

Molecular mechanisms underlying intraspecific variation in snake venom

Diana R. Amazonas^a, José A. Portes-Junior^a, Milton Y. Nishiyama-Jr^b, Carolina A. Nicolau^c, Hipócrates M. Chalkidis^d, Rosa H.V. Mourão^e, Felipe G. Grazziotin^h, Darin R. Rokytá^f, H. Lisle Gibbs^g, Richard H. Valente^c, Inácio L.M. Junqueira-de-Azevedo^{b,*}, Ana M. Moura-da-Silva^{a,*}

^a Laboratório de Imunopatologia, Instituto Butantan, Av. Vital Brazil, 1500, 05503-900 São Paulo, SP, Brazil

^b Laboratório Especial de Toxinologia Aplicada, Instituto Butantan, Av. Vital Brazil, 1500, 05503-900 São Paulo, SP, Brazil

^c Laboratório de Toxinologia, Instituto Oswaldo Cruz – IOC/FIOCRUZ, Av. Brasil, 4365, 21040-900 Rio de Janeiro, RJ, Brazil

^d Programa de Pós-Graduação em Recursos Naturais da Amazônia, Laboratório de Bioprospecção e Biologia Experimental, Universidade Federal do Oeste do Pará – UFOPA, Rua Vera Paz, s/n, 68035-110 Santarém, PA, Brazil

^e Faculdades Integradas do Tapajós/Faculdade da Amazônia – FIT/UNAMA, Rua Rosa Vermelha, 335, 68010-200 Santarém, PA, Brazil

^f Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA

^g Department of Evolution, Ecology and Organismal Biology Ohio State University, Columbus, OH 43210, USA

^h Laboratório Especial de Coleções Zoológicas, Instituto Butantan, Av. Vital Brazil, 1500, 05503-900 São Paulo, SP, Brazil

ARTICLE INFO

Keywords:

Snake venom
Evolution
Transcriptome
Proteome
Plasticity
Toxin function

ABSTRACT

Elucidating the molecular mechanisms underlying snake venom variability provides important clues for understanding how the biological functions of this powerful toxic arsenal evolve. We analyzed in detail individual transcripts and venom protein isoforms produced by five specimens of a venomous snake (*Bothrops atrox*) from two nearby but genetically distinct populations from the Brazilian Amazon rainforest which show functional similarities in venom properties. Individual variation was observed among the venoms of these specimens, but the overall abundance of each general toxin family was conserved both in transcript and in venom protein levels. However, when expression of independent paralogues was analyzed, remarkable differences were observed within and among each toxin group, both between individuals and between populations. Transcripts for functionally essential venom proteins (“core function” proteins) were highly expressed in all specimens and showed similar transcription/translation rates. In contrast, other paralogues (“adaptive” proteins) showed lower expression levels and the toxins they coded for varied among different individuals. These results provide support for the inferences that (a) expression and translational differences play a greater role in defining adaptive variation in venom phenotypes than does sequence variation in protein coding genes and (b) convergent adaptive venom phenotypes can be generated through different molecular mechanisms.

Significance: Analysis of individual transcripts and venom protein isoforms produced by specimens of a venomous snake (*Bothrops atrox*), from the Brazilian Amazon rainforest, revealed that transcriptional and translational mechanisms contribute to venom phenotypic variation. Our finding of evidence for high expression of toxin proteins with conserved function supports the hypothesis that the venom phenotype consists of two kinds of proteins: conserved “core function” proteins that provide essential functional activities with broader relevance and less conserved “adaptive” proteins that vary in expression and may permit customization of protein function. These observations allowed us to suggest that genetic mechanisms controlling venom variability are not restricted to selection of gene copies or mutations in structural genes but also to selection of the mechanisms controlling gene expression, contributing to the plasticity of this important phenotype for venomous snakes.

1. Introduction

Snake venoms show widespread variation in composition not only within and between species but also at higher taxonomic levels [1–3].

This variation is widely interpreted to represent adaptive variation that allows subduing and digesting prey or for defense against predators [4–6]. Venom heterogeneity also has important medical consequences by causing significant differences in venom toxicity and pathogenicity

* Corresponding authors.

E-mail addresses: inacio.azevedo@butantan.gov.br (I.L.M. Junqueira-de-Azevedo), ana.moura@butantan.gov.br (A.M. Moura-da-Silva).

<https://doi.org/10.1016/j.jprot.2018.03.032>

Received 8 January 2018; Received in revised form 8 March 2018; Accepted 31 March 2018

Available online 03 April 2018

1874-3919/ © 2018 Published by Elsevier B.V.

[7–9]. For example, broad differences in venom composition have been observed at multiple levels in rattlesnakes. In general, rattlesnake venoms can be classified as either type I or II according to the predominant expression of SVMP or PLA₂ phenotypes, respectively [10]. The occurrence of type I or II phenotypes has been reported in closely related species [11], within the same species (depending on their geographical distribution) [12,13] or during their ontogeny [14]. This phenotypic variation results in functional variation: the phospholipases in type II venoms exhibit neurotoxic activity leading to high levels of venom lethality [10] whereas type I venoms are considered coagulotoxic due to their high abundance in SVMPs, characterized as tissue-damaging digestive enzymes that affect the hemostasis of prey [15]. Consequently, these shifts in rattlesnake venom phenotypes alter the mechanisms by which prey are subdued (e.g. venoms with neurotoxic versus hemostatic effects) with clear consequences for the pathology of human envenomations.

In *Bothrops* pit vipers, PLA₂ molecules do not exhibit neurotoxic activity; instead, they show a tissue-damaging myotoxic activity [16]. Perhaps as a consequence, the majority of *Bothrops* species show a rattlesnake-like type I venom phenotype with high SVMP content [17], with only a few examples of type II venoms, such as *B. jararacussu* venom [18] which is usually less toxic (higher LD₅₀), at least in laboratory mouse models. Nevertheless, interspecific and intraspecific variability have also been reported for venoms from *Bothrops* species that show an SVMP-rich type I phenotype. Such variation arises through differences in the abundance of specific venom proteins in each toxin group, leading to a quantitative variability in venom activities. The studies on venom composition for *Bothrops* species report that this type of variation can be influenced by factors such as ontogeny [19,20], gender [21], and broad scale geographical variation [22,23]. However, to our knowledge, no previous studies in *Bothrops* have attempted to characterize venom variation among specimens belonging to the same ontogenetic stage, gender, and geographical location. The use of quantitative transcriptomics and proteomics approaches have been used for addressing questions about the nature of links between genotypes and phenotypes for adaptive traits like venom [24,25] and for devising strategies used for antivenom development [26].

Here, we used transcriptomic and proteomic approaches to characterize in detail venom variation for five individuals of the common lancehead (*B. atrox*) sampled from two relatively close populations located on the North and South banks of the Amazon River in the west of Pará State, Brazil. Previous analyses of pooled venom from individuals from these populations showed broad similarity in HPLC profiles but also differences in the sizes of specific venom peaks. Despite these differences, functional analyses of pooled venoms from these populations showed that they are indistinguishable in terms of their functional properties [27]. Other work has shown that these populations are genetically distinct and diverged from each other < 0.5 MYR bp but remained connected through low levels of gene flow [28]. Thus, these populations exhibit similar venom phenotype likely due to adaptation due to similar prey communities in similar habitats and therefore offer the opportunity to use our ability to dissect in detail the molecular variation in venom to assess whether common adaptive phenotypes have the same or different molecular bases [29].

In the present work, we have also, we quantified venom variation in unprecedented detail, both at individual and population levels, and then understood its molecular basis in terms of differences in expression and translation of different toxin proteins. We used transcriptomic and proteomic analyses to quantify patterns of paralogue/isoform expression in each venom followed by linking the observed paralogues with protein sequences, for which function had been previously determined, to understand the role of each isoform in determining the observed venom phenotype. We then used this information to assess three broad questions about the molecular basis of adaptive phenotypic variation using venom as model: 1) What role does expression and translation play in generating phenotypic variation in a complex molecular

adaptation? 2) Is there a difference in patterns of variation among different functional categories of venom proteins? 3) Do functionally similar venom phenotypes have the same or different molecular bases?

2. Material and methods

2.1. Snakes and venoms

Individual *B. atrox* adult female snakes were collected from two locations in the western part of the State of Pará, Brazil. Two specimens were collected from a recently cleared pasture area, which was previously an upland forest in the municipality of Oriximiná, on the North shore of the Amazon River (S 01°46'03.39"; W 55°43'53.31") and three specimens were collected in Floresta Nacional do Tapajós, a National Forest located in the municipality of Belterra, next to the Tapajós River, about 80 km south of the Amazon River (S 03°03'59.03"; W 54°58'8.94"), under ICMBio/SISBio license 32098-1. After capture, the snakes were immediately transferred to the Herpetarium of Faculdades Integradas do Tapajós, in Santarém, and anesthetized using CO₂ for venom extraction. Venom samples were collected using manual extraction techniques and individually freeze-dried for proteomic studies. For transcriptomic analysis, total RNA was extracted from venom gland tissue collected four days after venom extraction. The snakes were anesthetized with sodium pentobarbital (30 mg·kg⁻¹, s.c.) and decapitated. Venom glands were then dissected and immediately frozen in liquid nitrogen followed by storage at -80 °C for further mRNA isolation. Animal care and procedures used in the handling of snakes were in accordance with the guidelines of the Ethical Committee for Animal Research of Instituto Butantan (1244/14).

2.2. Venom fractionation by reverse phase chromatography

Individual *B. atrox* venoms were fractionated using reversed-phase high-performance liquid chromatography (RP-HPLC) following previously described methods [17]. Briefly, 5 mg of crude lyophilized venom was dissolved in 250 µL of 0.1% trifluoroacetic acid (TFA) and injected onto a Vydac C18 column (250 mm × 4.6 mm, 10 µm particle size) coupled to a Shimadzu LC 20 - AT HPLC system. Proteins were eluted at 2 mL/min with a gradient of 0.1% TFA in water (solution A) and 0.1% TFA in acetonitrile (solution B) (5% B for 5 min, 5–15% B over 10 min, 15–45% B over 60 min, 45–70% B over 10 min, 70–100% over 5 min, and 100% B for 10 min). The fractionation was monitored at 214 nm.

2.3. Proteomic characterization by shotgun mass spectrometry

Replicates of each individual venom sample (50 µg of protein) were reduced and alkylated before treatment with trypsin solution (0.2 µg/µL), as previously described [30]. The tryptic digests were desalted using in-house made columns packed with Poros R2 resin (Life Technologies, USA) and subjected to reversed-phase nanochromatography coupled to nanoelectrospray high resolution mass spectrometry for peptide analysis. Each digest was analyzed in triplicate in the mass spectrometer, as previously described for proteome analysis of several bothropic venoms [27].

Tandem mass spectra were processed and searched against an in-house database using the search tools Mascot (Matrix Science, London, UK; version 2.4.1) and X! Tandem (The GPM, thegpm.org; version CYCLONE (2010.12.01.1)). The database used to identify the MS/MS spectra is composed of the full-length precursor proteins predicted from the transcriptomes of the same five specimens of *B. atrox* described in this manuscript. Protein identification was based on the presence of proteotypic peptides relating to each venom protein isoform; search parameters were as described in the literature [27].

Scaffold (version Scaffold 4.6.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein

identifications. Peptide identifications were accepted if they could be established at > 99.0% probability by the Scaffold local FDR algorithm. Peptide identifications were also required to exceed specific database search engine thresholds, such as Mascot ion scores > 40.0 and/or X! Tandem $-\log(E\text{-value})$ scores > 2.0. Protein identifications were accepted if they could be established at > 99.0% probability to achieve an FDR < 1.0%. Protein probabilities were assigned by the Protein Prophet algorithm [31].

Quantitative values of protein for identifications at isoform and protein class levels for venom analyses (Supplementary Table 2) were expressed as exclusive spectral count, corresponding to the number of spectra attributed to proteotypic peptides of a given protein entry present in the database. Spectra for different charge states and/or different allowed modifications (e.g. deamidation of asparagine) related to proteotypic peptides were also computed.

2.4. Transcriptomic analyses

Total RNA was extracted from independent frozen glands homogenized in a tube containing Trizol (Invitrogen, Life Technologies Corp.) as described [32]. The mRNA was purified by affinity to magnetic beads containing oligo (dT) using the Dynabeads® mRNA DIRECT kit (Invitrogen, Life Technologies Corp) and used to prepare independent cDNA libraries for left and right venom glands from each snake. Ten cDNA libraries (2 per individual snake, right and left venom glands) were prepared following the protocol for TruSeq™ RNA Sample Preparation Kits v2 from Illumina, and sequenced using the HiSeq1500 Illumina technology, generating strand-specific paired-reads of 2×150 bp. The raw sequencing reads were pre-processed by an “in house” pipeline for sequencing quality control, to trim and remove reads with low-complexity and homopolymer enriched regions, poly-A/T/N tails, the adapter sequences and low-quality bases with the software fastq-mcf 1.04.662 [33] and bowtie2 2.2.5 [34]. Reads were filtered out if > 90% of a read corresponded to homopolymer or low-complexity regions, and they were trimmed if the mean quality score was lower than 25 in a window size equal 15. After trimming, all reads smaller than 40 bp were discarded.

After pre-processing the sequence reads, the remaining high quality sequences were used for transcriptome assembly and annotation. Two approaches were applied for assembly: we used the assembler SeqManNGen 11.2.1 (DNASTar Genomics Suite) following Rokyta et al. [35], using default parameters, with minimum identity of 90% and keeping only transcripts with at least 100 reads. Afterwards, we used the Extender program [36] for the De novo assembly, in multiple cyclic rounds, of seeds formed by 1000 random reads to extend the toxin transcripts partially generated by SeqManGen assembler. After the assembly and annotation, a database consisting of *Bothrops atrox* genes from specific toxin families was created from transcripts of the venom glands of five specimens of *B. atrox*, analyzed as described above. The MasterSet was curated using the CLC Genomics Workbench 7.5.2 program, the transcripts were aligned within each toxin family and groups of sequences showing > 99% identity were considered copies of the same gene.

To estimate the relative abundance of gene expression for specific transcripts, a subset of 1.5 million reads (paired-end) was created for each library and mapped back to the MasterSet genes using bowtie2 software [34], with parameters: “–no-mixed –no-discordant –gbar 1000 –end-to-end –q -X 800”. The expression profile analysis was initially performed considering each library individually (left and right glands separately). Nonetheless, since they showed similar results, the expression profile analysis was based on combined counts of the two glands of each individual, resulting in five total samples (one per individual). To calculate the number of RNA-seq reads or fragments derived from transcripts, we used the RSEM (RNA-seq by Expectation Maximization) software [37], and then computed transcript abundance (Expected count), and the normalized expression values as FPKM

(fragments per kilobase of transcript per million mapped reads). To estimate the differentially expressed transcripts in at least one pairwise sample comparison, we used the edgeR software [38] and used the Expected_count values option. To compare expression levels of different transcripts or genes across samples we performed an additional TMM (trimmed mean of M-values) scaling normalization that aimed to account for differences across all samples [38,39]. Given a set of differentially expressed transcripts, the normalized expression values (TMM-TPKM) for each distinct sequence was used to perform hierarchical clustering analysis, resulting in grouping together those samples that had similar expression profiles according to transcripts and to group transcripts with similar expression patterns across samples, which were at least fourfold differentially expressed with a false discovery rate of $1e-3$.

2.5. Sequence alignments and gene tree analyses

Prediction of biological function for specific *B. atrox* sequences was inferred using phylogenetic analysis to identify homologous venom sequences in other species for which detailed experimental functional analysis had been completed. Complete cDNA sequences were translated into their predicted proteins (expasy.ch, translate tools) and subjected to blastp search (ncbi.nlm.nih.gov) in order to find the closest primary-structure-related toxins which had their function experimentally characterized. Prior to phylogenetic constructions, the cDNA-translated amino acid sequences were partitioned into six groups according to their major toxin family and sub-classification within the family: 1 - PIII class SVMPs; 2 - PI class SVMPs; 3 - SVSPs, 4 - PLA2; 5 - CTL α -chains and; 6 - CTL β -chains. In each group, we included toxin sequences from other species' venoms with well characterized biological activities: Jararhagin, *B. jararaca* (P30431.1); Hemorrhagic Factor 3, *B. jararaca* (Q98UF9.3); Vascular apoptosis-inducing protein 1, *Crotalus atrox* (Q9DGB9.1); Berythracivase, *Bothrops erythromelas* (Q8UVG0.1); VaF1, *Vipera ammodytes ammodytes* (AJC52543); BaP1, *B. asper* (P83512.2); Atroxlysin-1, *B. atrox* (P85420.2); Batroxobin, *B. atrox* (AAA48553.1); PA-BJ, *B. jararaca* (P81824.2); KR-E-1, *Gloydius ussuriensis* (Q7SZE2.1); Basic PLA2 myotoxin III, *B. asper* (P20474.2); Basic PLA2 myotoxin II, *B. asper* (P24605.3); BthA-I-PLA2, *B. jararacussu* (Q8AXY1.1); Factor IX-binding protein, *Protobothrops flavoviridis* (Q7LZ71.1); Rhodocetin, *Calloselasma rhodostoma* (D2YW40.2); GPIB-binding protein, *B. jararaca* (Q9PSM5.1); Botrocetin, *B. jararaca* (P22030.2); Bothrojaracin, *B. jararaca* (Q56EB0.1); Agkicetin-C, *Deinagkistrodon acutus* (Q9DEA1.1); Jerdonibitin, *Protobothrops jerdonii* (D1MGU1.1); EMS16, *Echis multisquamatus* (Q7T2Q1.1). We also included other structurally-related, but phylogenetically more divergent proteins to root the phylogenetic trees. For each group of toxins, the sequences used as root were the following: for PIII class SVMPs, ADAM 2 from *Mus musculus* (Q60718); for PI class SVMPs, Jararhagin from *B. jararaca* (P30431); for SVSPs, thrombin from *Gekko gecko* (AAA49309); for PLA2, PA2c from *Notechis scutatus* (AAB20783); for CTL α -chains and β -chains, Lectin-C type from *Notechis scutatus* (ABP94117).

Protein sequences were aligned with the iterative refinement method E-INS-i using a BLOSUM62 scoring matrix as implemented in MAFFT v7.309 [40]. The best-fit models of evolution were selected for each alignment using corrected Akaike Information Criterion (AICc) implemented in PartitionFinder2 [41]. Maximum likelihood (ML) phylogenetic trees were estimated using RAxML v.8.2.3 [42]. The rapid bootstrap algorithm performed by RAxML (–f a) was used to generate 1000 pseudoreplications of non-parametric bootstrap aiming to estimate support values (BS) for each clade. This algorithm also conducts a complete search of ML using each 5th bootstrap tree as a starting tree for the rapid hill-climbing search (total of 200 starting trees). The non-parametric Shimodaira-Hasegawa-Like (SHL) test of topologies [43,44] was also performed as implemented in RAxML v.8.2.3 (–f E). BS and SHL tend to be considered as complementary methods to estimate branch supports [44], with SHL performing better on short branches or

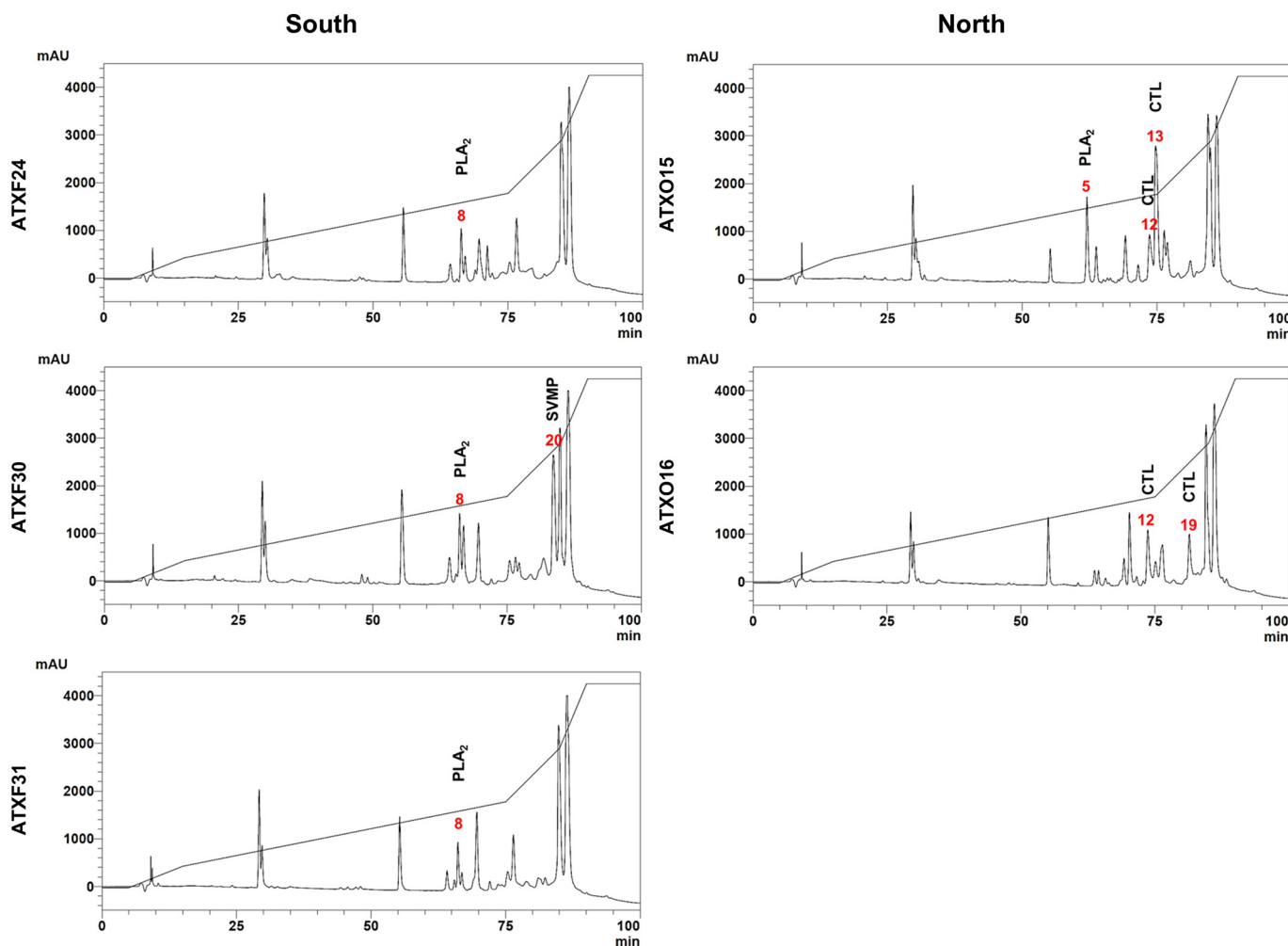


Fig. 1. Comparison of the elution profiles of venoms from individual snakes. Samples containing 5 mg of crude lyophilized venom from individual snakes were applied to a Vydac C-18 column (4.6 × 250 mm, 10-μm particle size) coupled to an Agilent 1100 HPLC system. The fractions were eluted at 2 mL/min, with a gradient of 0.1% TFA in water (solution A) and 0.1% TFA in acetonitrile (solution B) (5% B for 5 min, 5–15% B over 10 min, 15–45% B over 60 min, 45–70% B over 10 min, 70–100% B over 5 min, and 100% B for 10 min). The separations were monitored at 214 nm. Peaks that displayed noticeable changes are numbered and their most abundant protein family indicated according to the standard MS/MS analysis of *B. atrox* venom chromatographic fractions [27].

polytomies.

3. Results

3.1. Individual chromatographic profile comparisons

Comparisons of the chromatographic profiles of the venom from individual snakes collected from sites North (ATXO 15 and ATXO 16) and South (ATXF 24, ATXF 30 and ATXF 31) of the Amazon River indicate differences among individuals within and between these populations. In Fig. 1 we show the chromatograms obtained for each venom fractionation by RP-HPLC, performed according to a standard procedure already used for comparisons of other *B. atrox* venoms [17,27]. The venom profiles of the five snakes were conserved with respect to position of most peaks eluting at similar retention times. However, the heights and areas of some peaks were variable among the venoms, suggesting quantitative differences in the expression levels of some toxins. Although these comparisons were not subjected to statistical analyses, due to small sample size, some differences in % area of individual peaks could be noted among the venoms. For example, peak 12 was present only in the two samples from the North side of the Amazon (ATXO 15–4.7% and ATXO 16–7.3%) whereas peak 8 was evident in venoms from snakes collected on the South side of the river (ATXF

24–4.8%, ATXF 30–4.7% and ATXF 31–3.8%). Moreover, the venom of snake ATXO 15 was unique in showing detectable areas under peaks 5 (7.3%) and 13 (16.7%); ATXO 16 presented a larger peak 19 (5.7%); ATXF 30 was the only one with a high peak 20 (14.5%). These fractions had their composition identified in a parallel study from our laboratory [27] and include predominantly important toxins such as PLA₂s (fractions 5 and 8), CTLs (fractions 12, 13, and 19) and SVMPs (fraction 20). The data suggest differences in functionally relevant venom proteins among individual snakes and between the populations even though the overall functional properties of pooled venoms from each population are similar.

3.2. Transcriptome analyses

To assess the molecular mechanisms that underlie this variability, we analyzed the transcriptome of each snake venom gland and the individual venom proteomes, through the use of the translations of the transcriptomic datasets for protein identification. These data were used to evaluate transcriptional and translational levels of each independent toxin gene. High coverage transcriptomes of each specimen were obtained through Illumina cDNA sequencing (at least 60,000,000 reads for each snake) from ten cDNA libraries prepared from mRNA isolated from right and left glands of the snake. These sequences were assembled

and annotated for identification and quantification of specific toxin groups. After a preliminary assembling and annotation, we obtained 390 sequences related to toxins across all individuals. The sequences with at least 99% identity were grouped, resulting in a venom master-set of 152 different clusters (unigenes) composed of the complete sequences of all major toxin groups. The resulting unigenes were deposited in the GenBank TSA repository (Gene Bank SRA SRP056745) linked to Bioproject PRJNA279987 and individual sequences are under the accession numbers JAV01810.1 – JAV01961.1.

Distinct transcripts were present for major toxin groups: the largest number of paralogues was observed for the C-type lectins (CTL) group, which contained 29 and 26 sequences of alpha and beta chain representatives, respectively. Snake venom metalloproteinases (SVMP) accounted for 42 sequences, including 6 from P-I, 8 from P-II, and 28 for P-III classes of SVMPs. Twenty snake venom serine proteinase (SVSP) and 6 phospholipase A₂ (PLA₂) sequences were also identified. A reduced number of different transcripts were observed for Kunitz type proteins (KUNITZ - 3 sequences), vascular endothelial growth factor (VEGF - 5 sequences), and L-amino acid oxidase (LAAO - 2 sequences). Also, single transcripts of cysteine-rich secretory protein (CRISP), nerve growth factor (NGF), hyaluronidase (HYAL), nucleotidase (NUC), phosphodiesterases (PDE), and waprins-like proteins (WAP) were detected. Thirteen transcripts of bradykinin-potentiating peptides (BPP) were identified, but these corresponded mostly to truncated forms of similar sequences. The annotation of these sequences and their functional inferences is detailed in Supplementary Table 1.

The expression level for each transcript was evaluated by a normalized count of the reads mapping to each transcript (expected counts). Based on these data (Supplementary Table 2), we first analyzed the relative abundance of total transcripts coding for each of the toxin families in each snake, followed by a comparison of the expression level of each independent transcript. With respect to the abundance of the

toxin families, the overall composition of transcriptomes from glands of each snake was similar (Fig. 2), with a predominance of SVMPs (60.8–79.8%) followed by CTLs (4.9–19.5%), PLA₂s (4.4–7.3%), and BPPs (3.3–4.7%). SVSP (1.9–3.8%) and LAAO (1.6–2.8%) transcripts were less abundant in all glands with NUC, CRISP, PDE, and NGF contributing to < 1% of the total transcripts found in the venom glands. However, differences between individuals were also present such as the predominance of SVMPs transcripts in ATXO 15, ATXF 30, and ATXF 31 venom glands, with a major P-III class SVMP distribution in ATXF 31, and the highest proportion of P-I rather than P-III class SVMP transcripts in ATXO 15. In parallel, the percentage of CTLs was higher among ATXO 16 and ATXF 24 transcripts.

We then used heatmap analysis to assess similarities among the overall expression levels of venom glands, right and left glands from the same snake, and between individuals and populations. As shown in Fig. 3, there was a high correlation of expression between the right and left glands of the same snake, which confirms that there are no significant differences in expression between the left and right glands of the same snake. However, there were differences between individuals, especially between those from different populations on either side of the Amazon River, which suggests relevant differences in transcription levels between populations.

We then analyzed the patterns of transcription for individual paralogues in each snake (Supplementary Table 2). Among the 152 toxin sequences identified, only 28 had similar expression levels in all individuals. These transcripts corresponded to the groups represented by a single sequence (CRISP, HYAL, NGF, NUC, and PDE) or a few sequences (KUNITZ, LAAO, and VEGF). However, uniform expression was also noted for some of the transcripts coding for toxins belonging to the major toxin groups. Of interest, in these groups, the uniform distribution was restricted to the most abundant transcripts of PLA₂s (BATXPLA2), SVMPs (BATXSVMPPII28), and SVSPs (BATXSVSP2),

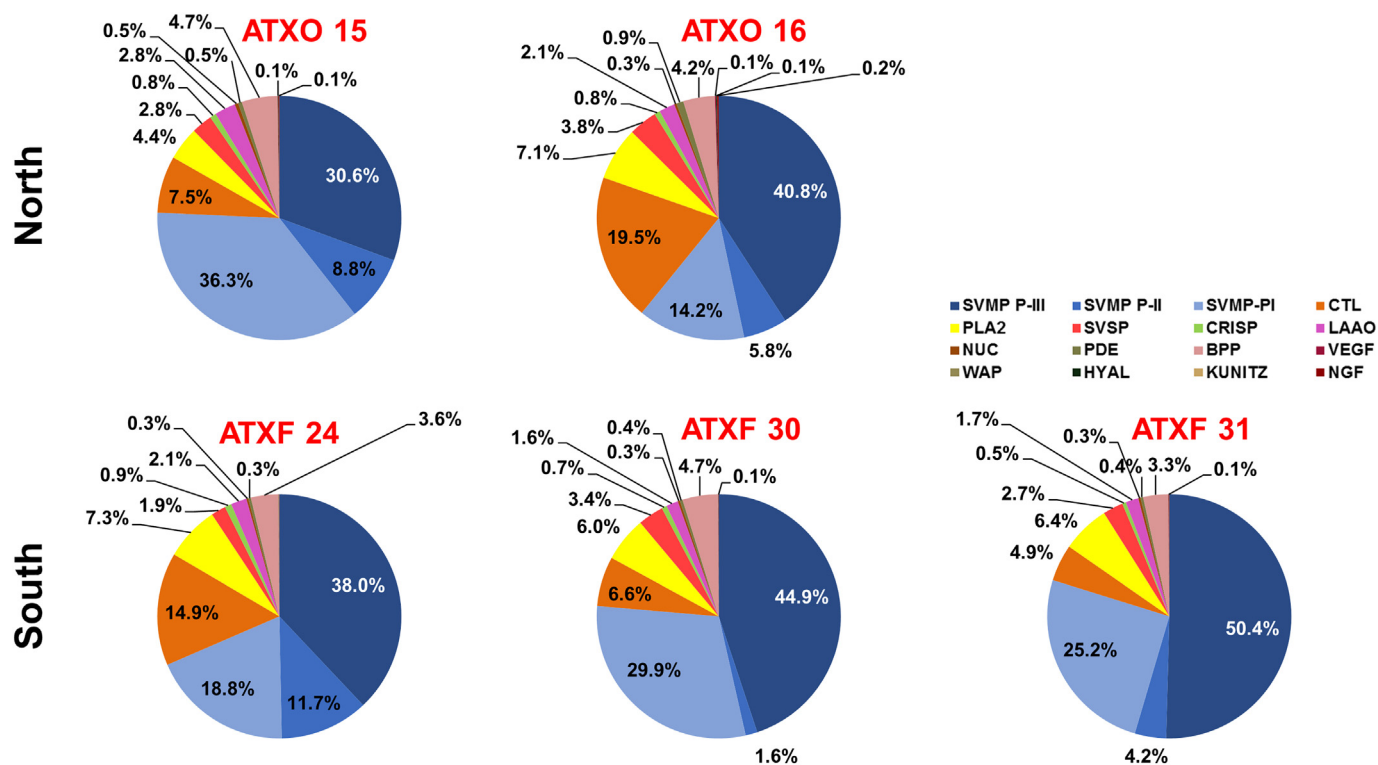


Fig. 2. Transcriptomic profile of the venom glands of *Bothrops atrox*. Relative expression indicated by the percentage of total expression values (expected counts) of transcripts identified in the snakes. Transcripts were grouped according to the toxin families: BPP - bradykinin-potentiating peptide; CRISP - cysteine-rich secretory protein; CTL - C-type lectin; HYAL - hyaluronidase; KUNITZ - Kunitz type protein; LAAO - L-amino acid oxidase; NGF - nerve growth factor; NUC - nucleotidase; PDE - phosphodiesterase; PLA₂ - phospholipase A₂; SVMP - snake venom metalloproteinase, classes PI, PII or PIII; SVSP - snake venom serine proteinase; VEGF - vascular endothelial growth factor; WAP - waprins-like protein.

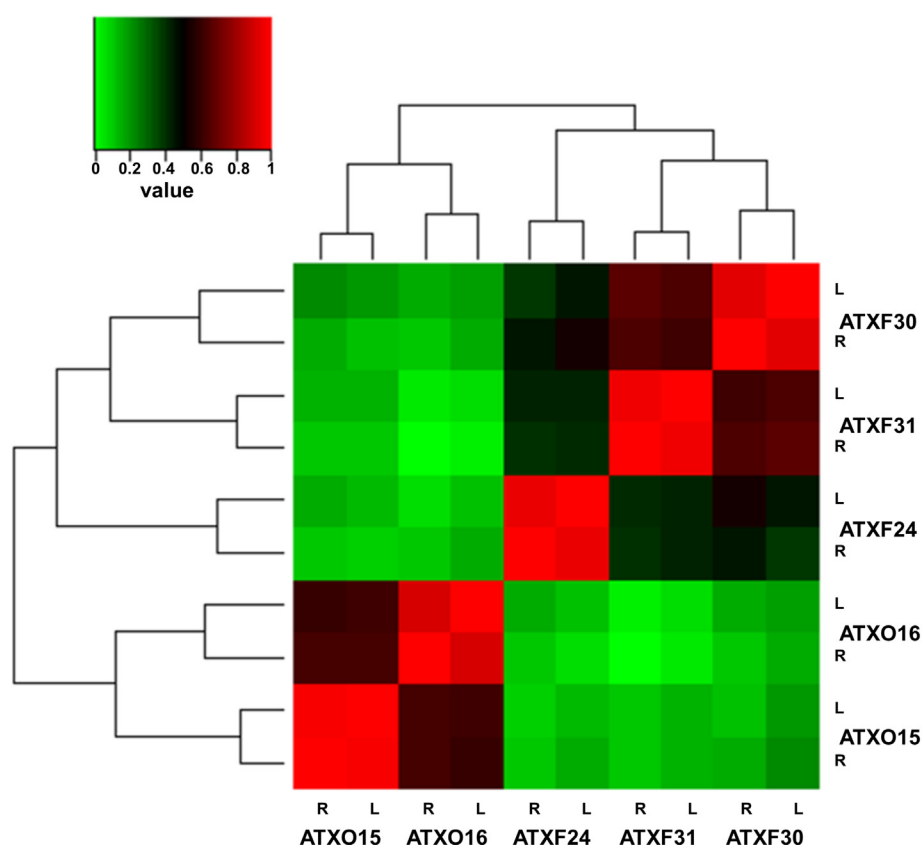


Fig. 3. Heatmap showing the hierarchically clustered Spearman correlation matrix resulting from comparing the transcript expression values (TMM-normalized FPKM) for each pair of samples.

whereas most of the less abundant paralogs showed more variable expression levels. The expression levels of the other 124 transcripts that were differentially expressed in the five subjects were then compared. In Fig. 4 we show the heatmap of the expression levels of individual transcripts for each snake. Differences in expression levels were observed both individually and among specimens from each population. According to the distribution of expression levels among individuals, transcripts were grouped in three major clusters, which included paralogues from all the major toxin families. The transcripts that made up the first cluster were common to venom glands of snakes collected on the south side of the Amazon River, while the transcripts from the second cluster were more prominent in the venom glands of northern snakes. Three sequences were grouped in cluster 3 since were almost entirely restricted to the snake ATXO 16. These observations support the idea that different paralogues are differentially expressed in the snakes from populations on the North or South sides of the Amazon River; however, we understand that the sample size used in this study is still limited and the feature should be further confirmed within a larger sample.

3.3. Correlations between transcription and translation

Our next step was to identify the proteins in the venom samples collected from each snake and to evaluate the correlation between levels of transcript expression and protein abundance. Venoms were subjected to “shotgun” proteomic analysis, in which approximately 9000 MS/MS spectra were identified in each sample using the transcriptome data described above as source for protein identification (Supplementary Tables 2 and 3). The relative distribution of the identified groups of toxins is shown in Fig. 5 and the correlation between levels of transcription and expression in the venoms is presented in Fig. 6. All the major toxin groups represented in the transcriptomes

were detected in the proteomes, and SVMs were also the most abundant proteins identified; however, their relative abundance was lower, when compared to other protein groups, than what was observed in the transcriptome data. In contrast, the relative percentage of CTLs, SVSPs, and LAOs in the venom (Fig. 5) increased significantly as compared to the transcriptomes (Fig. 2). The scatter plot in Fig. 6 shows a correlation between the total number of transcripts encoded by each gene family and their cumulative proteomic abundance in each snake specimen analyzed. For almost all comparisons, the toxin families abundant in the individual transcriptome were also abundant in the same individual proteome as well as toxin families with low abundance of transcripts were poorly represented in the proteome. The exception was the BPP protein family. The transcripts coding for the members of this family predict precursor proteins that undergo post-translational processing leading to small bioactive peptides, which are not optimally detected by the MS analysis and data interpretation pipeline used in this work. This was clearly demonstrated in the literature for the determination of *Bothrops jararaca* venom proteopeptidome, where the authors used distinct mass spectrometry acquisition and data analysis conditions for the proteomic and peptidomic analyses [45]. When analyzing the correlation between transcription and venom abundance of each gene paralog expressed in each snake, the lack of correlation between transcriptomic and proteomic levels was evident across all five snakes (Fig. 7), suggesting that mechanisms affecting protein translation, such as changes in mRNA/protein turnover, may act differentially on each transcript within a single toxin family. The most striking differences were noticed for CTL paralogs, for which a number of relatively highly transcribed genes exhibit little to no proteomic representation. In contrast, the toxins with highest expression levels, such as SVMs, showed similar transcriptional and translational activities (Fig. 7).

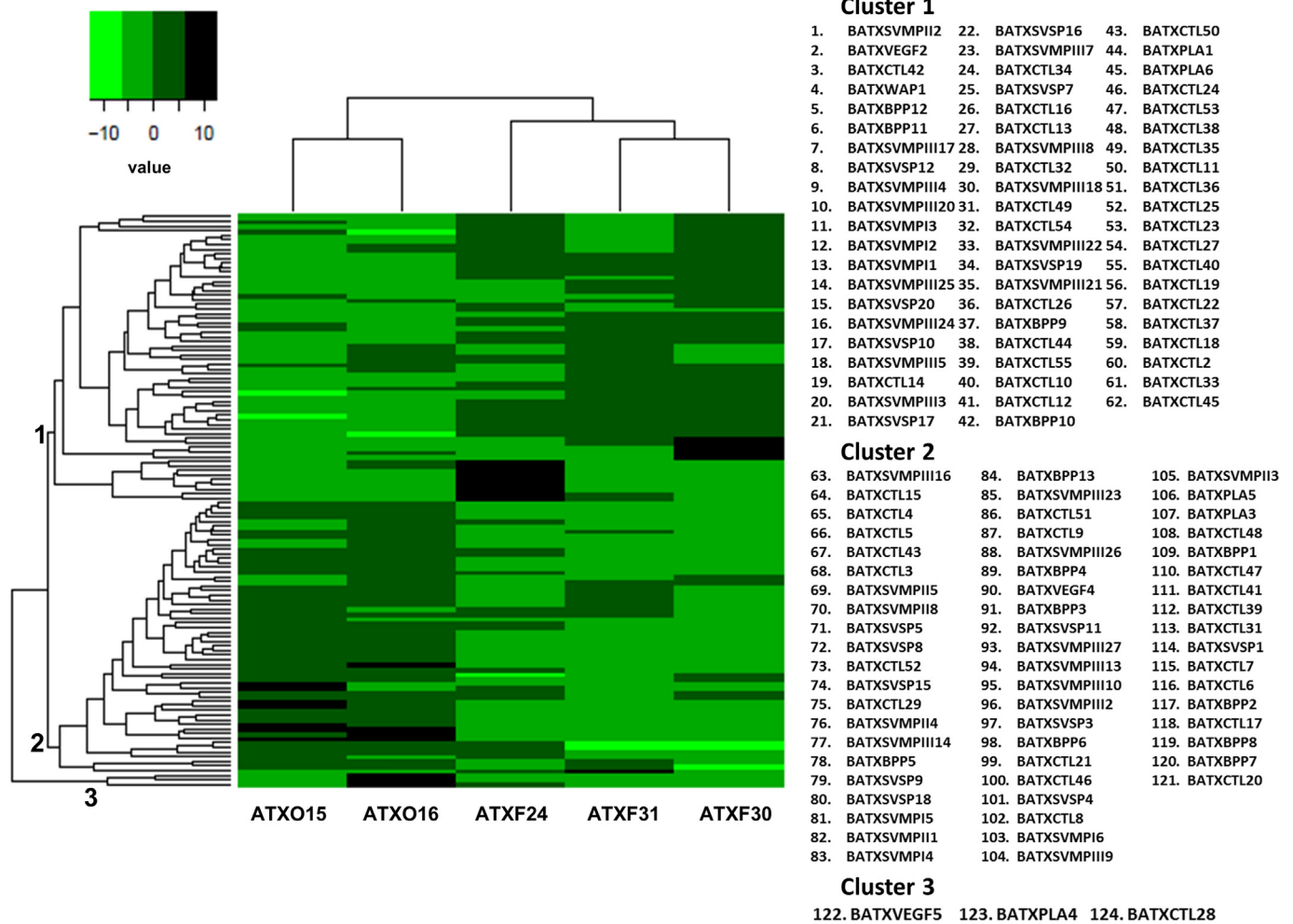


Fig. 4. Hierarchical clustering of transcripts and samples. Shown is a heatmap with the relative expression levels of each transcript (rows) in each sample (column). Rows and columns are hierarchically clustered (WARD method). Expression values (TMM-normalized FPKM) are log2-transformed and then mean-centered by transcript.

3.4. Functional inferences

The next step was to evaluate whether the functional role of each toxin could explain differential expression of each paralog. For this purpose, we interpreted close similarity of any *B. atrox* transcripts with transcripts from other species as evidence for conserved venom function across species. Thus, we selected sequences of each paralog of the main toxin groups (SVMPs, SVSPs, CTLs, and PLA₂s) and compared to sequences of toxins with well-established functions in other venomous snakes as assessed by previous studies using experimental approaches. In this regard, it is important to note that SVMPs, SVSPs, CTLs, and PLA₂s are the most abundant toxin families in snake venoms and also the proteins commonly related to impacts on prey hemostasis. *Bothrops* venoms are understood as coagulotoxic and other low abundance toxin families in the venom, such as CRISPs, LAAOs, and other minor enzymatic components have not yet been related to serious disturbances in the coagulation system, at least for mammals, and were analyzed in this part of our study.

Fig. 8 shows the ML tree that best represents the phylogenetic relationships between PIII class SVMP sequences identified in this study, including sequences of functionally-characterized proteins, along with the proteomic expression level of each paralog. The most abundant *B. atrox* paralog, BATXSVMPIII 28, which was equally abundant in all snakes, groups together with the previously well-characterized *Bothrops* SVMP jararhagin with high SHL and bootstrap values (100 and 99,

respectively) to support that claim. This group is also supported as an independent clade by high values of SHL (Fig. 8). BATXSVMPIII 28 presents 96% identity with *Bothrops* SVMP jararhagin, and conversely the identity varies between 54% and 68% within other sequenced paralogs. The sequence predicted by BATXSVMPIII 28 was identical to the sequence experimentally defined for batrotoxrhagin, a hemorrhagic SVMP recently isolated in our laboratory from venoms of *B. atrox* collected in natural environment or kept under captivity [46]. Jararhagin is a P-III class SVMP isolated from *B. jararaca* venom [47] which is a versatile toxin displaying many different biological functions, such as hemorrhagic, pro-inflammatory, anti-platelet, and apoptosis inducing activities [48]; batrotoxrhagin also shows similar activities [46]. Interestingly, jararhagin-like SVMPs were isolated from other venoms from *Bothrops* snakes such as *B. atrox* and *B. alternatus* [46,49]. In a blast search, proteins with high identity to jararhagin were also detected in venoms from other vipers, including *Crotalus durissus* (Q2QA02.1), *Crotalus viridis* (C9E1R8.1), *Crotalus atrox* (Q9082.1), *Gloydus brevicaudus* (P0C7B0.2), and *Deinagkistrodon acutus* (Q9W6M5.2). The presence of similar and possibly homologous proteins in different genera suggests that the high expression level of this gene copy may be an adaptive advantage among some viper snakes. The remaining branches include sequences with a greater number of paralogous transcripts but lower expression levels (Fig. 8). Differential expression levels in the venoms were clearly noted in the branches with sequences similar to HF3 and berythracinase. HF3 is a highly hemorrhagic SVMP from *B.*

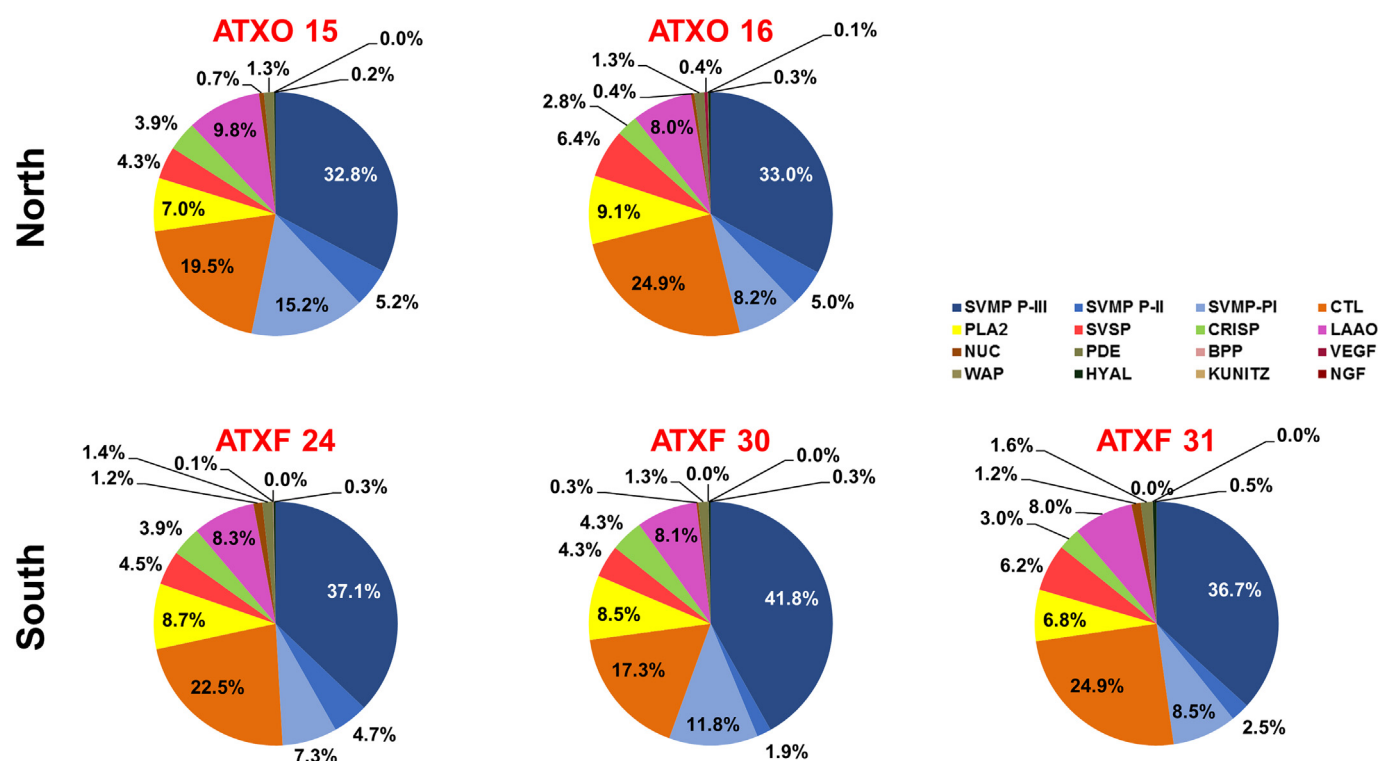


Fig. 5. Proteomic profile of the individual venom samples. Relative expression indicated by the percentage of total expression values (exclusive spectral count) of toxins identified in the venoms. Toxin isoforms were grouped according to the toxin families: BPP - bradykinin-potentiating peptide; CRISP - cysteine-rich secretory protein; CTL - C-type lectin; HYAL - hyaluronidase; LAAO - L-amino acid oxidase; NGF - nerve growth factor; NUC - nucleotidase; PDE - phosphodiesterase; PLA₂ - phospholipase A₂; SVMP - snake venom metalloproteinase, classes PI, PII or PIII; SVSP - snake venom serine proteinase; VEGF - vascular endothelial growth factor.

jararaca venom [50] while berythracivase is a non-hemorrhagic procoagulant SVMP, from *B. erythromelas* venom [51]. These SVMPs interact with different physiological targets enhancing the overall spectrum of action of SVMPs. Interestingly, a greater number of HF3- and berythracivase-related transcripts were found in the venom, most of them with differential expression levels. These observations suggested to us that the expression of each paralog may be controlled by different mechanisms in which the functional importance of the toxin may be a relevant factor that results in its individual expression level tuning.

Similar patterns arose from the analysis of the other groups of toxins indicating that the most expressed venom toxins are limited to a few paralogs showing homology to functionally characterized toxins. In

class P-I SVMPs (Supplementary Fig. 1), only one paralog (BATXPI3) had significant and uniform expression levels in all venoms and was homologous to atroxlysin-I, a hemorrhagic P-I class SVMP that contributes significantly to the biological activity of *B. atrox* venom [52,53]. With respect to SVSPs (Supplementary Fig. 2), the paralog (BATXSVSP2) coding for an isoform of batroxobin [54] was the most expressed. Interestingly, two other copies (BATXSVSP3 and BATXSVSP1) grouped together with BATXSVSP2 in a highly supported clade, but were only expressed in snakes collected on the northern shore of the Amazon River, ATXO 15 and ATXO 16. In the other branches of the SVSP tree, three other paralogs showed noticeable expression in the venoms, one of them, BATXSVSP13, is homologous to KR-E-1, a kinin-

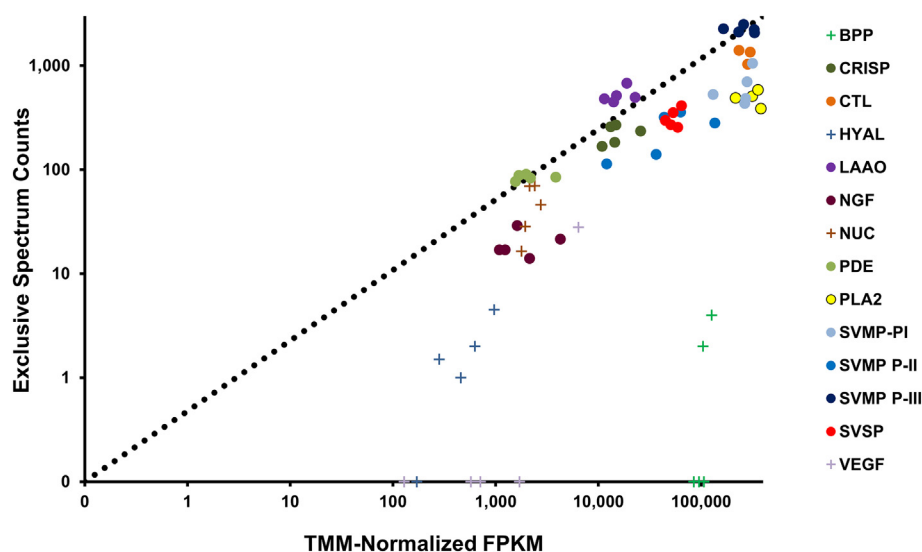


Fig. 6. Comparisons of gene transcription (TMM-normalized FPKM) in the venom gland and protein abundance (exclusive spectral counts) in the secreted venom of each of the five individual snakes analyzed. Each spot represents detected transcription and translation levels for a specific toxin family from one specimen, as described in Fig. 5. For visual purposes, dotted lines indicate equal gene and protein abundance.

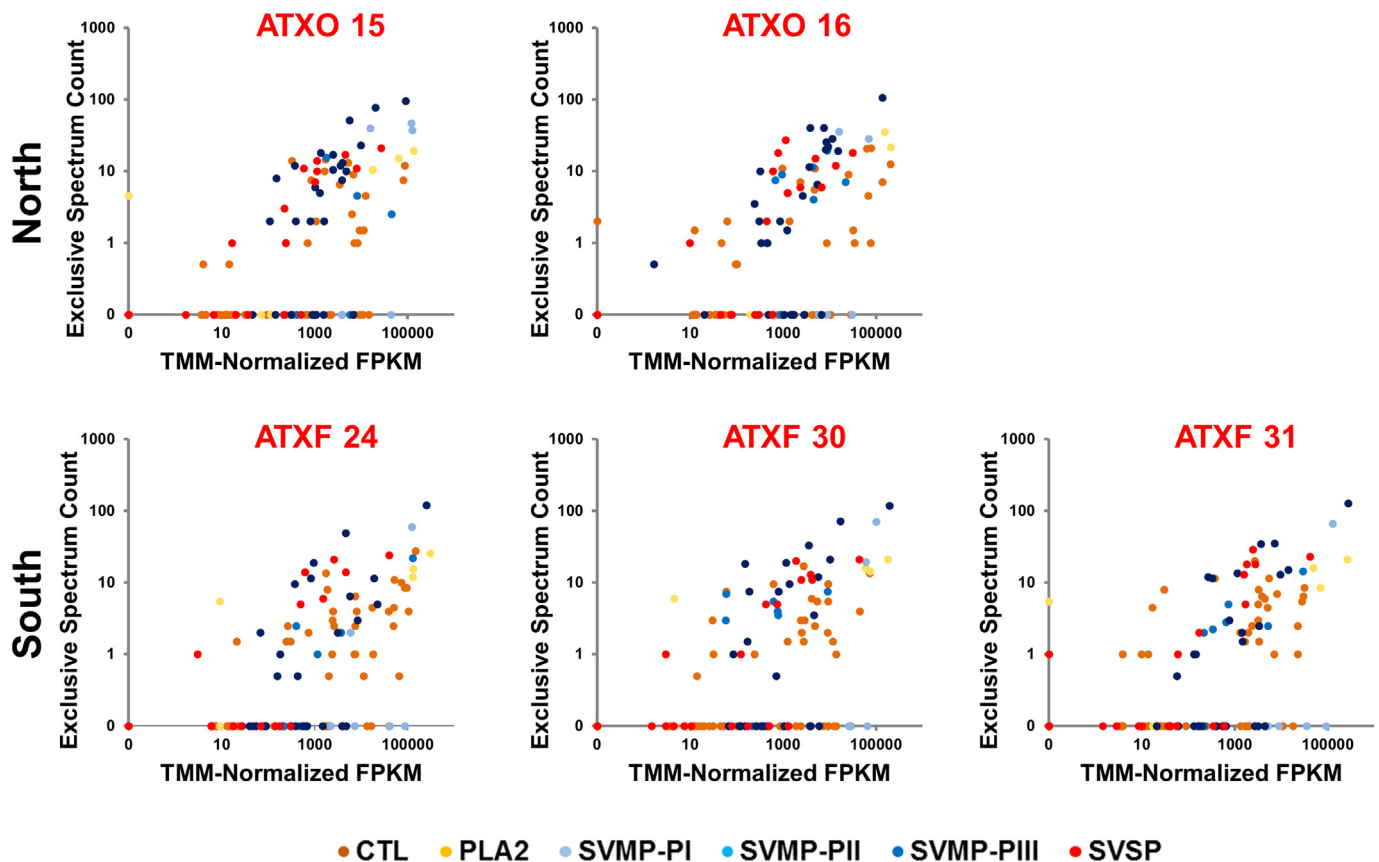


Fig. 7. Comparisons of individual gene paralogs abundance for each toxin family for each venom sample. Comparisons of gene transcription (TMM-normalized FPKM) in the venom gland and protein abundance (exclusive spectral count) in secreted venoms are represented independently for each snake and classification of each paralog/isoform according to their toxin families are represented by the color of the spots, as indicated in the figure.

releasing enzyme [55] that contributes to differential physiological targets of SVSPs. With respect to PLA₂s (Supplementary Fig. 3), all snakes showed substantial expression levels of the BATXPLA2 transcript, which is homologous to Mtx-II from *B. asper* venom and other K49 PLA₂, catalytic inactive myotoxins, widespread in venoms of *Bothrops* snakes [56]. Other paralogue, BATXPLA6, coding for the catalytically active D49 myotoxic PLA₂ homologous to MtxIII from *B. asper* venom [56], was highly abundant in individuals from the southern population while the expression of BATXPLA4, homologous to BthAI, predominated in one individual of the northern population.

The CTL group showed the highest variability in all parameters evaluated. For the CTL group, we found the highest number of transcripts with relevant expression levels in the venom (Supplementary Fig. 4A and B), with two transcripts with high expression levels coding for alpha and beta chains of Bothrojaracin and one transcript coding for Bothrocetin alpha chain, only. However, differently from the other groups, at least 7 paralogs coded for α -chains and 9 for β -chains homologous to Bothrojaracin α - and β -chains, respectively, with low and variable level of expression in each individual snake. Bothrojaracin is a potent thrombin inhibitor isolated from *B. jararaca* venom [57]. Considering the expression of β -chain transcripts, only BATXCTL14, responsible for the β -chain of Bothrojaracin, was expressed in all tested individuals. Expression of all the remaining transcripts was low and restricted to a few individuals in both sides of the Amazon River.

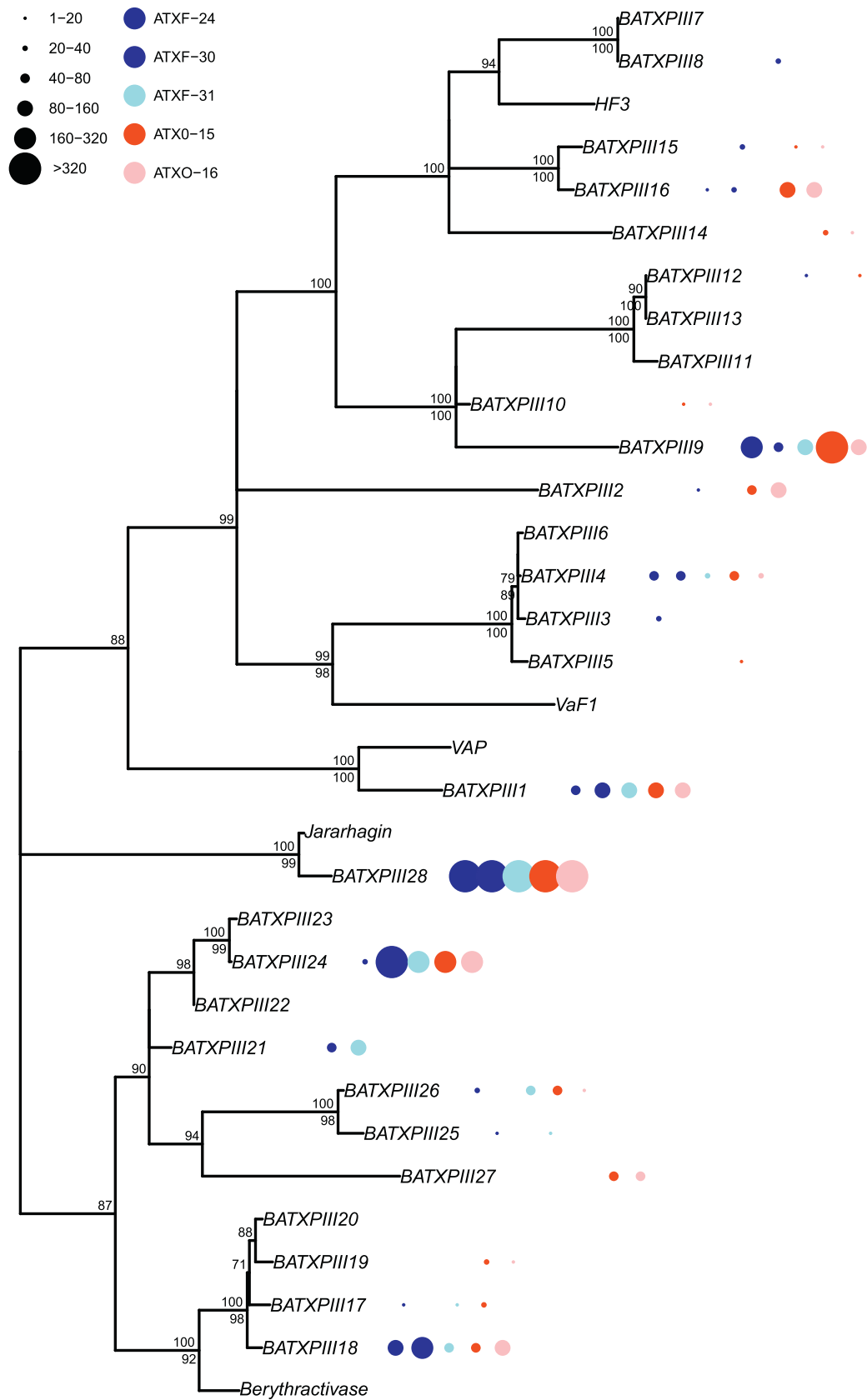
In summary, for each toxin family we detected the presence of isoforms with sequence similarity to functionally important toxins that are widespread in venoms of diverse viperid snakes and which were present in high abundance in the venom proteomes of all five snakes analyzed. We suggest these isoforms act as generalized core function toxins, which have high expression levels in the venoms of viperid

snakes and are responsible for key functional activities. In contrast, a greater number of different transcripts in each toxin family are low and differentially expressed, making a greater contribution to venom variability, possibly serving as a reservoir of adaptive variation in the toxic venom arsenal that can be used in rapid adaptive shifts in response to environmental changes: these are adaptive toxins.

4. Discussion

Previous studies have documented variability in composition of *B. atrox* venom over broad geographic scales but with a limited understanding of the molecular mechanisms responsible for this variation and its functional significance. For example, important shifts during ontogenetic development are observed in populations from Colombia and Venezuela for which the venom profile of juveniles is dominated by PIII-class SVMPs while adult snakes express predominantly PLA₂s and PI-class SVMPs [22,58]. In the Brazilian Amazon, *B. atrox* snakes do not undergo ontogenetic shift show the PIII-class SVMP-rich venom characteristic of juveniles from Colombia and Venezuela [22]. Our study, although limited in sample size, provides the most detailed analysis of venom variation within this species to date. It shows that there is significant variation in expression levels of venom proteins among individuals and between geographically close populations on either side of the Amazon River.

We observed quantitative differences in several groups of toxins among the five individuals and these differences occurred even between individuals within the same population (Fig. 1). Toxin groups that play recognized roles in venom toxicity such as SVMP, SVSP, PLA₂, and CTL showed greater representation in the transcriptomes than protein groups such as LAAO, CRISP, VEGF, NGF, HYAL, NUC, PDE, KUNITZ,



(caption on next page)

Fig. 8. Functional inference of PIII-class SVMP isoforms present in the venoms of the different snake specimens. Complete deduced protein sequences of PIII-class SVMPs were aligned together with sequences of jararhagin, *Bothrops jararaca* venom (P30431.1); hemorrhagic factor 3, *B. jararaca* venom (Q98UF9.3); vascular apoptosis-inducing protein 1, *Crotalus atrox* venom (Q9DGB9.1); berythracivase, *B. erythromelas* venom (Q8UVG0.1); VaF1, *Vipera ammodytes ammodytes* venom (AJC52543), which are PIII-class SVMPs isolated from different venoms with experimentally defined biological activities. The maximum likelihood phylogenetic tree was generated using RAxML. Numbers above and below branches represent SHL and bootstrap support values, respectively. Values smaller than 70% are not showed and nodes presenting values of SHL smaller than 70% are collapsed to a polytomy. The size of the circles on the right indicates the number of exclusive spectrum counts of MS/MS in each venom sample, which are represented by the different colors of the circles, as indicated in the figure.

and WAP (not yet related to major venom effects within this species) that showed a low number or single transcripts (Fig. 2). This observation supports some previous assumptions that rates of gene duplication are higher in functionally important gene families than in the so-called “ancillary” toxin families, as LAAO and CRISP or other groups not so important for venom toxicity [59].

With respect to transcript expression levels, our analysis revealed a highly variable pattern among individuals (Figs. 3 and 4). Notably, the most extensive quantitative variation was observed in the larger multi-transcript groups, whereas the small (or single) transcript toxin groups approached a uniform distribution among the individuals. However, an interesting aspect observed was a variation of the number of transcripts within each of the toxin groups. For SVMP, SVSP, and PLA₂, the most abundant transcripts were uniformly present among the transcriptomes of the five snakes. Likewise, these transcripts coded for abundant isoforms in venom proteomes (Fig. 7). We interpret this as representing expression patterns of “core function” venom loci, which encode multifunctional proteins that would be functionally important across species for effectively subduing, killing or digesting different prey. This is the case of BATXSVMPIII 28 that codes for the multifunctional toxin batroxhagin [46]. However, the wide range of targets that these toxins interact with may imply in medium to low affinity binding to prey receptors. In this case, the presence of other highly potent toxins would be essential for a fine tuning of receptors interaction and prey specificity in different environments. This function could be related to isoforms BATXSVMPIII 7 or 8, which are similar to HF3, a highly hemorrhagic SVMP [50]. Interestingly, the “core function” loci are also the genes coding for toxins that undergo ontogenetic shifts in *B. atrox* [20], supporting the hypothesis that they are controlled by distinct promoters, that may also be sensitive to the action of hormones involved in ontogenetic differentiation.

In terms of functional significance, the predominant toxin transcript/toxin isoform detected in the transcriptome/proteome corresponds to batroxhagin, a multifunctional hemorrhagic PIII-class isolated from venoms of *B. atrox* snakes from nature or captivity [46]. Batroxhagin shows functional and structural similarity to key toxins from venoms of different species from *Bothrops*, *Crotalus* and *Deinagkistrodon* genera: (1) 96% identity with the *B. jararaca* SVMP jararhagin [48]; (2) 87% identity with catrocollastatin from *Crotalus atrox* [60]; (3) 85% identity with acutolysin from *Deinagkistrodon acutus* venom [61]. In addition, the most abundant SVSP transcript, also uniformly expressed among the individuals, codes for batroxobin, a fibrinogenolytic toxin isolated from *B. atrox* venom that has been used in clinical approaches [54], and that is 95% identical to bothrombin, a fibrinogenolytic enzyme isolated from *B. jararaca* venom [62]. The predominant PLA₂ transcript codes for a toxin with 90% identity to bothropstoxin-I, a K49 phospholipase *B. jararacussu* homologue responsible for venom myotoxicity [63], and that was also uniformly expressed and present in equivalent rates for transcription/translation. Thus, in contrast to the perception that many venom genes are under positive diversifying selection [64], our results suggest that genes coding for core function toxins are apparently highly conserved in venom glands of viperid snakes, supporting our hypothesis that different mechanisms of gene evolution are acting depending on the function of the isoform encoded by each paralogue.

CTLs in *B. atrox* showed a different pattern of molecular variation. This group of toxins accounted for a large number of transcripts (29 for

α- and 26 for β-chains), indicating that gene duplications are also occurring in the genes of this toxin family. However, most transcripts presented low expression levels and were differentially expressed among the individuals tested, encountering low correlation between transcription/translation rates. Only a few transcripts code for proteins with high identity to toxins already isolated and identified in other venoms. Transcripts BATXCTL 39, 40, and 41 and BATXCTL 6 and 7 code for the α and β chains, respectively, of a toxin with > 90% identity to botrocetin, which was isolated from *B. jararaca* venom, and is able to affect platelet functions [65]. Botrocetin-like coding transcripts were observed predominantly in the snakes collected on the northern margin of the Amazon River and may represent an intrinsic characteristic of this population of *B. atrox* snakes. Notwithstanding these two botrocetin-like transcripts, most of the detected CTL transcripts code for proteins with low identity levels (< 80%) to toxins already isolated and identified in other venoms, suggesting that CTLs are not conserved in venoms of viperid snakes.

Our results have evolutionary implications for understanding the evolution of venom as an adaptive molecular phenotype. First, they provide evidence for a positive link between levels of expression and conserved function in venom proteins. This supports the hypothesis, first proposed by [4], that broadly speaking the complex venom phenotype contains two functional classes of toxins that are related to differences in their functional role in the overall venom phenotype in terms of killing and digesting different prey [66]. In particular, high abundance venom proteins found in most snakes in most populations (labeled here as “core function” venom proteins) may be analogous to highly expressed “housekeeping” proteins which perform a variety of essential cellular functions [67] and perform generic killing and digestive functions that are not prey specific. In contrast, toxins present at low abundance (e.g. CTLs) may be more plastic in either evolutionary or ecological timescales and serve to “customize” an individual snake's venom to feeding on particular prey requiring a specific venom protein [68].

Furthermore, our results suggest a molecular mechanism for individual and population variation that implies the use of different gene loci or alternate expression of different venom alleles in different populations, as described in other studies [69,70]. The lack of correlation between transcriptomic and proteomic levels indicates that mechanisms affecting protein translation may also influence the expression of the venom phenotype. The demonstration of this work adds to the emerging view that expression and translational differences play a greater role in defining adaptive variation in venom phenotypes than does sequence variation in protein coding genes [25].

Finally, the fact that whole venoms in each genetically distinct population are functionally similar, yet show substantial variation in transcript identity and abundance, provides evidence that adaptive phenotypes may have different molecular bases in terms of the identity and expression levels of specific transcripts. In other words, different combinations of transcripts and expression levels may be responsible for the same adaptive phenotype in different populations. This supports the general idea that, even at the intraspecific level, similar adaptive phenotypes can have different molecular mechanisms [71,72].

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jprot.2018.03.032>.

Funding

This study was supported by CAPES (063/2010-Toxinology – AUXPE 1209/2011, 1224/2011, and 1519/2011), FAPESP (2012/16277-9, 2014/26058-8 and 2016/50127-5) and CNPq (131820/2014-1). DRA was a Ph.D. student in the Science Graduate Program – Toxinology (Instituto Butantan), with a CAPES scholarship (BEX no. 2520/14-1) and CAN in the Cellular and Molecular Biology Graduate Program (IOC, FIOCRUZ), with a CAPES scholarship (AUXPE 1214/2011). JAPJ is recipient of a fellowship from CAPES (AUXPE 1209/2011).

Data accessibility

The resulting unigenes were deposited in the GenBank TSA repository (Gene Bank SRA SRP056745) linked to Bioproject PRJNA279987 and individual sequences are under the accession numbers JAV01810.1 – JAV01961.1.

Author contributions

AMMS, ILMJA, RHV designed research; DRA, JAPJ, MYNJ, CAN, DRR, HMC performed research; AMMS, ILMJA, RHV, DRA, RHVM, JAPJ, MYNJ, FGG, DRR, HLG analyzed data; AMMS, ILMJA, DRA, RHV, HLG wrote the paper; RHVM, DRA, HMC collected and maintained snakes.

References

- [1] J.P. Chippaux, V. Williams, J. White, Snake venom variability: methods of study, results and interpretation, *Toxicon* 29 (1991) 1279–1303.
- [2] B.J. Wooldridge, G. Pineda, J.J. Banuelas-Ornelas, R.K. Dagda, S.E. Gasanov, E.D. Rael, C.S. Lieb, Mojave rattlesnakes (*Crotalus scutulatus scutulatus*) lacking the acidic subunit DNA sequence lack Mojave toxin in their venom, *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 130 (2001) 169–179.
- [3] S.P. Mackessy, Venom composition in rattlesnakes: trends and biological significance, in: W.K. Hayes, K.R. Beaman, M.D. Cardwell, S.P. Bush (Eds.), *The Biology of Rattlesnakes*, Loma Linda University Press, Loma Linda, CA, USA, 2008, pp. 495–510.
- [4] H.L. Gibbs, S.P. Mackessy, Functional basis of a molecular adaptation: prey-specific toxic effects of venom from *Sistrurus rattlesnakes*, *Toxicon* 53 (2009) 672–679.
- [5] N.R. Casewell, W. Wüster, F.J. Vonk, R.A. Harrison, B.G. Fry, Complex cocktails: the evolutionary novelty of venoms, *Trends Ecol. Evol.* 28 (2013) 219–229.
- [6] J.J. Calvete, Snake venomomics: from the inventory of toxins to biology, *Toxicon* 75 (2013) 44–62.
- [7] J.J. Calvete, G.B. Domont, Omic technologies to fight the neglect, *J. Proteome* 74 (2011) 1483–1484.
- [8] J.M. Gutiérrez, B. Lomonte, G. León, A. Alape-Girón, M. Flores-Díaz, L. Sanz, Y. Angulo, J.J. Calvete, Snake venomomics and antivenomics: proteomic tools in the design and control of antivenoms for the treatment of snakebite envenoming, *J. Proteome* 72 (2009) 165–182.
- [9] D.J. Williams, J.M. Gutiérrez, J.J. Calvete, W. Wüster, K. Ratanabangkoorn, O. Paiva, N.I. Brown, N.R. Casewell, R.A. Harrison, P.D. Rowley, M. O'Shea, S.D. Jensen, K.D. Winkel, D.A. Warrell, Ending the drought: new strategies for improving the flow of affordable, effective antivenoms in Asia and Africa, *J. Proteome* 74 (2011) 1735–1767.
- [10] S.P. Mackessy, Evolutionary trends in venom composition in the western rattlesnakes (*Crotalus viridis sensu lato*): toxicity vs. tenderizers, *Toxicon* 55 (2010) 1463–1474.
- [11] J.J. Calvete, E. Fasoli, L. Sanz, E. Boschetti, P.G. Righetti, Exploring the venom proteome of the western diamondback rattlesnake, *Crotalus atrox*, via snake venomomics and combinatorial peptide ligand library approaches, *J. Proteome Res.* 8 (2009) 3055–3067.
- [12] D.J. Massey, J.J. Calvete, E.E. Sánchez, L. Sanz, K. Richards, R. Curtis, K. Boesen, Venom variability and envenoming severity outcomes of the *Crotalus scutulatus scutulatus* (Mojave rattlesnake) from southern Arizona, *J. Proteome* 75 (2012) 2576–2587.
- [13] J. Durban, P. Juárez, Y. Angulo, B. Lomonte, M. Flores-Díaz, A. Alape-Girón, M. Sasa, L. Sanz, J.M. Gutiérrez, J. Dopazo, A. Conesa, J.J. Calvete, Profiling the venom gland transcriptomes of Costa Rican snakes by 454 pyrosequencing, *BMC Genomics* 12 (2011) 259.
- [14] J. Durban, A. Pérez, L. Sanz, A. Gómez, F. Bonilla, S. Rodríguez, D. Chacón, M. Sasa, Y. Angulo, J.M. Gutiérrez, J.J. Calvete, 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.
- [15] A.M. Moura-da-Silva, D. Butera, I. Tanjoni, Importance of snake venom metalloproteinases in cell biology: effects on platelets, inflammatory and endothelial cells, *Curr. Pharm. Des.* 13 (2007) 2893–2905.
- [16] A.M. Moura-da-Silva, H. Desmond, G. Laing, R.D.G. Theakston, Isolation and comparison of Myotoxins isolated from venoms of different species of *Bothrops* snakes, *Toxicon* 29 (1991) 713–723.
- [17] L.F. Sousa, C.A. Nicolau, P.S. Peixoto, J.L. Bernardoni, S.S. Oliveira, J.A. Portes-Junior, R.H. Mourão, I. Lima-dos-Santos, I.S. Sano-Martins, H.M. Chalkidis, R.H. Valente, A.M. Moura-da-Silva, Comparison of phylogeny, venom composition and neutralization by antivenom in diverse species of *Bothrops* complex, *PLoS Negl. Trop. Dis.* 7 (2013) e2442.
- [18] C. Correa-Netto, R. Teixeira-Araujo, A.S. Aguiar, A.R. Melgarejo, S.G. De-Simone, M.R. Soares, D. Foguel, R.B. Zingali, Immunome and venom of *Bothrops jararacussu*: a proteomic approach to study the molecular immunology of snake toxins, *Toxicon* 55 (2010) 1222–1235.
- [19] A. Zelanis, A.K. Tashima, A.F. Pinto, A.F. Leme, D.R. Stuginski, M.F. Furtado, N.E. Sherman, P.L. Ho, J.W. Fox, S.M. Serrano, *Bothrops jararaca* venom proteome rearrangement upon neonate to adult transition, *Proteomics* 11 (2011) 4218–4228.
- [20] R.A. Guércio, A. Shevchenko, J.L. López-Lozano, J. Paba, M.V. Sousa, C.A. Ricart, Ontogenetic variations in the venom proteome of the Amazonian snake *Bothrops atrox*, *Proteome Sci.* 4 (2006) 11.
- [21] M.C. Menezes, M.F. Furtado, S.R. Travaglia-Cardoso, A.C. Camargo, S.M. Serrano, Sex-based individual variation of snake venom proteome among eighteen *Bothrops jararaca* siblings, *Toxicon* 47 (2006) 304–312.
- [22] J.J. Calvete, L. Sanz, A. Pérez, A. Borges, A.M. Vargas, B. Lomonte, Y. Angulo, J.M. Gutiérrez, H.M. Chalkidis, R.H. Mourão, M.F. Furtado, A.M. Moura-da-Silva, Snake population venomomics and antivenomics of *Bothrops atrox*: Paedomorphism along its transamazonian dispersal and implications of geographic venom variability on snakebite management, *J. Proteome* 74 (2011) 510–527.
- [23] R.J. Jorge, H.S. Monteiro, L. Gonçalves-Machado, M.C. Guarnieri, R.M. Ximenes, D.M. Borges-Nojosa, K.P. Luna, R.B. Zingali, C. Corrêa-Netto, J.M. Gutiérrez, L. Sanz, J.J. Calvete, D. Pla, Venomomics and antivenomics of *Bothrops erythromelas* from five geographic populations within the Caatinga ecoregion of northeastern Brazil, *J. Proteome* 114 (2015) 93–114.
- [24] D.R. Rokyta, M.J. Margres, K. Calvin, Post-transcriptional mechanisms contribute little to phenotypic variation in snake venoms, *G3 (Bethesda)* 5 (2015) 2375–2382.
- [25] D.R. Rokyta, K.P. Wray, J.J. McGivern, M.J. Margres, The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the timber rattlesnake (*Crotalus horridus*), *Toxicon* 98 (2015) 34–48.
- [26] N.R. Casewell, S.C. Wagstaff, W. Wüster, D.A. Cook, F.M. Bolton, S.I. King, D. Pla, L. Sanz, J.J. Calvete, R.A. Harrison, Medically important differences in snake venom composition are dictated by distinct postgenomic mechanisms, *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) 9205–9210.
- [27] L.F. Sousa, J.A. Portes-Junior, C.A. Nicolau, J.L. Bernardoni, M.Y. Nishiyama-Jr, D.R. Amazonas, L.A. Freitas-de-Sousa, R.H. Mourão, H.M. Chalkidis, R.H. Valente, A.M. Moura-da-Silva, Functional proteomic analyses of *Bothrops atrox* venom reveals phenotypes associated with habitat variation in the Amazon, *J. Proteome* 159 (2017) 32–46.
- [28] L. Gibbs, M. Sovic, D. Amazonas, H. Chalkidis, D. Salazar-Valenzuela, A. Moura-da-Silva, Recent lineage diversification in a venomous snake through dispersal across the Amazon River, *Biol. J. Linn. Soc.* 123 (2018) 1–15.
- [29] K.R. Elmer, A. Meyer, Adaptation in the age of ecological genomics: insights from parallelism and convergence, *Trends Ecol. Evol.* 26 (2011) 298–306.
- [30] V. Bastos, J. Salles, R. Valente, I. Leon, J. Perales, R. Dantas, R. Albano, F. Bastos, D. Bastos, Cytosolic glutathione peroxidase from liver of pacu (*Piaractus mesopotamicus*), a hypoxia-tolerant fish of the Pantanal, *Biochimie* 89 (2007) 1332–1342.
- [31] A.I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, A statistical model for identifying proteins by tandem mass spectrometry, *Anal. Chem.* 75 (2003) 4646–4658.
- [32] P. Chomczynski, N. Sacchi, Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction, *Anal. Biochem.* 162 (1987) 156–159.
- [33] S. Anders, W. Huber, Differential expression analysis for sequence count data, *Genome Biol.* 11 (2010) R106.
- [34] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, *Nat. Methods* 9 (2012) 357–359.
- [35] D.R. Rokyta, K.P. Wray, A.R. Lemmon, E.M. Lemmon, S.B. Caudle, A high-throughput venom-gland transcriptome for the eastern diamondback rattlesnake (*Crotalus adamanteus*) and evidence for pervasive positive selection across toxin classes, *Toxicon* 57 (2011) 657–671.
- [36] 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.
- [37] B. Li, C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome, *BMC Bioinformatics* 12 (2011) 323.
- [38] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data, *Bioinformatics* 26 (2010) 139–140.
- [39] M.A. Dillies, A. Rau, J. Aubert, C. Hennequet-Antier, M. Jeanmougin, N. Servant, C. Keime, G. Marot, D. Castel, J. Estelle, G. Guernec, B. Jagla, L. Jouneau, D. Laloe, C. Le Gall, B. Schaeffer, S. Le Crom, M. Guedj, F. Jaffrézic, F.S. Consortium, A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis, *Brief. Bioinform.* 14 (2013) 671–683.
- [40] K. Katoh, D.M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability, *Mol. Biol. Evol.* 30 (2013) 772–780.
- [41] R. Lanfear, P.B. Frandsen, A.M. Wright, T. Senfeld, B. Calcott, PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses, *Mol. Biol. Evol.* 34 (2017) 772–773.
- [42] A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis

- of large phylogenies, *Bioinformatics* 30 (2014) 1312–1313.
- [43] M. Anisimova, O. Gascuel, Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative, *Syst. Biol.* 55 (2006) 539–552.
- [44] S. Guindon, J.F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0, *Syst. Biol.* 59 (2010) 307–321.
- [45] C.A. Nicolau, P.C. Carvalho, I.L. Junqueira-de-Azevedo, A. Teixeira-Ferreira, M. Junqueira, J. Perales, A.G. Neves-Ferreira, R.H. Valente, An in-depth snake venom proteopeptidome characterization: benchmarking *Bothrops jararaca*, *J. Proteome* 151 (2017) 214–231.
- [46] L.A. Freitas-de-Sousa, D.R. Amazonas, L.F. Sousa, S.S. Sant'Anna, M.Y. Nishiyama Jr., S.M.T. Serrano, I.L.M. Junqueira-de-Azevedo, H.M. Chalkidis, A.M. Moura-da-Silva, R.H.V. Mourao, Comparison of venoms from wild and long-term captive *Bothrops atrox* snakes and characterization of Batroxrhagin, the predominant class PIII metalloproteinase from the venom of this species, *Biochimie* 118 (2015) 60–70.
- [47] M.J. Paine, H.P. Desmond, R.D. Theakston, J.M. Crampton, Purification, cloning, and molecular characterization of a high molecular weight hemorrhagic metalloprotease, jararhagin, from *Bothrops jararaca* venom. Insights into the disintegrin gene family, *J. Biol. Chem.* 267 (1992) 22869–22876.
- [48] G.D. Laing, A.M. Moura-da-Silva, Jararhagin and its multiple effects on hemostasis, *Toxicon* 45 (2005) 987–996.
- [49] D.H.F. Souza, M.R.C. Iemma, L.L. Ferreira, J.P. Faria, M.L.V. Oliva, R.B. Zingali, S. Niewiarowski, H.S. Selistre-de-Araujo, The disintegrin-like domain of the snake venom metalloprotease alternagin inhibits alpha 2 beta 1 integrin-mediated cell adhesion, *Arch. Biochem. Biophys.* 384 (2000) 341–350.
- [50] M.T. Assakura, A.P. Reichl, F.R. Mandelbaum, Comparison of immunological, biochemical and biophysical properties of three hemorrhagic factors isolated from the venom of *Bothrops jararaca* (jararaca), *Toxicon* 24 (1986) 943–946.
- [51] M.B. Silva, M. Schattner, C.R. Ramos, I.L. Junqueira-de-Azevedo, M.C. Guarnieri, M.A. Lazzari, C.A. Sampaio, R.G. Pozner, J.S. Ventura, P.L. Ho, A.M. Chudzinski-Tavassi, A prothrombin activator from *Bothrops erythromelas* (jararaca-da-seca) snake venom: characterization and molecular cloning, *Biochem. J.* 369 (2003) 129–139.
- [52] E.F. Sanchez, F.S. Schneider, A. Yarleque, M.H. Borges, M. Richardson, S.G. Figueiredo, K.S. Evangelista, J.A. Eble, The novel metalloproteinase atroxlysin-I from Peruvian *Bothrops atrox* (Jergón) snake venom acts both on blood vessel ECM and platelets, *Arch. Biochem. Biophys.* 496 (2010) 9–20.
- [53] L.A. Freitas-de-Sousa, M. Colombini, M. Lopes-Ferreira, S.M.T. Serrano, A.M. Moura-da-Silva, Insights into the mechanisms involved in strong hemorrhage and dermonecrosis induced by Atoxylisin-Ia, a PI-class snake venom metalloprotease, *Toxins (Basel)* 9 (2017).
- [54] K. Stocker, G.H. Barlow, The coagulant enzyme from *Bothrops atrox* venom (batroxobin), *Methods Enzymol.* 45 (1976) 214–223.
- [55] E. Oyama, T. Fukuda, H. Takahashi, Amino acid sequence of a kinin-releasing enzyme, KR-E-1, from the venom of *Agkistrodon caliginosus* (Kankoku-mamushi), *Toxicon* 52 (2008) 651–654.
- [56] B. Francis, J.M. Gutierrez, B. Lomonte, I.I. Kaiser, Myotoxin II from *Bothrops asper* (Terciopelo) venom is a lysine-49 phospholipase A2, *Arch. Biochem. Biophys.* 284 (1991) 352–359.
- [57] H.C. Castro, M. Fernandes, R.B. Zingali, Identification of bothrojaracin-like proteins in snake venoms from *Bothrops* species and *Lachesis muta*, *Toxicon* 37 (1999) 1403–1416.
- [58] M.M. Saldarriaga, R. Otero, V. Núñez, M.F. Toro, A. Díaz, J.M. Gutiérrez, Ontogenetic variability of *Bothrops atrox* and *Bothrops asper* snake venoms from Colombia, *Toxicon* 42 (2003) 405–411.
- [59] F.J. Vonk, N.R. Casewell, C.V. Henkel, A.M. Heimberg, H.J. Jansen, R.J. McCleary, H.M. Kerckamp, R.A. Vos, I. Guerreiro, J.J. Calvete, W. Wüster, A.E. Woods, J.M. Logan, R.A. Harrison, T.A. Castoe, A.P. de Koning, D.D. Pollock, M. Yandell, D. Calderon, C. Renjifo, R.B. Currier, D. Salgado, D. Pla, L. Sanz, A.S. Hyder, J.M. Ribeiro, J.W. Arntzen, G.E. van den Thillart, M. Boetzer, W. Pirovano, R.P. Dirks, H.P. Spink, D. Duboule, E. McGlinn, R.M. Kini, M.K. Richardson, 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.
- [60] T. Igarashi, S. Araki, H. Mori, S. Takeda, Crystal structures of catrocollastatin/VAP2B reveal a dynamic, modular architecture of ADAM/adamalsin/reprolysin family proteins, *FEBS Lett.* 581 (2007) 2416–2422.
- [61] Q. Liu, W. Xu, X. Cheng, G. Jin, X. Shen, H. Lou, J. Liu, Molecular cloning and sequence analysis of cDNA encoding haemorrhagic toxin acutolysin A from *Agkistrodon acutus*, *Toxicon* 37 (1999) 1539–1548.
- [62] S. Nishida, Y. Fujimura, S. Miura, Y. Ozaki, Y. Usami, M. Suzuki, K. Titani, E. Yoshida, M. Sugimoto, A. Yoshioka, Purification and characterization of batroximin, a fibrinogen-clotting serine protease from the venom of *Bothrops jararaca*, *Biochemistry* 33 (1994) 1843–1849.
- [63] A.C. Cintra, S. Marangoni, B. Oliveira, J.R. Giglio, Bothropstoxin-I: amino acid sequence and function, *J. Protein Chem.* 12 (1993) 57–64.
- [64] V.J. Lynch, Inventing an arsenal: adaptive evolution and neofunctionalization of snake venom phospholipase A2 genes, *BMC Evol. Biol.* 7 (2007) 2.
- [65] R.K. Andrews, W.J. Booth, J.J. Gorman, P.A. Castaldi, M.C. Berndt, Purification of botrocetin from *Bothrops jararaca* venom. Analysis of the botrocetin-mediated interaction between von Willebrand factor and the human platelet membrane glycoprotein Ib-IX complex, *Biochemistry* 28 (1989) 8317–8326.
- [66] H.L. Gibbs, L. Sanz, J.E. Chiucchi, T.M. Farrell, J.J. Calvete, Proteomic analysis of ontogenetic and diet-related changes in venom composition of juvenile and adult dusky pigmy rattlesnakes (*Sistrurus miliarius barbouri*), *J. Proteome* 74 (2011) 2169–2179.
- [67] J.A. Warrington, A. Nair, M. Mahadevappa, M. Tsyganskaya, Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes, *Physiol. Genomics* 2 (2000) 143–147.
- [68] S.P. Mackessy, N.M. Sixberry, W.H. Heyborne, T. Fritts, Venom of the brown treesnake, *Boiga irregularis*: ontogenetic shifts and taxa-specific toxicity, *Toxicon* 47 (2006) 537–548.
- [69] K. Sunagar, E.A. Undheim, H. Scheib, E.C. Gren, C. Cochran, C.E. Person, I. Koludarov, W. Kelln, W.K. Hayes, G.F. King, A. Antunes, B.G. Fry, Intraspecific venom variation in the medically significant southern Pacific rattlesnake (*Crotalus oreganus helleri*): biodiversity, clinical and evolutionary implications, *J. Proteome* 99 (2014) 68–83.
- [70] K. Sunagar, D. Morgenstern, A.M. Reitzel, Y. Moran, Ecological venomomics: how genomics, transcriptomics and proteomics can shed new light on the ecology and evolution of venom, *J. Proteome* 135 (2016) 62–72.
- [71] J. Arendt, D. Reznick, Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* 23 (2008) 26–32.
- [72] M. Manceau, V.S. Domingues, C.R. Linnen, E.B. Rosenblum, H.E. Hoekstra, Convergence in pigmentation at multiple levels: mutations, genes and function, *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 365 (2010) 2439–2450.