilis*: role of procoagulant proteins in rapid rodent prey incapacitation, PLoS ONE 8 (2013) e63988, http://dx.doi.org/10.1371/journal.pone.0063988.
[60] M.A. Reza, T.N. Minh Le, S. Swarup, R. Manjunatha Kini, Molecular evolution caught in action: gene duplication and evolution of molecular isoforms of prothrombin activators in *Pseudonaja textilis* (brown snake), J. Thromb. Haemost. 4 (2006) 1346–1353, http://dx.doi.org/10.1111/j.1538-7836.2006.01969.x.
[61] J.J. Calvete, Next-generation snake venomics: protein-locus resolution through venom proteome decomplexation, Expert Rev. Proteomics 11 (2014) 315–329, http://dx.doi.org/10.1586/14789450.2014.900447.
[62] L.D. Florea, S.L. Salzberg, Genome-guided transcriptome assembly in the age of next-generation sequencing, IEEE/ACM Trans. Comput. Biol. Bioinform. 10 (2013) 1234–1240.
[63] S.D. Aird, Y. Watanabe, A. Villar-Briones, M.C. Roy, K. Terada, A.S. Mikheyev, Quantitative high-throughput profiling of snake venom gland transcriptomes and proteomes (*Ovophis okinavensis* and *Protobothrops flavoviridis*), BMC Genomics 14 (2013) 790, http://dx.doi.org/10.1186/1471-2164-14-790.
[64] D.R. Rokyta, A.R. Lemmon, M.J. Margres, K. Aronow, The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*), BMC Genomics 13 (2012) 312, http://dx.doi.org/10.1186/1471-2164-13-312.
[65] M.J. Margres, J.J. McGivern, K.P. Wray, M. Seavy, K. Calvin, D.R. Rokyta, Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*), J. Proteome 96 (2014) 145–158, http://dx.doi.org/10.1016/j.jprot.2013.11.001.
[66] S. Flight, P. Mirtschin, P.P. Masci, Comparison of active venom components between eastern brown snakes collected from South Australia and Queensland, Ecotoxicology Lond. Engl. 15 (2006) 133–141, http://dx.doi.org/10.1007/s10646-005-0047-z.
[67] J. Durban, A. Pérez, L. Sanz, A. Gómez, F. Bonilla, S. Rodríguez, et al., Integrated "omics" profiling indicates that miRNAs are modulators of the ontogenetic venom composition shift in the central American rattlesnake, *Crotalus simus* simus, BMC Genomics 14 (2013) 234, http://dx.doi.org/10.1186/1471-2164-14-234.
[68] F.J. Vonk, N.R. Casewell, C.V. Henkel, A.M. Heimberg, H.J. Jansen, R.J.R. McCleary, et al., The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 20651–20656, http://dx.doi.org/10.1073/pnas.1314702110.
[69] B.G. Fry, N.R. Casewell, W. Wüster, N. Vidal, B. Young, T.N.W. Jackson, The structural and functional diversification of the toxicofera reptile venom system, Toxicon 60 (2012) 434–448, http://dx.doi.org/10.1016/j.toxicon.2012.02.013.
[70] R. Shine, Constraints, allometry, and adaptation: food habits and reproductive biology of Australian brownsnakes (Pseudonaja: Elapidae), Herpetologica 45 (1989) 195–207.
[71] J.A. Pearson, M.I. Tyler, K.V. Retson, M.E.H. Howden, Studies on the subunit structure of textilotoxin, a potent presynaptic neurotoxin from the venom of the Australian common brown snake (*Pseudonaja textilis*). 2. The amino acid sequence and toxicity studies of subunit D, Biochim. Biophys. Acta Protein Struct. Mol. Enzymol. 1077 (1991) 147–150, http://dx.doi.org/10.1016/0167–4838(91)90051-Z.
[72] J. Gajewiak, L. Azam, J. Imperial, A. Walewska, B.R. Green, P.K. Bandyopadhyay, et al., A disulfide tether stabilizes the block of sodium channels by the conotoxin μO§-GVIIJ, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 2758–2763, http://dx.doi.org/10.1073/pnas.1324189111.