



## The tale of a resting gland: Transcriptome of a replete venom gland from the scorpion *Hottentotta judaicus*

David Morgenstern<sup>a,d,\*,1</sup>, Bettina H. Rohde<sup>b,1</sup>, Glenn F. King<sup>a</sup>, Tzachy Tal<sup>d</sup>, Daniel Sher<sup>c,d</sup>, Eliahu Zlotkin<sup>d,2</sup>

<sup>a</sup> Institute for Molecular Bioscience, The University of Queensland, St. Lucia QLD 4072, Australia

<sup>b</sup> GATC Biotech AG, Jakob-Stadler Platz 7, Konstanz 78467, Germany

<sup>c</sup> Department of Marine Biology, Leon H Charney School of Marine Sciences, University of Haifa, Mt. Carmel 31905, Israel

<sup>d</sup> Department of Cell and Animal Biology, Silberman Institute of Life Sciences, Hebrew University, Jerusalem 91904, Israel

### ARTICLE INFO

#### Article history:

Received 6 October 2010

Received in revised form 25 January 2011

Accepted 8 February 2011

Available online 15 February 2011

#### Keywords:

Scorpion

Transcriptome

Venom

Venom gland

Venom regeneration

Toxin death

*Hottentotta judaicus*

Proteases

### ABSTRACT

cDNA libraries are increasingly being used for high-throughput interrogation of animal venomes. Most previous studies have focused on discovery of new venom toxins, whereas the dynamics of toxin transcription and associated cellular processes have received much less attention. Here we provide, for the first time, an analysis of a transcriptome from the venom gland of a scorpion (*Hottentotta judaicus*) that is not actively engaged in regenerating its venom. We demonstrate a low abundance of toxin-encoding transcripts coupled with a previously unobserved proliferation of protease sequences. Additionally, we identified several low abundance, toxin-like sequences that may represent decommissioned toxins that are unlikely to be translated. These sequences are not evenly distributed across all toxin families, but rather appear more frequently in transcripts related to  $\alpha$ -toxins and  $\beta$ -toxins that are known to target voltage-gated sodium channels. The transcriptomic profile of the replete venom gland is very different to that obtained previously from scorpion venom glands actively engaged in venom regeneration, and it highlights our lack of knowledge as to how the dynamics of transcription changes as the gland progresses from venom regeneration to a "resting" state. This study therefore provides an important foundation for future studies into the dynamics of transcription in the venom glands of scorpions and other venomous animals.

© 2011 Elsevier Ltd. All rights reserved.

### 1. Introduction

Like many other venomous organisms, most scorpions depend completely on their venom apparatus—the stinger (telson) and venom gland—to survive. They use their venom to catch prey, defend themselves against predators, and even for intraspecific communication (Polis, 1990).

Most studies of scorpion venoms have focused on the identification and characterization of various toxins, often because of their medical, pharmacological and industrial importance (e.g., see (Mamelak and Jacoby, 2007; Wulff et al., 2007; Chippaux and Goyffon, 2008; Lokeshwar and Selzer, 2008; Remijsen et al., 2010; Zhao et al., 2010)). This has resulted in the identification and functional characterization of hundreds of toxins, and for many of these the relationship between their structure and function has been characterized in detail during the past six decades. The combination of proteomic approaches (e.g., (Nascimento et al., 2006; Batista et al., 2007)) and transcriptomic methods (e.g., (Schwartz and Diego-Garcia, 2007; Zhao et al., 2010)) has revealed that scorpion

\* Corresponding author.

E-mail addresses: [d.morgenstern@uq.edu.au](mailto:d.morgenstern@uq.edu.au) (D. Morgenstern), [b.rohde@gatc-biotech.com](mailto:b.rohde@gatc-biotech.com) (B.H. Rohde), [glenn.king@imb.uq.edu.au](mailto:glenn.king@imb.uq.edu.au) (G.F. King), [tzachyt@gmail.com](mailto:tzachyt@gmail.com) (T. Tal), [dsher@sci.haifa.ac.il](mailto:dsher@sci.haifa.ac.il) (D. Sher).

<sup>1</sup> These authors contributed equally to this work.

<sup>2</sup> Deceased May 18, 2008.

venom is immensely complex, often containing many hundreds of proteins (Rodriguez de la Vega et al., 2010).

During the past decade, the availability of large-scale DNA sequencing has facilitated the identification of many toxin-like peptides, greatly accelerating the pace of toxin discovery (Schwartz and Diego-Garcia, 2007; Kozminsky-Atias et al., 2008; D'Suze et al., 2009; Ma et al., 2009, 2010; Silva et al., 2009; Zhao et al., 2010). The use of cDNA libraries has also provided an opportunity for a more holistic investigation of scorpion venomes; for example, it has provided an opportunity to compare the relative abundance of toxin transcripts, thereby allowing an assessment of the relative importance of the various venom components. Moreover, the rapid increase in the number of available scorpion toxin sequences (King et al., 2008), as well as the number of species for which sequence information is available, facilitates investigation of the molecular evolution of the scorpion toxin proteome (Zhu et al., 2004a, 2010; Kozminsky-Atias et al., 2008; Tian et al., 2008; Weinberger et al., 2010). However, current methodology suffers from a major problem, as there is a tendency to interpret every novel family of protein sequences as a “venom peptide”, without at least confirming their appearance in the secreted venom. Indeed, it was recently demonstrated that some of these protein families are not seen in the proteome at detectable levels (Ma et al., 2010), advocating for a more conservative analysis of transcriptomic data.

The aforementioned cDNA libraries were all produced from scorpion venom glands that were actively engaged in the process of venom regeneration. In contrast, the aim of the research presented here was to characterize, for the first time, the transcriptomic profile of a “replete” scorpion venom gland that was not actively engaged in regenerating venom. This venom gland presented a unique transcriptomic profile that comprised a low abundance of toxin transcripts and a previously unobserved proliferation of protease sequences.

## 2. Methods

### 2.1. Animals and venom collection

*Hottentotta judaicus* specimens were collected from Southern Galilee near the village of Bir el-Maksur. The animals were housed in the laboratory under constant conditions. Scorpions were fed every 21 days with crickets and watered every 14 days. Prior to library construction, a single scorpion was starved for 14 days before the telson was removed, and it was not stimulated to produce venom during this period of time.

### 2.2. cDNA library construction

Total RNA was extracted from the venom gland and a tail fragment of the same scorpion using the RNeasy Mini Kit (QIAGEN). Double-stranded cDNAs were synthesized from 1 µg of total RNA using the SMART cDNA synthesis kit (Clontech). A subtracted cDNA library was prepared from the double-stranded cDNA fragments using the PCR Select

cDNA Subtraction Kit (Clontech) according to the manufacturer's instructions.

Double-stranded cDNA fragments were purified using the PCR Purification Kit (QIAGEN), cloned into the pCR2.1-TOPO vector and transfected into *Escherichia coli* SURE cells. 1152 randomly chosen cDNA clones were cultured overnight, and plasmid DNA was isolated using a standard alkaline lysis protocol. The cDNA inserts were sequenced from both ends with T7 and M13RP primers, respectively, using a 96-capillary 3730XL DNA analyzer (Applied Biosystems) according to the manufacturer's instructions.

### 2.3. Sequence analysis

Sequences were assembled at a stringency of >98% into contigs using DNASTAR Lasergene software (DNASTAR, Inc., USA). The library was composed of 604 clones. After removal of low quality sequences, 537 high quality clones were left, composed of 189 singlets and 94 contigs. These sequences underwent BLAST and annotation (Gene Ontology, InterPro and Kegg) using Blast2Go (Conesa et al., 2005). All other analyses were performed using CLC Main Workbench (CLC bio, Denmark). Toxin nomenclature is as suggested by King et al. (2008). Nucleotide sequences obtained in this work have been deposited in GenBank (accession numbers HQ288054–HQ288209 and HQ906937–HQ906945) and DbEST (accession numbers HS977518–HS977546). Where information is available, toxins are identified by either their UniProt (up) or GenBank (gb) accession numbers.

## 3. Results

### 3.1. Transcriptome of replete venom gland from *Hottentotta judaicus*

We analyzed a subtracted cDNA library constructed from the venom gland of a specimen of the Israeli buthid scorpion *H. judaicus* (Hj) (Arachnida: Scorpiones: Buthidae) that had not been milked or induced to produce venom for two weeks; we refer to this as a “replete” venom gland. All previously constructed scorpion venom gland cDNA libraries were obtained from recently milked specimens where the gland was actively engaged in venom regeneration; we refer to this as a “regenerating” venom gland. Previous analyses of regenerating venom glands from buthid scorpions (i.e., members of the Buthidae family) have yielded transcriptomes dominated by transcripts encoding peptide toxins (defined as toxins smaller than 10 kDa) (see Table 1) (Kozminsky-Atias et al., 2008; D'Suze et al., 2009; Zhao et al., 2010). In contrast, the transcriptome of the replete venom gland from *H. judaicus* contained a much lower abundance of toxin transcripts (24% of total transcripts) (Fig. 1a, Table 1) than found in the transcriptome of regenerating venom glands from other buthids (44–78% of total transcripts). Indeed, the level of toxin transcripts was even lower than found in regenerating venom glands of non-buthid scorpions (36–44% of total transcripts) (Schwartz and Diego-Garcia, 2007; Ma et al., 2009, 2010; Silva et al., 2009) which are less reliant than buthids on their venom, as they use their pedipalps more

**Table 1**

Relative weight of protein families in various venom-derived cDNA libraries.

|                         | Hj  | Boi <sup>a</sup> | Td  | Lm  | Sj <sup>a</sup> | Oc  | Hg <sup>a</sup> |
|-------------------------|-----|------------------|-----|-----|-----------------|-----|-----------------|
| Toxins & venom peptides | 24% | 78%              | 50% | 55% | 42%             | 36% | 44%             |
| Cellular proteins       | 39% | n/a <sup>b</sup> | 13% | 21% | 34%             | 39% | 26%             |
| Unknown & hypothetical  | 37% | n/a <sup>b</sup> | 37% | 29% | 24%             | 25% | 30%             |

The abbreviations used are: Hj, *Hottentotta judaicus* (this study); Boi, *Buthus occitanus israelis* (Kozminsky-Atias et al., 2008); Td, *Tityus discrepans* (D'Suze et al., 2009); Sj, *Scorpiops jendeki* (Ma et al., 2009); Lm, *Lychas mucronatus* (Zhao et al., 2010); Oc, *Opisthacanthus cayaporum* (Silva et al., 2009); Hg, *Hadrurus gertschi* (Schwartz and Diego-Garcia, 2007).

<sup>a</sup> Does not include non-ORF sequences.

<sup>b</sup> Data not available.

extensively for prey capture (Polis, 1990). Moreover, the low abundance observed is likely to be a significant overestimate of the actual proportion of toxin transcripts. Unlike previously constructed scorpion venom gland cDNA libraries, the library used here was constructed using an RNA subtraction protocol in order to enrich for transcripts specific to the venom gland and to reduce the proportion of “housekeeping” transcripts commonly found in other tissues (Diatchenko et al., 1996).

The other categories of transcripts found in the library can be divided into functionally annotated housekeeping-related components (i.e., those with an associated gene ontology (GO) term) and those whose function is unknown (no GO terms). Analysis of the functionally annotated

components reveals two main transcript populations: (i) transcripts encoding proteases. These transcripts comprise a few high-copy transcripts that represent ~16% of the library (Fig. 1a,b). Notably, protease transcripts have never been described before in such abundance in any scorpion venom gland cDNA library; (ii) other housekeeping transcripts, representing ~25% of the library. These transcripts (Supp. Fig. 1) have relatively low abundance and are each represented by very few clones. They may be the result of transcriptional variability due to the subtraction process rather than being uniquely transcribed in the venom gland.

The last group of transcripts are unique in that have never been described before or are not affiliated to any function. These transcripts make up 37% of the library, an amount similar to the level of unique transcripts found in most other scorpion venom gland cDNA libraries. SignalP and TMHMM analyses revealed that most proteins encoded by these transcripts (59%) are predicted to be either secreted or have a transmembrane domain.

In the following sections we will focus our attention on toxin-encoding transcripts, as they have been the focus of previous studies of scorpion venom gland cDNA libraries.

### 3.2. Toxin-like sequences

#### 3.2.1. Sodium channel toxins

One of the major differences between buthid and non-buthid scorpion venoms is the abundance and dominance

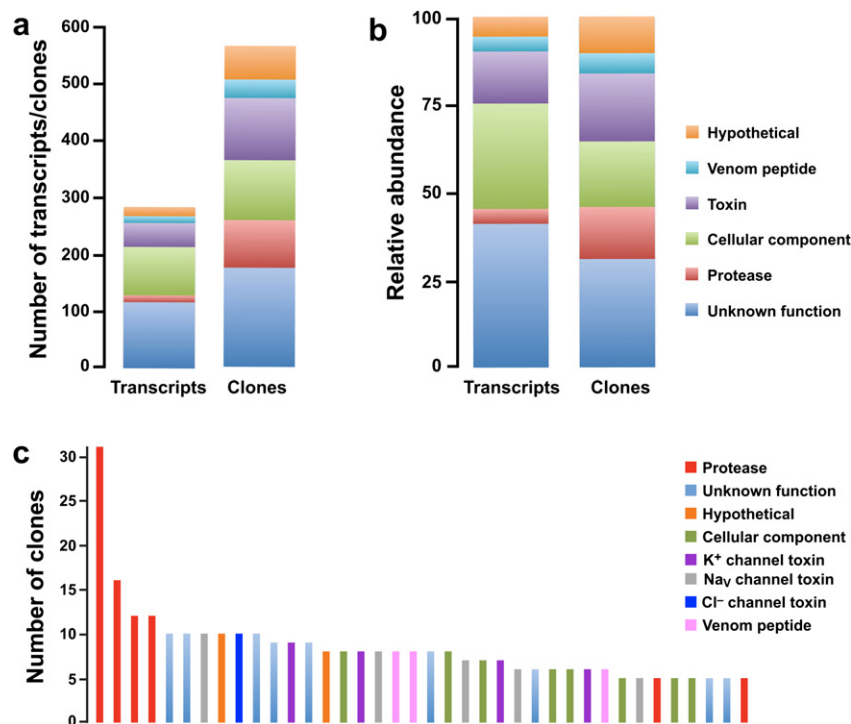


**Fig. 1.** Distribution of venom gland transcripts. Schematic showing the absolute quantity (a) and the relative proportion (b) of each category of transcript. Both the total number of clones and the number of unique transcripts is provided. “Unknown” refers to transcripts that yielded no obvious sequence homologs through BLAST analysis. “Hypothetical” indicates sequences that show homology to previously described transcripts encoding proteins whose function is not known. “Venom peptides” are venom-related sequences that have no known toxic activity. (c) Histogram showing the abundance of various transcripts; the number of clones obtained for each transcript is indicated. The functional affiliation of each component is indicated by the color codes.

of so-called “long-chain” sodium channel toxins (NaScTxS) in buthid venoms compared with the almost complete lack of them in non-buthid scorpions. This has been emphasized quantitatively in transcriptomic analysis of buthid venoms (Kozminsky-Atias et al., 2008; D'Suze et al., 2009).

Surprisingly, the abundance of NaScTx transcripts in the cDNA library from the replete venom gland from *H. judaicus* is relatively low (Fig. 2). The majority of toxin transcripts in replete venom gland transcriptome are potassium channel toxins (KScTxS), which is uncharacteristic of buthid venoms. Moreover, the relative proportion of the two major classes of NaScTxS, namely the  $\alpha$  scorpion toxins ( $\alpha$ NaScTxS) and  $\beta$  scorpion toxins ( $\beta$ NaScTxS), is unusual. These two toxin classes occur in a ratio of 60:40 ( $\alpha$ NaScTx: $\beta$ NaScTx) in a library constructed from the regenerating venom gland of *Buthus occitanus israelis* (Boi), which occurs in an adjacent geographical area (Kozminsky-Atias et al., 2008). In contrast,  $\alpha$ NaScTx transcripts are under-represented in the transcriptome of the replete venom gland from *H. judaicus*; only ~20% of all toxin transcripts encode  $\alpha$ NaScTxS and the  $\alpha$ NaScTx: $\beta$ NaScTx ratio is 24:76.

Other toxin classes found in the transcriptome of the replete venom gland include modulators of ryanodine receptors (RyRs) and chloride channels, and other putative venom peptides including Cysteine Rich Secretory Proteins (CRISPs) and antimicrobial peptides (Fig. 2). The category of “putative venom peptides” refers to disulfide-rich and non-disulfide bridged peptides that are commonly associated with venom, but are not known to be toxic; a full list of these peptides can be found in Supplementary Table 1.

### 3.2.2. Abundance of previously identified *H. judaicus* toxins

Eight different toxins have been previously isolated and characterized from the venom of *H. judaicus*; all of these toxins modulate either voltage-gated sodium (Nav) channels or RyRs. Of these eight known toxins, we detected transcripts for five, namely BjIT2 (up:P24336), cardiotoxic peptide (up:Q2TSD9), BjTx2 (up:P83406), BjIT1 (up:Q56TT9) and BjV (Tal, T. and Ben-Shaanan, T., unpublished results). Two of these (BjIT2, BjTx2) were more abundant, each comprising ~3.5–5.5% of total toxin transcripts. The toxins for which we could not identify corresponding transcripts were BjIT1 (up:P56637) and BjTx1 (up:P83405), a homolog of BjTx2. Intriguingly, BjIT1 is considered to be a major component of the venom, comprising ~1% of the dry weight of the crude milked venom (Froy et al., 1999), yet no transcripts encoding this toxin were identified in the *H. judaicus* cDNA library. In contrast, two other  $\alpha$ NaScTxS, which have not previously been described, appear in higher abundance—six and four clones each—representing 4.2% and 2.8% of the total toxin transcripts, respectively. Both are homologs of  $\alpha$  toxins that bind to site 3 on Nav channels. In total, six novel NaScTx transcripts were detected (Supplementary Table 1), most of which belong to the  $\beta$  depressant toxin family (Possani et al., 1999).

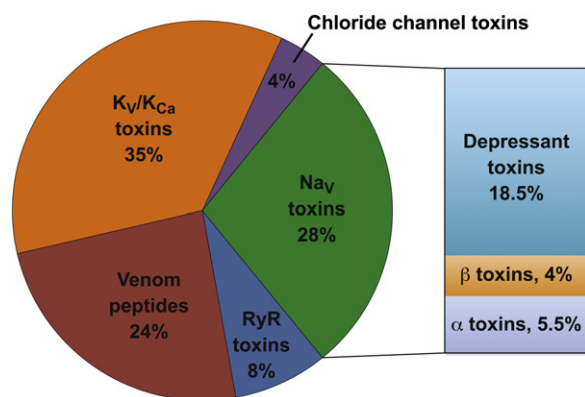


**Fig. 2.** Relative proportion of different categories of toxin transcripts in the *H. judaicus* venom gland transcriptome. The cDNA library contained 124 toxin-encoding sequences that have been divided into various functional classes. The Nav channel toxins (NaScTx) have been further divided into three major subcategories, namely  $\alpha$  toxins,  $\beta$  depressant toxins, and putative  $\beta$  toxins.

(up:Q56TT9) and BjV (Tal, T. and Ben-Shaanan, T., unpublished results). Two of these (BjIT2, BjTx2) were more abundant, each comprising ~3.5–5.5% of total toxin transcripts. The toxins for which we could not identify corresponding transcripts were BjIT1 (up:P56637) and BjTx1 (up:P83405), a homolog of BjTx2. Intriguingly, BjIT1 is considered to be a major component of the venom, comprising ~1% of the dry weight of the crude milked venom (Froy et al., 1999), yet no transcripts encoding this toxin were identified in the *H. judaicus* cDNA library. In contrast, two other  $\alpha$ NaScTxS, which have not previously been described, appear in higher abundance—six and four clones each—representing 4.2% and 2.8% of the total toxin transcripts, respectively. Both are homologs of  $\alpha$  toxins that bind to site 3 on Nav channels. In total, six novel NaScTx transcripts were detected (Supplementary Table 1), most of which belong to the  $\beta$  depressant toxin family (Possani et al., 1999).

### 3.2.3. Potassium channel toxins

The majority of toxin transcripts in the *H. judaicus* cDNA library (42 clones representing 18 distinct toxin sequences) encoded KScTxS, belonging to eight distinct families (Fig. 3). Most of these transcripts (26 clones, five distinct toxins) represented three families of toxins that modulate the activity of small-conductance, calcium-activated potassium (K<sub>Ca</sub>) channels ( $\alpha$ KTx5,  $\alpha$ KTx8, and  $\alpha$ KTx14). While data as to their role in the venom's toxicity is scarce (Romi-Lebrun et al., 1997; Basis et al., 2005), their abundance in the venom warrants further investigation. One of the most potent known insecticidal toxins isolated from spider venom is a potent blocker of insect K<sub>Ca</sub> channels (Gunning et al., 2008), which suggests that the putative K<sub>Ca</sub> toxins in *H. judaicus* venom may play an important role in insect predation.

### 3.2.4. Ryanodine receptor toxins

The action of scorpion venoms on RyRs has been attributed almost exclusively to toxins with an inhibitor cystine knot (ICK) motif, a protein scaffold that is relatively rare in

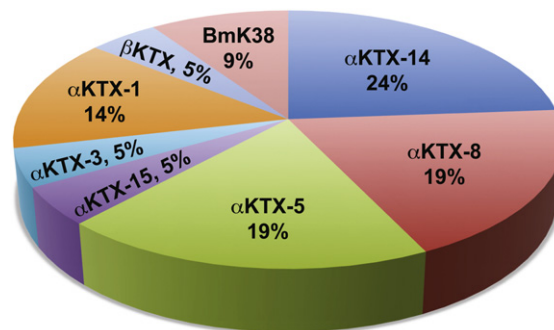


**Fig. 3.** Relative proportion of transcripts encoding different classes of potassium channel toxins. The vast majority (86%) of the transcripts encoding potassium channel toxins (KTx) are  $\alpha$  toxins, with only 5% belonging to the long chain  $\beta$  toxin class. The three dominant KTx families are  $\alpha$ KTx-5,  $\alpha$ KTx-8, and  $\alpha$ KTx-14. These toxin classes are associated with activity on small-conductance, calcium-activated potassium (K<sub>Ca</sub>) channels.

### 3.3. “Decommissioned” toxins

### 3.3.1. *U<sub>3m</sub>*-Buthitoxin-Hi1a—skipped stop codon

U<sub>3m</sub>-Buthitoxin-Hj1a (U<sub>3m</sub>-BUTX-Hj1a; gb:HQ288091) based on the rational nomenclature proposed by King et al. (2008), appears to have an additional C-terminal “tail” of 14 amino acid residues. The calculated molecular mass of 8793/9087 Da (depending on the start of the mature toxin sequence) makes U<sub>3m</sub>-BUTX-Hj1a the heaviest NaScTx described to date. The additional amino acids may be caused by loss of a conserved stop codon found in other, similar toxins, probably though a G→A substitution in nucleotide position 209, which alters the TGA stop codon to a Trp-encoding TGG codon. U<sub>3m</sub>-BUTX-Hj1a is quite distinct from other scorpion toxins, with the highest level of identity being 51% and 52% to AaH6 (up:P56743) (Hassani et al., 1999) from *Androctonus australis* hector and TSVI/TSNTxP (Marangoni et al., 1990; Guatimosim et al., 1999) from *T. serrulatus*. While the differences between this toxin and its closest relatives do not predict *a priori* loss of function, it is intriguing that the peptides most similar to U<sub>3m</sub>-BUTX-Hj1a either have low or no demonstrated toxicity.

U<sub>1m</sub>-BUTX-Hj1c (gb: HQ288099) is an example of a heavily mutated toxin sequence (Fig. 5; see assembly in Supplementary Fig. 2). CSαβ toxins from scorpion venom have a highly conserved cystine scaffold comprising three strictly conserved disulfide bridges, with some toxins having an additional, less conserved disulfide-bridge (most NaScTx, γKTx, and chloride channel toxins). This motif can be identified from the primary structure by the presence of two distinct motifs, CX<sub>3</sub>C and CXC, where X denotes any amino acid. The CX<sub>3</sub>C motif is not present in the amino acid sequence of U<sub>1m</sub>-BUTX-Hj1c whereas the CXC motif is retained. However, a protein BLAST reveals high homology with a three-disulfide βNaScTx from *Androctonus crassicauda*, called Acra III-2 (up:POC299), and its homologs (Caliskan et al., 2006). However, the sequence similarity is limited to the C-terminus of the protein, from position 36 onwards, with 66% identity and 72% similarity (Fig. 5). The positions 3' to this section have a considerably lower similarity to the Acra toxins, and in fact, the protein structure is lost completely; based on the alignment, three

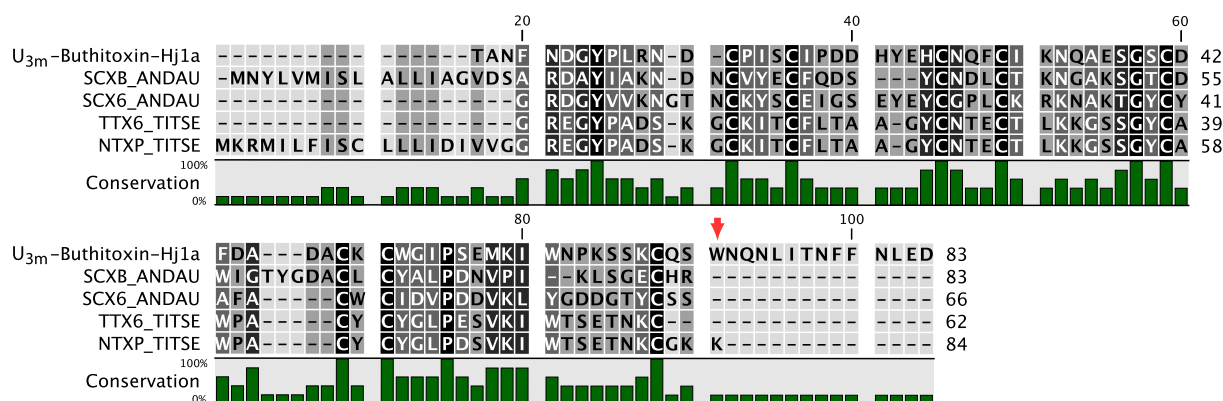


**Fig. 4.** Loss of stop codon in U<sub>3m</sub>-BUTX-Hj1a results in a long C-terminal tail. U<sub>3m</sub>-BUTX-Hj1a aligned with a group of weakly active or non-toxic peptides from *A. australis* (North Africa) and *T. serrulatus* (South America). Note that U<sub>3m</sub>-Buthitoxin-Hj1a contains 14 additional C-terminal residues (starting at the point marked by the arrow) compared to the other sequences due to the loss of a stop signal. In contrast, the core region of the mature toxin is highly homologous with the other toxin sequences.

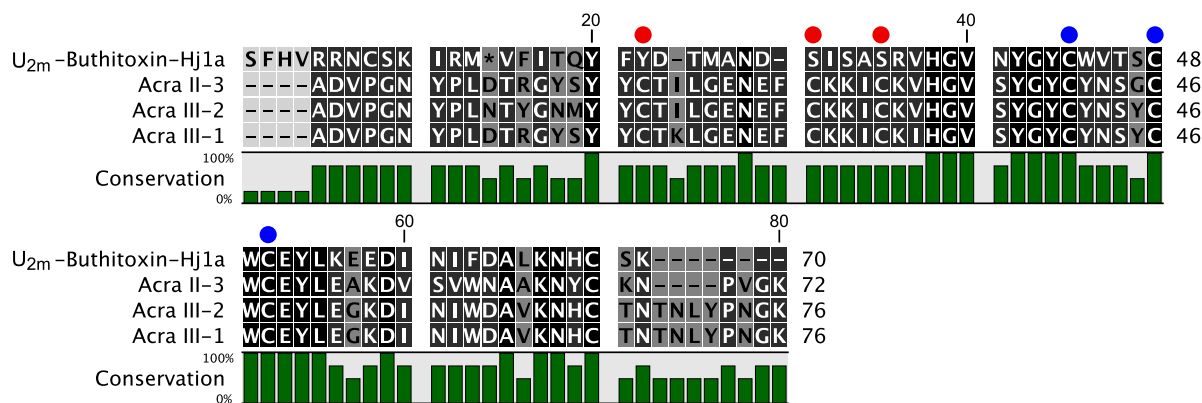


**Fig. 5.** Loss of structurally important cysteine residues in the N-terminal region of U<sub>1m</sub>-BUTH-Hj1a. Alignment of the primary structure of U<sub>1m</sub>-BUTH-Hj1a with Acra toxins from *Androctonus* sp. The N-terminal region of the *H. judaicus* sequence (i.e., residues 1–36) is highly mutated relative to the Acra toxins, leading to loss of structural cysteines at the sites indicated by the dark circles. In contrast, the C-terminal region shows a high level of sequence identity with the Acra toxins and the structurally important cysteine residues, indicated by the pale circles, are conserved in this region.

cysteines have been lost, most probably due to point substitutions: C22→Y22 (TGT→TAT), C31→S31 (TGC→TCC), and C35→S35 (TGC→TCC). Additionally, the sequence has completely lost its reading frame, with the signal peptide unidentifiable due to multiple stop codons. Essentially, this sequence has undergone a series of deleterious mutations, accompanied by probable structural degradation occurring over a very long time.

### 3.3.3. $\kappa_m$ -BUTH-Hj1c—loss of signal peptide coding sequence

$\kappa_m$ -BUTH-Hj1c (gb: HQ288097; Fig. 6; see assembly in Supplementary Fig. 3) is an ortholog of  $\alpha$ KTx 8.2 (up:Q9U8D2), a toxin from the scorpion *Mesobuthus martinsei* (Romi-Lebrun et al., 1997; Wu et al., 1999) that targets K<sub>Ca</sub> channels. Transcripts encoding two isoforms of this toxin were observed in this cDNA library, one with eight clones ( $\kappa$ -BUTH-Hj1a) and one with two clones ( $\kappa_m$ -BUTH-Hj1c) (Fig. 3C).  $\kappa$ -BUTH-Hj1a (gb:HQ288061) and  $\kappa$ -BUTH-Hj1c are clearly paralogs but the rarer transcript encoding  $\kappa$ -BUTH-Hj1c lacks a 20-residue segment corresponding to most of the signal peptide. A comparison between the known genomic organization of  $\alpha$ KTx8.2 (Wu et al., 1999) and the transcript encoding  $\kappa$ -BUTH-Hj1c suggests that the missing segment corresponds to exon 1. Moreover, the sequence of the mature  $\kappa$ -BUTH-Hj1a toxin is identical to that of  $\kappa$ -BUTH-Hj1c, with the only difference between them being the lack of this exon. Thus, it appears that  $\kappa$ -BUTH-Hj1c has lost most of the sequence encompassed in exon 1, including the whole coding sequence and a part of the 5' untranslated region (UTR), making it unlikely to be processed during translation.

## 4. Discussion

### 4.1. The unique transcriptional profile of a replete venom gland

The transcriptome of the replete venom gland from *H. judaicus* is clearly different from that described previously for regenerating venom glands from other buthid scorpions: the relative abundance of toxin transcripts is low,

and most of the toxin transcripts encode K<sub>V</sub>/K<sub>Ca</sub> toxins. Additionally, most of the transcripts encoding K<sub>V</sub>/K<sub>Ca</sub> toxins belong to the  $\alpha$ KTx5,  $\alpha$ KTx8 and  $\alpha$ KTx14 families; the role of these toxins in the venom remains to be elucidated as they act on ion channels that are not considered critical for the insecticidal activity of the venom.

Taken together, these observations tend to suggest that *H. judaicus* does not possess highly potent venom but rather is similar to non-buthids in relying more on physical subjugation of its prey using its pedipalps. In fact, contrary to this supposition, although *H. judaicus* venom is not fatal towards vertebrates, it is highly potent against arthropods that are its natural prey (Zlotkin et al., 1978, 1982). Therefore, other explanations must be sought for the unusual pattern and distribution of transcripts in the cDNA library of the replete venom gland.

One possible explanation is that the venom gland transcriptome does not reflect the actual composition of the venom. In this scenario, the most active “core peptides” are more abundant in the venom itself than in the cDNA library. A difference in protein profile between the scorpion venom transcriptome and proteome has already been discerned (Ma et al., 2010), and it is plausible that there may be differences in both the range and quantity of toxins.

An alternative hypothesis is that the transcriptome does indeed reflect the venom content of *H. judaicus* or, more likely, a specific *H. judaicus* population. Population variation in the venom content in venomous animals has been identified previously, even between populations living within a short distance from one another (Abdel-Rahman et al., 2009). This hypothesis can be tested through additional sampling of *H. judaicus* venom gland transcriptomes, including comparison of venom gland transcriptomes from different sub-populations. If true, one wonders what environmental pressures (most likely prey or predators) could drive the unusual patterns of transcript abundances shown here, especially the bias for K<sub>V</sub>/K<sub>Ca</sub> channel versus Na<sub>V</sub> channel toxins.

A third, and perhaps most likely, hypothesis stems from the fact that the *H. judaicus* specimen from which the venom gland was collected was starved for ~3 weeks and

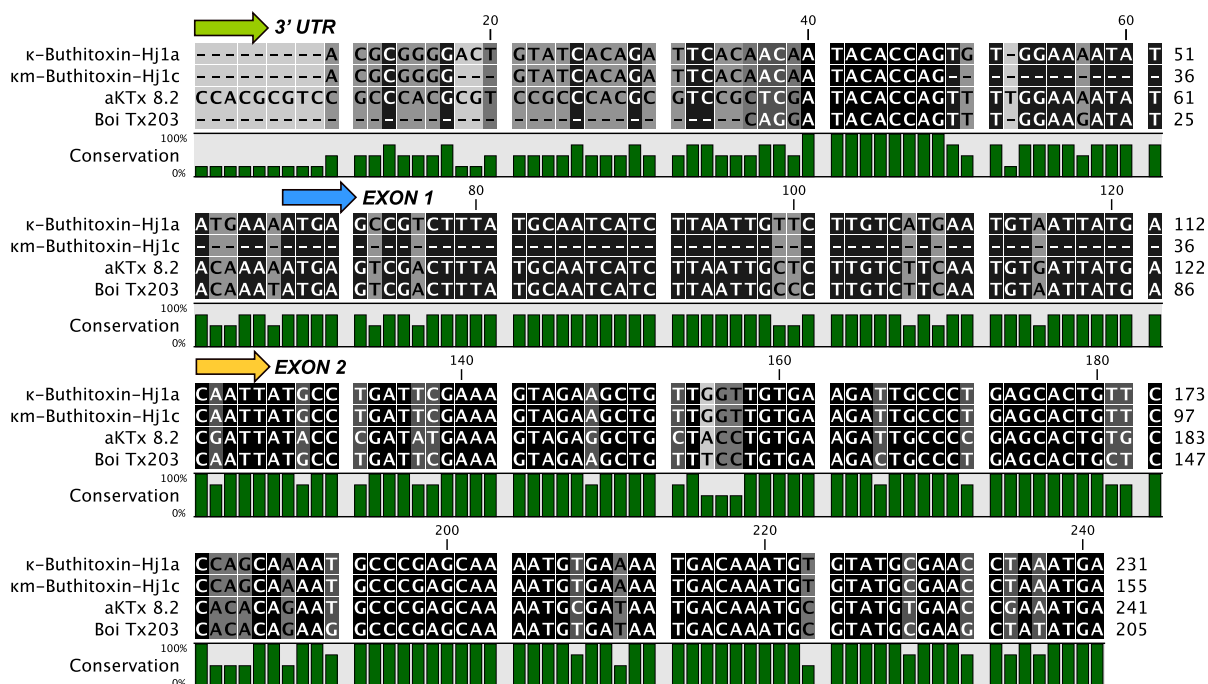


**Fig. 6.** Loss of 3' coding sequence and UTR in  $\kappa$ -BUTX-Hj1b. Nucleic acid alignment of transcripts encoding  $\kappa$ -BUTX-Hj1a and  $\kappa$ -BUTX-Hj1c with transcripts encoding  $\alpha$ KTx8.2 from *Mesobuthus martensii* and Tx203 from *Buthus occitanus israelis*. Although the mature toxin region of  $\kappa$ -BUTX-Hj1c is identical to that of  $\kappa$ -BUTX-Hj1a, the  $\kappa$ -BUTX-Hj1c transcript lacks most of exon 1, which encodes most of the signal peptide, thus preventing translation. The arrows indicate the start points of the 3' UTR and exons 1 and 2.

wasn't milked prior to the extraction of the mRNA. In contrast, most venom gland cDNA libraries have been produced from scorpions 2–5 days after milking in order to enrich for venom-related transcripts. Thus, the *H. judaicus* transcriptome presented here may be representative of a “resting” or replete venom gland, compared to the “regenerating” venom gland represented by other cDNA libraries. This hypothesis is consistent with the relatively low level of toxin transcripts compared to those involved in cellular metabolism; the observed level of toxin transcripts in the *H. judaicus* library was lower than previously obtained from regenerating venom glands from both buthid and non-buthid scorpions (Table 1), even after enrichment. In this regard, it is possible that the difference in transcript abundance between the  $K_v/K_{Ca}$  and  $N_{av}$  channel toxins may not reflect the ratio of these two classes of peptides in the venom, but rather the transcript levels necessary to maintain the venom gland in a “ready to act” state. It is possible that the  $K_v/K_{Ca}$  channel toxins are less stable than the  $N_{av}$  channel toxins and thus incur a higher steady-state turnover, which in turn necessitates a higher level of mRNA transcripts in order to maintain the desired concentration of toxins in the venom.

Resting venom glands have been the focus of structural research on this tissue and it has been shown that even resting glands maintain a certain level of protein transport into secretory vacuoles via secretory vesicles (Keegan and Lockwood, 1971). Moreover, research on regeneration of spider venom has revealed that regeneration of toxicity and protein is a time-dependent process that is most probably

reliant on gradual transcription of toxin components (Boevé et al., 1995). Quantitative analysis of scorpion venom gland transcriptomes before and during venom regeneration should provide significant insight into the mechanism and dynamics of toxin transcription.

The uniquely high abundance of protease transcripts in the *H. judaicus* cDNA library may provide additional hints as to underlying causes of the unusual transcriptomic profile. While there is no evidence for high abundance of serine proteases and metalloproteases in scorpion venom, these proteins play important functions in the every day cellular processes such as cell signaling, cell adhesion, invasion and motility, and wound healing (Jones and Riley, 2005). The abundance of these transcripts is a curious observation, but further work is needed to determine whether their relatively high abundance is a general phenomenon for resting venom glands, or whether they are simply a reflection of the physiology or health of the specific scorpions used for this study.

#### 4.2. Decommissioned toxins

Research on scorpion venoms has revealed an extensive occurrence of toxin paralogs (Valdez-Cruz et al., 2004; Strugatsky et al., 2005; Zhijian et al., 2006). This phenomenon is explained by a model whereby progenitor toxin genes were duplicated multiple times followed by strong positive selection (Zhu et al., 2004a). While this might explain the path of birth of new toxins, there is very little evidence as to the path of toxin “death”. Li and co-workers

(Li et al., 2005) studied toxin “death” in the degenerating venom secreting system of a snake that gradually shifted its diet from fish to fish eggs. They suggested that the observed truncation of a three-finger toxin (a major component in the venom of sea snakes (Fry et al., 2003)) which led to elimination of this toxin activity from the venom resulted from a loss of selection pressure caused by the dietary switch. However, loss of selection pressure can occur even in a functional venom gland. Differences in prey and geographical location (Daltry et al., 1996; Barlow et al., 2009; Casewell et al., 2009) are likely to lead to different selection pressures (or lack thereof) on different toxins in the venom. Some toxins may lose selection pressure as a result of such changes, allowing the accumulation of deleterious mutations that would have otherwise been purified.

The same thing may occur during gene duplication (Sollod et al., 2005; Zhijian et al., 2006). When a duplicated toxin gene suffers deleterious mutations, it can be either purified if it affects the fitness of the scorpion, or if it has lost selection pressure it can continue to accumulate mutations over time. Some of these toxin transcripts may maintain an ORF yet undergo a series of mutations leading to loss of function due to changes in either the composition of the sequence or through additions and deletions of the sequence; these toxins are referred to as “disabled toxins”.

Other sequences, known as “expressed pseudogenes”, may be transcribed into mRNA that cannot be translated due to loss of an ORF or other translation and processing signals. Expressed pseudogenes constitute a higher degree of toxin “decommissioning” than disabled toxins since disabled toxins may still appear in the venom even though they might not contribute to its activity. In this study we presented three examples of transcripts encoding decommissioned toxins. Two of these appear to be pseudogenes as they contain no ORF due to the presence of a premature stop codon or the loss of start signal. This phenomenon is not unique to our library; several similar sequences have been presented in previous studies (Fazal et al., 1989; Martin-Eauclaire et al., 2005), although their anomalous features were not further elaborated. These pseudogenes and disabled toxins may provide a “snapshot” of the process of toxin death, and this may explain the presence of such transcripts in venom gland cDNA libraries. Moreover, identification of toxin clades that are rich or poor with such sequences may provide “smoking gun” evidence of the pathway of toxin evolution (Fleishman et al., 2003).

## 5. Conclusions

We have presented here, for the first time, the transcriptome of a replete scorpion venom gland. The transcriptional profile of the replete venom gland is uniquely different from previously described transcriptomes from regenerating scorpion venom glands. Surprisingly, the replete venom gland maintains a low level of toxin transcription that is presumably designed to facilitate a low level of toxin production, transport, and storage. Our results indicate that analysis of scorpion venom gland transcriptomes before and during the process of venom regeneration should provide important insights into the

mechanism and dynamics of toxin transcription. Moreover, the identification of pseudogenes and deactivated toxins in these transcriptomes may provide a snapshot of the process of toxin decommissioning and death during scorpion venom evolution.

## Acknowledgements

We thank Mr. Johannes Regenbogen for assistance in production of the cDNA library. DM acknowledges the support of a University of Queensland Research Scholarship (UQRS) and an Endeavor International Postgraduate Research Scholarship (IPRS). The work presented here was funded by the German “BMBF (Federal Ministry of Education and Research), support code 0312616. EZ was supported by the “Committee of German–Israeli Cooperation in Biotechnology. This work is dedicated to the memory of the late Prof. Eliahu Zlotkin.

## Conflict of interest

None declared.

## Ethical statement

The authors declare that this manuscript complies with the Elsevier Ethical Guidelines for Journal Publication and acknowledge the Duties of Authors described therein.

## Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [10.1016/j.toxicon.2011.02.001](http://dx.doi.org/10.1016/j.toxicon.2011.02.001).

## References

- Abdel-Rahman, M.A., Omran, M.A.A., Abdel-Nabi, I.M., Ueda, H., McVean, A., 2009. Intraspecific variation in the Egyptian scorpion *Scorpio maurus palmatus* venom collected from different biotopes. *Toxicon* 53, 349–359.
- Barlow, A., Pook, C.E., Harrison, R.A., Wuster, W., 2009. Co-evolution of diet and prey-specific venom activity supports the role of selection in snake venom evolution. *Proc. R. Soc. B* 276, 2443–2449.
- Basis, F., Kapeliovich, M., Bentur, Y., 2005. Low atrial rhythm as a complication of *buthotus judaicus* (“black” scorpion) evenomation. *Isr. J. Emerg. Med.* 5, 60–62.
- Batista, C.V.F., Román-González, S.A., Salas-Castillo, S.P., Zamudio, F.Z., Gómez-Lagunas, F., Possani, L.D., 2007. Proteomic analysis of the venom from the scorpion *Tityus stigmurus*: Biochemical and physiological comparison with other *Tityus* species. *Comp. Biochem. Physiol. C Toxicol. Pharmacol.* 146, 147–157.
- Boevé, J.-L., Kuhn-Nentwig, L., Keller, S., Nentwig, W., 1995. Quantity and quality of venom released by a spider (*Cupiennius salei*, Ctenidae). *Toxicon* 33, 1347–1357.
- Caliskan, F., Garcia, B.I., Coronas, F.I., Batista, C.V., Zamudio, F.Z., Possani, L.D., 2006. Characterization of venom components from the scorpion *Androctonus crassicauda* of Turkey: peptides and genes. *Toxicon* 48, 12–22.
- Casewell, N., Harrison, R., Wuster, W., Wagstaff, S., 2009. Comparative venom gland transcriptome surveys of the saw-scaled vipers (Viperidae: *Echis*) reveal substantial intra-family gene diversity and novel venom transcripts. *BMC Genomics* 10, 564.
- Chippaux, J.P., Goyffon, M., 2008. Epidemiology of scorpionism: a global appraisal. *Acta Tropica* 107, 71–79.
- Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., Robles, M., 2005. Blast2GO: a universal tool for annotation, visualization and

- analysis in functional genomics research. *Bioinformatics* 21, 3674–3676.
- Daltry, J.C., Wuster, W., Thorpe, R.S., 1996. Diet and snake venom evolution. *Nature* 379, 537–540.
- Diatchenko, L., Lau, Y.F., Campbell, A.P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E.D., Siebert, P.D., 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl. Acad. Sci. U.S.A.* 93, 6025–6030.
- D'Suze, G., Schwartz, E.F., García-Gómez, B.I., Sevcik, C., Possani, L.D., 2009. Molecular cloning and nucleotide sequence analysis of genes from a cDNA library of the scorpion *Tityus discrepans*. *Biochimie* 91, 1010–1019.
- Fazal, A., Beg, O.U., Shafqat, J., Zaidi, Z.H., Jörnvall, H., 1989. Characterization of two different peptides from the venom of the scorpion *Buthus indicus*. *FEBS Lett.* 257, 260–262.
- Fleishman, S.J., Dagan, T., Graur, D., 2003. pANT: a method for the pairwise assessment of nonfunctionalization times of processed pseudogenes. *Mol. Biol. Evol.* 20, 1876–1880.
- Froy, O., Zilberberg, N., Gordon, D., Turkov, M., Gilles, N., Stankiewicz, M., Pelhate, M., Loret, E., Oren, D.A., Shaanan, B., Gurevitz, M., 1999. The putative bioactive surface of insect-selective scorpion excitatory neurotoxins. *J. Biol. Chem.* 274, 5769–5776.
- Fry, B.G., Lumsden, N.G., Wüster, W., Wickramaratna, J.C., Hodgson, W.C., Kini, R.M., 2003. Isolation of a neurotoxin ( $\alpha$ -colubritoxin) from a nonvenomous colubrid: evidence for early origin of venom in snakes. *J. Mol. Evol.* 57, 446–452.
- Guatimosim, S.C., Prado, V.F., Diniz, C.R., Chávez-Olortegui, C., Kalapothakis, E., 1999. Molecular cloning and genomic analysis of TsNTxp: an immunogenic protein from *Tityus serrulatus* scorpion venom. *Toxicon* 37, 507–517.
- Gunning, S.J., Maggio, F., Windley, M.J., Valenzuela, S.M., King, G.F., Nicholson, G.M., 2008. The Janus-faced atracotoxins are specific blockers of invertebrate  $K_{Ca}$  channels. *FEBS J.* 275, 4045–4059.
- Hassani, O., Loew, D., Van Dorsselaer, A., Papandréou, M.J., Sorokine, O., Rochat, H., Sampieri, F., Mansuelle, P., 1999. Aah VI, a novel, N-glycosylated anti-insect toxin from *Androctonus australis hector* scorpion venom: isolation, characterisation, and glycan structure determination. *FEBS Lett.* 443, 175–180.
- Jones, G., Riley, G., 2005. ADAMTS proteinases: a multi-domain, multi-functional family with roles in extracellular matrix turnover and arthritis. *Arthritis Res. Ther.* 7, 160–169.
- Keegan, H.L., Lockwood, W.R., 1971. Secretory epithelium in venom glands of two species of scorpion of the genus *Centruroides* Marx. *Am. J. Trop. Med. Hyg.* 20, 770–785.
- King, G.F., Tedford, H.W., Maggio, F., 2002. Structure and function of insecticidal neurotoxins from Australian funnel-web spiders. *Toxin Rev.* 21, 361–389.
- King, G.F., Gentz, M.C., Escoubas, P., Nicholson, G.M., 2008. A rational nomenclature for naming peptide toxins from spiders and other venomous animals. *Toxicon* 52, 264–276.
- Kozminsky-Atias, A., Bar-Shalom, A., Mishmar, D., Zilberberg, N., 2008. Assembling an arsenal, the scorpion way. *BMC Evol. Biol.* 8, 333.
- Li, M., Fry, B.G., Kini, R.M., 2005. Eggs-only diet: its implications for the toxin profile changes and ecology of the marbled sea snake (*Aipysurus eydouxii*). *J. Mol. Evol.* 60, 81–89.
- Lokeswar, V.B., Selzer, M.G., 2008. Hyaluronidase: both a tumor promoter and suppressor. *Semin. Canc. Biol.* 18, 281–287.
- Ma, Y., Zhao, R., He, Y., Li, S., Liu, J., Wu, Y., Cao, Z., Li, W., 2009. Transcriptome analysis of the venom gland of the scorpion *Scorpiops jendeki*: implication for the evolution of the scorpion venom arsenal. *BMC Genomics* 10, 290.
- Ma, Y., Zhao, Y., Zhao, R., Zhang, W., He, Y., Wu, Y., Cao, Z., Guo, L., Li, W., 2010. Molecular diversity of toxic components from the scorpion *Heterometrus petersii* venom revealed by proteomic and transcriptome analysis. *Proteomics* 10, 2471–2485.
- Mamelak, A.N., Jacoby, D.B., 2007. Targeted delivery of antitumoral therapy to glioma and other malignancies with synthetic chlorotoxin (TM-601). *Expert Opin. Drug Deliv.* 4, 175–186.
- Marangoni, S., Ghiso, J., Sampaio, S.V., Arantes, E.C., Giglio, J.R., Oliveira, B., Frangione, B., 1990. The complete amino acid sequence of toxin TsTX-VI isolated from the venom of the scorpion *Tityus serrulatus*. *J. Protein Chem.* 9, 595–601.
- Martin-Eauclaire, M.-F., Ceard, B., Bosmans, F., Rosso, J.-P., Tytgat, J., Bougis, P.E., 2005. New "Birtoxin analogs" from *Androctonus australis* venom. *Biochem. Biophys. Res. Commun.* 333, 524–530.
- Nascimento, D.G., Rates, B., Santos, D.M., Verano-Braga, T., Barbosa-Silva, A., Dutra, A.A.A., Biondi, I., Martin-Eauclaire, M.F., Lima, M.E.D., Pimenta, A.M.C., 2006. Moving pieces in a taxonomic puzzle: venom 2D-LC/MS and data clustering analyses to infer phylogenetic relationships in some scorpions from the Buthidae family (Scorpiones). *Toxicon* 47, 628–639.
- Pallaghy, P.K., Norton, R.S., Nielsen, K.J., Craik, D.J., 1994. A common structural motif incorporating a cystine knot and a triple-stranded  $\beta$ -sheet in toxic and inhibitory polypeptides. *Protein Sci.* 3, 1833–1839.
- Polis, G.A., 1990. *The Biology of Scorpions*. Stanford University Press, Stanford, CA, USA.
- Possani, L.D., Becerril, B., Delepierre, M., Tytgat, J., 1999. Scorpion toxins specific for  $Na^+$  channels. *Eur. J. Biochem.* 264, 287–300.
- Remijsen, Q., Verdonck, F., Willems, J., 2010. Parabutoporin, a cationic amphipathic peptide from scorpion venom: much more than an antibiotic. *Toxicon* 55, 180–185.
- Romi-Lebrun, R., Martin-Eauclaire, M.-F., Escoubas, P., Wu, F.Q., Lebrun, B., Hisada, M., Nakajima, T., 1997. Characterization of four toxins from *Buthus Martensi* scorpion venom, which act on apamin-sensitive calcium activated  $K^+$  channels. *Eur. J. Biochem.* 245, 457–464.
- Schwartz, E.F., Diego-García, E., Rodríguez de la Vega, R.C., Possani, L.D., 2007. Transcriptome analysis of the venom gland of the Mexican scorpion *Hadrurus gertschi* (Arachnida: Scorpiones). *BMC Genomics* 8, 119.
- Silva, E.C.N., Camargos, T.S., Maranhão, A.Q., Silva-Pereira, I., Silva, L.P., Possani, L.D., Schwartz, E.F., 2009. Cloning and characterization of cDNA sequences encoding for new venom peptides of the Brazilian scorpion *Opisthacanthus cayaporum*. *Toxicon* 54, 252–261.
- Sollod, B.L., Wilson, D., Zhaxybayeva, O., Gogarten, J.P., Drinkwater, R., King, G.F., 2005. Were arachnids the first to use combinatorial peptide libraries? *Peptides* 26, 131–139.
- Strugatsky, D., Zilberberg, N., Stankiewicz, M., Ilan, N., Turkov, M., Cohen, L., Pelhate, M., Gilles, N., Gordon, D., Gurevitz, M., 2005. Genetic polymorphism and expression of a highly potent scorpion depressant toxin enable refinement of the effects on insect  $Na$  channels and illuminate the key role of Asn-58. *Biochemistry* 44, 9179–9187.
- Tian, C., Yuan, Y., Zhu, S., 2008. Positively selected sites of scorpion depressant toxins: possible roles in toxin functional divergence. *Toxicon* 51, 555–562.
- Valdez-Cruz, N.A., Dávila, S., Licea, A., Corona, M., Zamudio, F.Z., García-Valdes, J., Boyer, L., Possani, L.D., 2004. Biochemical, genetic and physiological characterization of venom components from two species of scorpions: *Centruroides exilicauda* Wood and *Centruroides sculpturatus* Ewing. *Biochimie* 86, 387–396.
- Rodríguez de la Vega, R.C., Schwartz, E.F., Possani, L.D., 2010. Mining on scorpion venom biodiversity. *Toxicon* 56, 1155–1161.
- Weinberger, H., Moran, Y., Gordon, D., Turkov, M., Kahn, R., Gurevitz, M., 2010. Positions under positive selection—key for selectivity and potency of scorpion  $\alpha$ -toxins. *Mol. Biol. Evol.* 27, 1025–1034.
- Wu, J.-J., Dai, L., Lan, Z.-D., Chi, C.-W., 1999. Genomic organization of three neurotoxins active on small conductance  $Ca^{2+}$ -activated potassium channels from the scorpion *Buthus martensi* Karsch. *FEBS Lett.* 452, 360–364.
- Wulff, H., Kolski-Andreaco, A., Sankaranarayanan, A., Sabatier, J.-M., Shakkottai, V., 2007. Modulators of small- and intermediate-conductance calcium-activated potassium channels and their therapeutic indications. *Curr. Med. Chem.* 14, 1437–1457.
- Zhao, R., Ma, Y., He, Y., Di, Z., Wu, Y., Cao, Z., Li, W., 2010. Comparative venom gland transcriptome analysis of the scorpion *Lychas mucronatus* reveals intraspecific toxic gene diversity and new venomous components. *BMC Genomics* 11, 452.
- Zhijian, C., Feng, L., Yingliang, W., Xin, M., Wenxin, L., 2006. Genetic mechanisms of scorpion venom peptide diversification. *Toxicon* 47, 348–355.
- Zhu, S., Bosmans, F., Tytgat, J., 2004a. Adaptive evolution of scorpion sodium channel toxins. *J. Mol. Evol.* 58, 145–153.
- Zhu, X., Zamudio, F.Z., Olbinski, B.A., Possani, L.D., Valdivia, H.H., 2004b. Activation of skeletal ryanodine receptors by two novel scorpion toxins from *Buthus judaicus*. *J. Biol. Chem.* 279, 26588–26596.
- Zhu, S., Peigneur, S., Gao, B., Luo, L., Jin, D., Zhao, Y., Tytgat, J., 2009. Molecular diversity and functional evolution of scorpion potassium channel toxins. *Mol. Cell. Proteomics*, in press 10.1074/mcp.M110.002832.
- Zlotkin, E., Miranda, F., Rochat, H., 1978. Chemistry and pharmacology of Buthinae scorpion venoms. In: Bettini, S. (Ed.), *Arthropod Venoms*. Springer, Berlin Heidelberg, pp. 317–369.
- Zlotkin, E., Lester, D., Lazarovici, P., Pelhate, M., 1982. The chemistry and axonal action of two insect toxins derived from the venom of the scorpion. *Toxicon* 20, 323–331.

|                                  |                                                                                                                                                                       |     |     |     |     |     |
|----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----|-----|-----|-----|
|                                  | 20                                                                                                                                                                    | 40  | 60  | 80  | 100 | 120 |
| U <sub>2m</sub> -Buthitoxin-Hj1a | ACAGCAAAC <del>TT</del> CAATGATGGATATCCACTTAGAAACGATTGCCCAAT <del>TT</del> CATGCATACCGGATGATCATTATGAGCATTGCAACCAATTCTGCATAAAGAATCAAGCCGAATCAGG <del>TT</del> CATGTGAC |     |     |     |     |     |
| Telson_96-4_D02_T7.abi           | ACAGCAAAC <del>TT</del> CAATGATGGATATCCACTTAGAAACGATTGCCCAAT <del>TT</del> CATGCATACCGGATGATCATTATGAGCATTGCAACCAATTCTGCATAAAGAATCAAGCCGAATCAGG <del>TT</del> CATGTGAC |     |     |     |     |     |
| Telson_96-11_F11_T7.abi          | ACAGCAAAC <del>TT</del> CAATGATGGATATCCACTTAGAAACGATTGCCCAAT <del>TT</del> CATGCATACCGGATGATCATTATGAGCATTGCAACCAATTCTGCATAAAGAATCAAGCCGAATCAGG <del>TT</del> CATGTGAC |     |     |     |     |     |
| Telson_96-3_C08_T7.abi           | ACAGCAAAC <del>TT</del> CAATGATGGATATCCACTTAGAAACGATTGCCCAAT <del>TT</del> CATGCATACCGGATGATCATTATGAGCATTGCAACCAATTCTGCATAAAGAATCAAGCCGAATCAGG <del>TT</del> CATGTGAC |     |     |     |     |     |
| Telson_96-2_G02_M13-RP.abi       | ACAGCAAAC <del>TT</del> CAATGATGGATATCCACTTAGAAACGATTGCCCAAT <del>TT</del> CATGCATACCGGATGATCATTATGAGCATTGCAACCAATTCTGCATAAAGAATCAAGCCGAATCAGG <del>TT</del> CATGTGAC |     |     |     |     |     |
| Telson_96-12_G01_M13-RP.abi      | ACAGCAAAC <del>TT</del> CAATGATGGATATCCACTTAGAAACGATTGCCCAAT <del>TT</del> CATGCATACCGGATGATCATTATGAGCATTGCAACCAATTCTGCATAAAGAATCAAGCCGAATCAGG <del>TT</del> CATGTGAC |     |     |     |     |     |
| Telson_96-11_E06_M13-RP.abi      | ACAGCAAAC <del>TT</del> CAATGATGGATATCCACTTAGAAACGATTGCCCAAT <del>TT</del> CATGCATACCGGATGATCATTATGAGCATTGCAACCAATTCTGCATAAAGAATCAAGCCGAATCAGG <del>TT</del> CATGTGAC |     |     |     |     |     |
| Telson_96-12_G01_T7.abi          | ACAGCAAAC <del>TT</del> CAATGATGGATATCCACTTAGAAACGATTGCCCAAT <del>TT</del> CATGCATACCGGATGATCATTATGAGCATTGCAACCAATTCTGCATAAAGAATCAAGCCGAATCAGG <del>TT</del> CATGTGAC |     |     |     |     |     |
| Telson_96-4_D02_M13-RP.abi       |                                                                                                                                                                       |     |     |     |     |     |
| Telson_96-3_C08_M13-RP.abi       |                                                                                                                                                                       |     |     |     |     |     |
| Telson_96-2_G02_T7.abi           |                                                                                                                                                                       |     |     |     |     |     |
|                                  | 140                                                                                                                                                                   | 160 | 180 | 200 | 220 | 240 |
| U <sub>2m</sub> -Buthitoxin-Hj1a | TTCGACGCAGATGCATGTAAATGTTGGGGTATNCCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATNCCAGAGTTGGAACCAGAATNTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-4_D02_T7.abi           | TTCGACGCAGATGCATGTAAATGTTGGGGTATACCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATGCCAGAGTTGGAACCAGAATCTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-11_F11_T7.abi          | TTCGACGCAGATGCATGTAAATGTTGGGGTATACCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATGCCAGAGTTGGAACCAGAATCTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-3_C08_T7.abi           | TTCGACGCAGATGCATGTAAATGTTGGGGTATACCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATGCCAGAGTTGGAACCAGAATCTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-2_G02_M13-RP.abi       | TTCGACGCAGATGCATGTAAATGTTGGGGTATACCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATGCCAGAGTTGGAACCAGAATCTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-12_G01_M13-RP.abi      | TTCGACGCAGATGCATGTAAATGTTGGGGTATACCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATGCCAGAGTTGGAACCAGAATCTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-11_E06_M13-RP.abi      | TTCGACGCAGATGCATGTAAATGTTGGGGTATACCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATGCCAGAGTTGGAACCAGAATCTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-12_G01_T7.abi          | TTCGACGCAGATGCATGTAAATGTTGGGGTATACCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATGCCAGAGTTGGAACCAGAATCTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-4_D02_M13-RP.abi       | TTCGACGCAGATGCATGTAAATGTTGGGGTATACCATCTGAAATGAAAATATGGAACCCTAAATCATCAAAATGCCAGAGTTGGAACCAGAATNTCATAACAAATTTTTTTAACTTGAAGATTAA                                         |     |     |     |     |     |
| Telson_96-3_C08_M13-RP.abi       |                                                                                                                                                                       |     |     |     |     |     |
| Telson_96-2_G02_T7.abi           |                                                                                                                                                                       |     |     |     |     |     |

|                                  |                                                                                                                 |     |     |     |     |
|----------------------------------|-----------------------------------------------------------------------------------------------------------------|-----|-----|-----|-----|
|                                  | 20                                                                                                              | 40  | 60  | 80  | 100 |
| U <sub>2</sub> m-Buthitoxin-Hj1c | TTATTTACTACAATGGTTTTTAAGTGCATCGAAAATATTGATGTCTTCTTCTTTCAAATATTCACACCAACACGAGGTGACCCAACAATACCCATAATTCACTCCATG    |     |     |     |     |
| Telson-96-11_D05_M13-RP          | TTATTTACTACAATGGTTTTTAAGTGCATCGAAAATATTGATGTCTTCTTCTTTCAAATATTCACACCAACACGAGGTGACCCAACAATACCCATAATTCACTCCATG    |     |     |     |     |
| SKORPION1_H05_T7_044             | TTATTTACTACAATGGTTTTTAAGTGCATCGAAAATATTGATGTCTTCTTCTTTCAAATATTCACACCAACACGAGGTGACCCAACAATACCCATAATTCACTCCATG    |     |     |     |     |
|                                  | 120                                                                                                             | 140 | 160 | 180 | 200 |
| U <sub>2</sub> m-Buthitoxin-Hj1c | TACGCGGGAAGCAGAAATGGAATCATTGGCCATCGTGT CATAGAAGTATTGAGTGATGAAAAC TTACATTCTGATTTTGGAAACAATTCCTTCTGACGTGGAAACTTCA |     |     |     |     |
| Telson-96-11_D05_M13-RP          | TACGCGGGAAGCAGAAATGGAATCATTGGCCATCGTGT CATAGAAGTATTGAGTGATGAAAAC TTACATTCTGATTTTGGAAACAATTCCTTCTGACGTGGAAACTTCA |     |     |     |     |
| SKORPION1_H05_T7_044             | TAC                                                                                                             |     |     |     |     |

|                            |                                                                                |    |  |     |  |     |
|----------------------------|--------------------------------------------------------------------------------|----|--|-----|--|-----|
|                            |                                                                                | 20 |  | 40  |  | 60  |
| κ-Buthitoxin-Hj1b          | ACGCGGGGGTATCACAGATTCACAACAATACACCAGCAATTATGCCTGATTCGAAAGTAGAAGCTGTTGGTTGTGAAR |    |  |     |  |     |
| Telson_96-2_F03_T7.abi     | ACGCGGGGGTATCACAGATTCACAACAATACACCAGCAATTATGCCTGATTCGAAAGTAGAAGCTGTTGGTTGTGAA  |    |  |     |  |     |
| Telson_96-4_C08_M13-RP.abi | ACGCGGGGGTATCACAGATTCACAACAATACACCAGCAATTATGCCTGATTCGAAAGTAGAAGCTGTTGGTTGTGAAG |    |  |     |  |     |
| Telson_96-2_F03_M13-RP.abi | CCAGCAATTATGCCTGATTCGAAAGTAGAAGCTGTTGGTTGTGAA                                  |    |  |     |  |     |
|                            |                                                                                | 80 |  | 100 |  | 120 |
|                            |                                                                                |    |  |     |  | 140 |
| κ-Buthitoxin-Hj1b          | ATTGCCCTGAGCACTGTTCCAGCAAAATGCCCGAGCAAAATGTGAAAATGACAAATGTGTATGCGAACCTAAATGA   |    |  |     |  |     |
| Telson_96-2_F03_T7.abi     | ATTGCCCTGAGCACTGTTCCAGCAAAATGCCCGAGCAAAATGTGAAAATGACAAATGTGTATGCGAACCTAAATGA   |    |  |     |  |     |
| Telson_96-4_C08_M13-RP.abi | ATTGCCCTGAGCACTGTTCCAGCAAAATGCCCGAGCAAAATGTGAAAATGACAAATGTGTATGCGAACCTAAATGA   |    |  |     |  |     |
| Telson_96-2_F03_M13-RP.abi | ATTGCCCTGAGCACTGTTCCAGCAAAATGCCCGAGCAAAATGTGAAAATGACAAATGTGTATGCGAACCTAAATGA   |    |  |     |  |     |