




Article

# Dissecting Toxicity: The Venom Gland Transcriptome and the Venom Proteome of the Highly Venomous Scorpion *Centruroides limpidus* (Karsch, 1879)

Jimena I. Cid-Urbe <sup>1</sup>, Erika P. Meneses <sup>2</sup>, Cesar V. F. Batista <sup>2</sup>, Ernesto Ortiz <sup>1,\*</sup>  
and Lourival D. Possani <sup>1,\*</sup>

<sup>1</sup> Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Avenida Universidad 2001, Cuernavaca, Morelos 62210, Mexico; jcidu@ibt.unam.mx

<sup>2</sup> Laboratorio Universitario de Proteómica, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Avenida Universidad 2001, Cuernavaca, Morelos 62210, Mexico; pkas@ibt.unam.mx (E.P.M.); fbatista@ibt.unam.mx (C.V.F.B.)

\* Correspondence: erne@ibt.unam.mx (E.O.); possani@ibt.unam.mx (L.D.P.)

Received: 12 April 2019; Accepted: 26 April 2019; Published: 30 April 2019



**Abstract:** Venom glands and soluble venom from the Mexican scorpion *Centruroides limpidus* (Karsch, 1879) were used for transcriptomic and proteomic analyses, respectively. An RNA-seq was performed by high-throughput sequencing with the Illumina platform. Approximately 80 million reads were obtained and assembled into 198,662 putative transcripts, of which 11,058 were annotated by similarity to sequences from available databases. A total of 192 venom-related sequences were identified, including Na<sup>+</sup> and K<sup>+</sup> channel-acting toxins, enzymes, host defense peptides, and other venom components. The most diverse transcripts were those potentially coding for ion channel-acting toxins, mainly those active on Na<sup>+</sup> channels (NaScTx). Sequences corresponding to β- scorpion toxins active of K<sup>+</sup> channels (KScTx) and λ-KScTx are here reported for the first time for a scorpion of the genus *Centruroides*. Mass fingerprint corroborated that NaScTx are the most abundant components in this venom. Liquid chromatography coupled to mass spectrometry (LC-MS/MS) allowed the identification of 46 peptides matching sequences encoded in the transcriptome, confirming their expression in the venom. This study corroborates that, in the venom of toxic buthid scorpions, the more abundant and diverse components are ion channel-acting toxins, mainly NaScTx, while they lack the HDP diversity previously demonstrated for the non-buthid scorpions. The highly abundant and diverse antareases explain the pancreatitis observed after envenomation by this species.

**Keywords:** *Centruroides limpidus* Karch; proteome; scorpion; transcriptome; venom toxicity

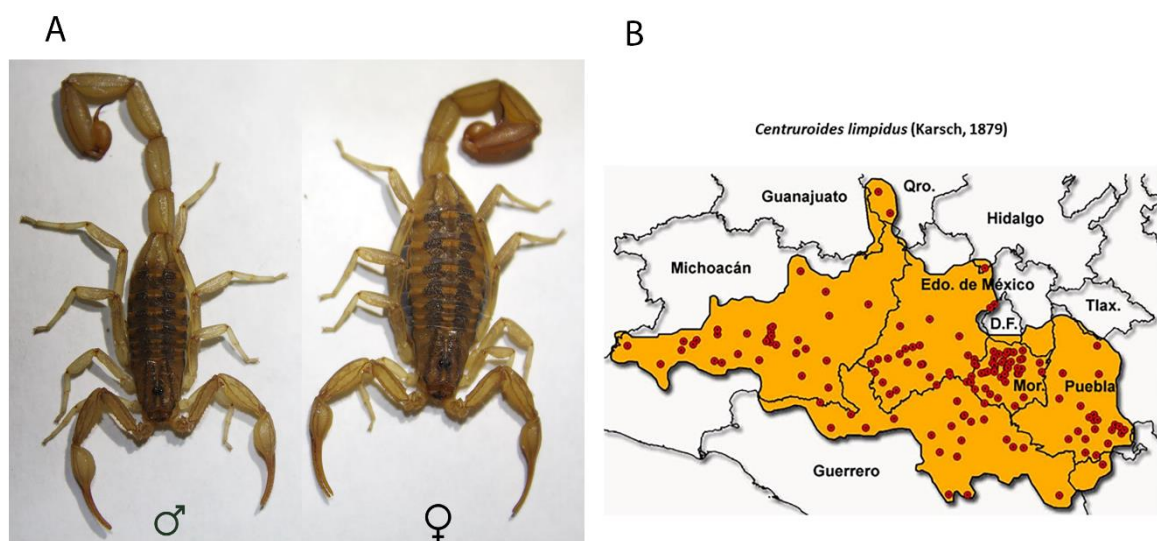
**Key Contribution:** A detailed molecular dissection of the venom of the highly toxic buthid scorpion *Centruroides limpidus* through transcriptomic and proteomic analyses is reported. Ion channel-acting toxins are shown to be the most abundant and diverse components of the venom, especially the NaScTx. Two new families of KScTx are reported for the first time in a scorpion of the genus *Centruroides*: β-KScTx and λ-KScTx. Zn-metalloproteases of the antarease family are present in the venom, providing the molecular basis for the observed pancreatitis after envenomation by this species.

## 1. Introduction

Scorpion venoms are known to contain hundreds of pharmacologically active components, affecting many other animals, which constitute their preys, competitors and/or predators [1]. They have a cosmopolitan distribution with 2415 distinct species reported up to December 2018 [2], which are classified into twenty different families. The scorpions belonging to the family Butidae produce the

most active toxins that affect mammals, including humans [3,4]. Mexico harbors approximately 12% of the world's diversity of scorpions, with 281 different species thus far described, of which 21 are dangerous to humans [5]. Their medical importance drove the initial research towards the isolation and biochemical characterization of the toxic compounds present in their venoms, and the identification of their physiological effects. The more recent development of high-throughput methods for massive sequencing and mass spectrometry has impacted the number of identified venom components [6–13].

In Mexico, the scorpions dangerous to humans belong to the genus *Centruroides*, comprising 42 different species [4,5]. The best studied species are: *Centruroides noxius*, *Centruroides suffusus*, *Centruroides tecomanus*, and *Centruroides limpidus*. *Centruroides limpidus* (Karsch, 1879) (Figure 1A) is widely distributed in densely populated areas of Central Mexico, including the States of Guerrero, Morelos, Mexico State, Michoacán, Queretaro, Hidalgo, and Puebla (Figure 1B) [14], and has a tendency to live side by side with humans [15]. It produces a potent venom which is highly toxic for mammals (the LD<sub>50</sub> in mice is approximately 15 µg/20 g [16]) and poses a serious threat to human life. Even though it is practically impossible to single out the species involved in each envenomation case, *C. limpidus* and other closely related species of the *Centruroides* genus with which it shares its habitat, are responsible for over 120,000 reported accidents with humans in those states alone every year—a third of all scorpionism cases in the country [17]. Those high morbidity numbers alone justify the efforts made toward the characterization of the venom components responsible for human intoxication. The identification of the toxins present in this venom should impact, for example, the research aimed at the production of antivenoms [18]. Thus far, for *C. limpidus* alone, nine peptides have been biochemically characterized, and nine precursors for other putative toxins have been described [19–23]. It is clear, however, that the characterization of this venom is far from complete, as information on novel, potentially lethal toxins from this species continues to emerge [23]. A more comprehensive study was therefore required.



**Figure 1.** Habitus and distribution of *Centruroides limpidus*. (A) The morphology of *C. limpidus*, male (left) and female (right). (B) Geographical distribution of *C. limpidus* in 10 Mexican States (red dots indicate the places of sampling). (B) Reproduced with permission from [14] Copyright 2009, Universidad Nacional Autónoma de México.

This communication reports the transcriptomic and proteomic analyses of the venom glands and the soluble venom of the scorpion *C. limpidus*, respectively. The transcriptomic results showed 192 transcripts encoding proteins/peptides with sequences identified as authentic venom components. Among these sequences, five main categories of precursors were identified: toxins, host defense peptides (HDPs), protease inhibitors, enzymes, and other components. The peptides encoded by the 46 transcripts identified in the transcriptome analysis were confirmed to be expressed in the venom

by LC-MS/MS sequencing. They showed sequence similarity with previously reported scorpion venom components.

The relative abundance and diversity of ion channel-acting toxins, in particular those active on Na<sup>+</sup> channels, is reported, confirming their responsibility in venom toxicity for mammals. The presence of several transcripts coding for antareases provide molecular support for the observed pancreatitis resulting from this species' envenomation. Toxins from families never before identified in *Centruroides* venoms are also reported.

## 2. Results and Discussion

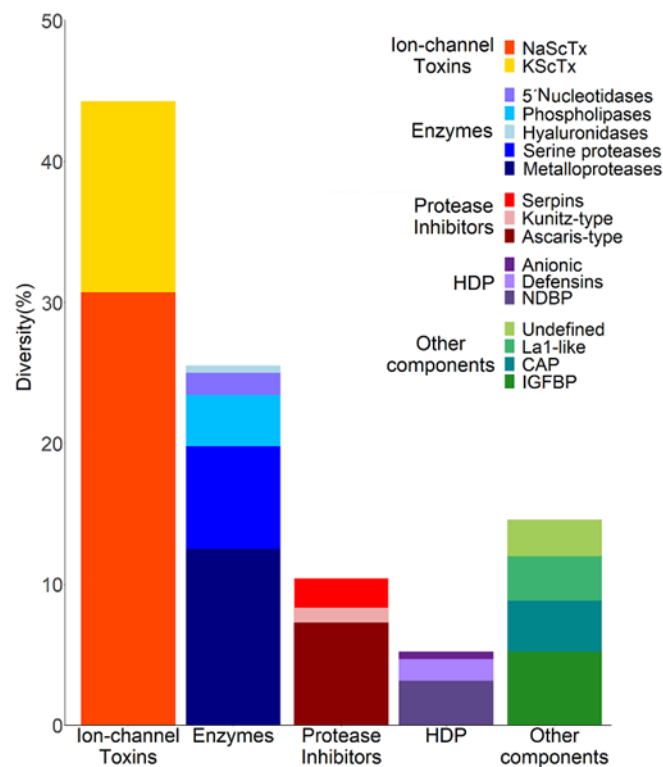
### 2.1. RNA Isolation, Sequencing, and Assembly

Two groups of scorpions separated by gender (5 males and 5 females) were used for total RNA isolation. A total of 9 µg (males) and 8 µg (females) of RNA was obtained and its quality was assessed with a Bioanalyzer 2100 (Agilent), as reported in other transcriptomic analyses [7,10,12,24,25]. The results indicated that the RNA samples were not degraded, even though a single RNA band corresponding to the mitochondrial 18S was observed, an effect previously reported for other organisms [7,10,12,24,25]. Paired-end cDNA libraries were prepared for males and females separately and sequenced using the Illumina platform (2 × 72 bp reads). The two genders were sequenced independently to fulfill the requirements of a related project that focuses on the differential expression of venom components in both genders [26], and which the results of will be published elsewhere. Two datasets with 38,364,311 and 41,366,601 reads were obtained for males and females, respectively, which were submitted to the European Nucleotide Archive (ENA), under project PRJEB31683. In order to have a global, species-specific transcriptomic analysis, comparable with other published scorpion transcriptomes, both datasets were merged and jointly analyzed henceforth. After de novo RNA-seq assembly with the Trinity software, 198,662 putative transcripts were obtained, with an N50 of 1611. The annotation was performed with the Trinotate software, resulting in 11,058 annotated transcripts that were identified by sequence similarity with sequences deposited in the Uniprot database. The divergence between the large number of assembled transcripts and the smaller subset of sequences with annotation, reflects the lack of information on many scorpion venom components, and reinforces the need for further biochemical and functional characterization of the scorpion venoms. Among the annotated transcripts, 366 sequences had similarity with arachnid sequences, 227 were specific for scorpions, and 192 sequences were from components related to venom, in particular. The 192 venom-related sequences were analyzed by BLAST to identify the closest related matches by sequence. The transcripts described in this study were labeled following the previously suggested standard for transcript nomenclature [10]. The species code for *C. limpidus* was set to "Cli", the existing family and subtype codes were observed, and new ones were added for components not previously found in scorpion venoms. A complete listing of the 192 sequence IDs, organized by type/subtype, the ID of the closest matches (with E values), and their translated Open Reading Frame (ORF) is provided in Supplementary Materials Table S1. For the reference sequences, the original names found in the databases were honored.

### 2.2. The Diversity of Transcripts Related to Venom Components in the Venom Gland of *C. limpidus*

The identification of transcripts putatively coding for venom peptides/proteins in *C. limpidus* was initially based on Pfam domains [27]. The search for matching domains was not limited to scorpions, but broadened to include all venomous animals [28]. Four major categories were identified by this method: toxins, host defense peptides, enzymes, and protease inhibitors. A fifth heterogeneous category denominated as "other components" was designated to include annotated sequences without defined domain and function, as well as sequences with well-conserved domains, but without a defined function in venoms. Figure 2 shows the distribution of the annotated venom-related sequences grouped into these five categories, as percentages of the total 192. For this figure, the number of

individual sequences was considered, rather than their relative abundance in the transcriptome (and therefore the term “Diversity” was used in the graph). Each major category was further divided into subcategories, determined by the more specific, putative structural or functional classification of the peptides/proteins. The transcripts were also globally classified in accordance to Gene Ontology (GO) terms [29] (Supplementary Materials Figure S1).



**Figure 2.** Relative diversity of transcripts related to venom components. For the graphic, only the number of different transcripts identified for each category.

### 2.2.1. Ion Channel-Acting Toxins

Hundreds of toxic components have been described to be present in scorpion venoms [1]. Many of these are peptides known to affect ion channels from mammals, birds or arthropods (e.g., arachnids, insects or crustaceans), including sodium, potassium, calcium and chloride ion channels [30].

As shown in Figure 2, this category was the most diverse in terms of sequence in the transcriptome of *C. limpidus*, which did not come as a surprise, given the known high toxicity of this species to animals from different taxa. Eighty-five distinct sequences with conserved motifs or domains for ion channel-acting toxins were found, including 59 transcripts potentially coding for sodium channel-acting toxins (NaScTx) and 26 for putative potassium channel-acting toxins (KScTx). This is by far the largest number of toxin sequences reported for a scorpion of any species to date. Supplementary Materials Table S2 contains all the transcripts putatively coding for ion channel-acting toxins, together with the reference protein/transcript (the best match by sequence similarity), the source of the reference, plus its function if known.

#### Toxins Active on Voltage-Gated Sodium Channels (NaScTx)

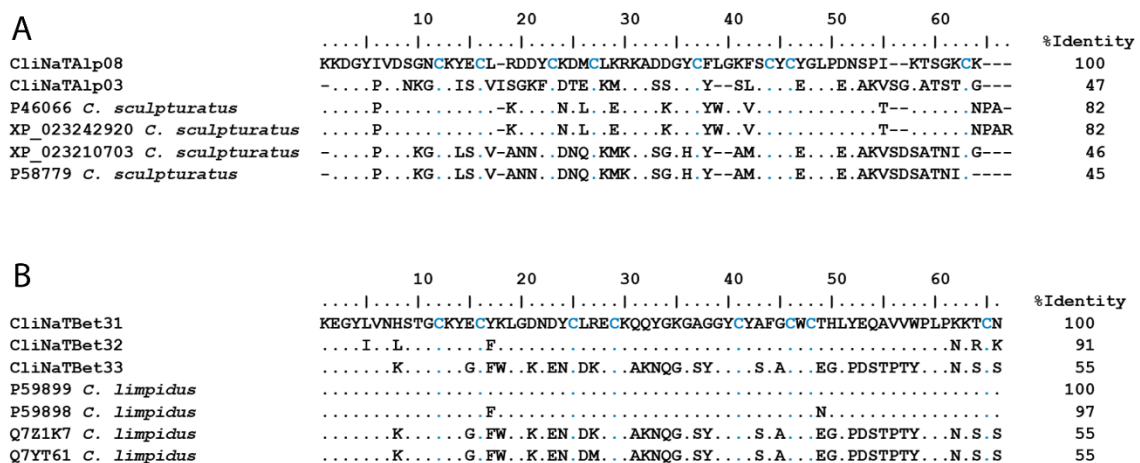
Due to their central role in the intoxication process following envenomation, the scorpion toxins that affect voltage-gated sodium channels are the best characterized venom components, both biochemically and functionally. They are broadly subdivided into  $\alpha$ -NaScTx and  $\beta$ -NaScTx, the former being known for slowing down the  $\text{Na}^+$  channel inactivation and binding to receptor site-3, while the latter shifts

the channels’ opening kinetics to more negative potentials and bind to receptor site-4 [31]. Scorpion  $\beta$ -NaScTx are further classified as classical (active on mammalian Na<sup>+</sup> channels), excitatory anti-insect, depressant anti-insect, and  $\beta$ -like (compete for binding sites on both insect and mammalian Na<sup>+</sup> channels) toxins [32].

The genome sequences from a scorpion of the same genus, *Centruroides sculpturatus*, have been recently released as BioProject PRJNA422877 on NCBI. As expected, many of the transcripts recovered in this study had sequence similarity with sequences from *C. sculpturatus*. Sequences matching those of other scorpions were also found, not only of the *Centruroides* genus, but also from old-world scorpions of the genera *Isometrus*, *Lychas* and *Parabuthus*. In the particular case of the putative NaScTx, 45 transcripts were found to be similar to sequences from *C. sculpturatus*, seven to previously reported sequences from *C. limpidus*, three to those from *C. noxius*, and one to a similar sequence from each of the following species, *C. exilicauda*, *Centruroides vittatus*, *Parabuthus transvaalicus*, and *Lychas mucronatus* (Supplementary Materials Table S2).

The NaScTx are usually more abundant in scorpions of the family Buthidae, which produce highly neurotoxic venoms. Three previous high-throughput sequencing-derived transcriptomic analyses with scorpions belonging to this family (*Tityus bahiensis*, *Centruroides hentzi*, and *C. noxius* [6,8,33]), reported 27 to 38 transcripts potentially coding for NaScTx. This study found 59 transcripts of this kind, which represents the largest diversity so far described for scorpions of this, or any other taxonomic family. Of them, 16 transcripts putatively code for  $\alpha$ -NaScTx and 43 for  $\beta$ -NaScTx (Supplementary Materials Table S2).

Figure 3A shows, as an example, a multiple-sequence alignment of CliNaTAlp03 and CliNaTAlp08, both annotated as possible  $\alpha$ -NaScTx, with their best matches in terms of sequence similarity. Only the sequence region corresponding to the predicted mature toxin was used for the alignment, since the signal and propeptide regions were missing in the sequences of the peptides obtained directly from the venom. The peptides encoded by CliNaTAlp03 and CliNaTAlp08 were confirmed to be present in the venom by the proteomic analysis, and were the  $\alpha$ -NaScTx with the highest identification scores by that analysis. CliNaTAlp03 matched with “precursor alpha-like toxin CsEv5” (XP\_023210703) and the peptide itself (P58779) [34]. CliNaTAlp08 matched with “precursor alpha-toxin CsE5” (XP\_023242920) and the peptide itself (P46066) [35]. Here, and in all further alignments, the newly reported *C. limpidus* sequences are shown on top of the alignment for clarity.



**Figure 3.** Examples of Na<sup>+</sup> channel-specific toxins. (A) Putative  $\alpha$ -NaScTx and (B) Putative  $\beta$ -NaScTx. The transcript-derived peptides are aligned to their reference sequences. The percentage of identity was calculated considering only the mature sequences. Accession numbers and species’ names of the references were taken from UniProt or GenBank. Conserved cysteine residues are highlighted in blue. Dots indicate identical residues, and dashes indicate gaps.

Three transcripts, CliNaTxBet31, CliNaTxBet32, and CliNaTxBet33 are shown in Figure 3B, aligned to the best matching reference sequences (mature peptides only). CliNaTxBet31 codes for CII2b (P59899), a peptide found in the *C. limpidus* venom [21], which was confirmed by the proteomic analysis in this study (see below). CII2b is toxic to mice and active on sodium and calcium channels in cultured chick dorsal root ganglion cells [21]. CliNaTxBet32 potentially codes for a peptide similar to CII2b, but with six amino acid changes. CII2b is very close in sequence to CII2 (P59898), one of the major toxins of the *C. limpidus* venom [18,23], having only two differences at the amino acid level. CII2 is a highly neurotoxic peptide for mammals [19]. CliNaTxBet33 matches the precursor of CII5b (Q7Z1K7) and was here confirmed to be present in the venom by the proteomic analysis. A very similar sequence, CII5c (Q7YT61), which differs from CII5b by just one residue, is also included in the alignment. The remaining 40 transcripts found in this study are described in Supplementary Materials Table S2. These include four sequences with similarity to other previously reported sequences from *C. limpidus*.

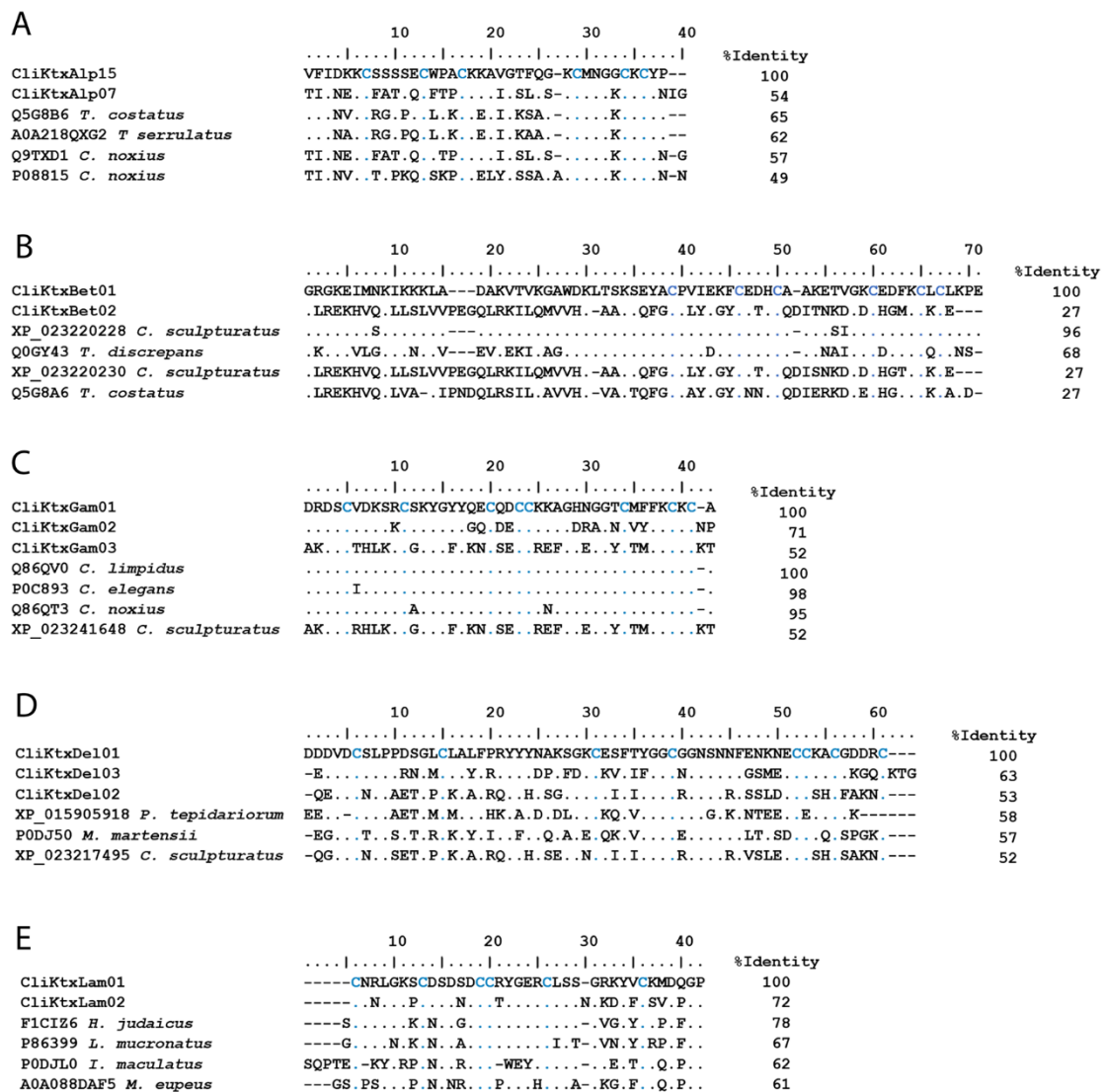
#### Toxins Active on Potassium Channels (KScTx)

Potassium channels are the most widely distributed type of ion channels, found in basically all living organisms, where they play fundamental roles in the physiology of the cells [36–38]. Scorpion venoms contain toxins that affect these channels (KScTx), as demonstrated with insect and mammalian channels [37]. The KScTx have been classified into seven families ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\kappa$ , and  $\lambda$ ) depending on their amino acid sequences, length, and 3D-structure [39–41]. They have been shown to be active on voltage-gated potassium channels (such as Kv1, Kv3, Kv4, Kv7, Kv11) and calcium-activated potassium channels (KCa1.1 or BK) [37].

This study reports 26 transcripts with sequence similarity to members of five families of KScTx: 15 sequences potentially coding for  $\alpha$ -KScTx, 2 for  $\beta$ -KScTx, 3 for  $\gamma$ -KScTx, 3 for  $\delta$ -KScTx, and 2 for  $\lambda$ -KScTx (Supplementary Materials Table S1).

Figure 4A shows two examples of the peptide sequences derived from the transcriptomic analysis in this study, which potentially correspond to  $\alpha$ -KScTx. CliKTxAlp15 codes for a peptide confirmed to be expressed in the venom of *C. limpidus* by the proteomic analysis here reported. It is similar to alpha-KTx4.5 from *Tityus costatus* (Q5G8B6), a toxin found to inhibit with low potency the Kv1.1, Kv1.2, Kv1.3, and Kv11.1 (ERG1) channels [42]. CliKTxAlp07 potentially codes for a peptide similar to Noxiustoxin-2 (Q9TXD1), having only one difference at the amino acid level. Noxiustoxin-2 has a paralyzing effect on crickets but is not toxic to mice or crustaceans [43]. Two other similar sequences were included in the alignment, A0A218QXG2, a nucleotide sequence reported from *Tityus serrulatus* (only the predicted, translated mature region was included), and Noxiustoxin (P08815), a toxin from *C. noxius* that blocks several Kv and KCa channels [44]. The remaining 14 transcripts found in this study are summarized in Supplementary Materials Table S2, including CliKTxAlp10, which is similar to a previously reported toxin from this scorpion, ClITx1 (P45629).

One very relevant finding of this work is the description, for the first time, of potential  $\beta$ -KScTx in a scorpion from the *Centruroides* genus. Toxins of this family have been found in the venoms of buthid and non-buthid scorpions from the genera *Androctonus*, *Euscorpions*, *Heterometrus*, *Hoffmannihadrurus*, *Liocheles*, *Mesobuthus*, *Pandinus*, and *Tityus*, but never before in *Centruroides*. Toxins of the  $\beta$ -KScTx family are active on Kv1.1, Kv1.3 or Kv4.2 mammalian channels [45]. Two transcripts were found in the venom gland of *C. limpidus* coding for peptides that were confirmed to be expressed in the venom by the proteomic analysis. CliKTxBet01 displayed sequence similarity with a DNA sequence from *C. sculpturatus* (XP\_023220228) and with toxin TdiKIK (Q0GY43) from *Tityus discrepans*. CliKTxBet02 was similar to another DNA sequence from *C. sculpturatus* (XP\_023220230) and to the “Scorpine-like peptide Tco 41.46-2” from *T. costatus*. The putative  $\beta$ -KScTx here described are shown aligned to the reference sequences (only the predicted mature sequences) in Figure 4B.



**Figure 4.** Examples of K<sup>+</sup> channel-specific toxins. (A–E) show the alignments of the transcript-derived sequences of members of the  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ , and  $\lambda$ -KScTx families, respectively, with the reference sequences. The percentage of identity was calculated considering only the mature sequences. Accession numbers and species' names of the references were taken from UniProt or GenBank. Conserved cysteine residues are highlighted in blue. Dots indicate identical residues, and dashes indicate gaps.

The  $\gamma$ -KScTx are short-chain peptides of 36–47 amino acid residues. Structurally, they have a cysteine-stabilized  $\alpha\beta$  (CS $\alpha/\beta$ ) motif, with three or four disulfide bridges. The  $\gamma$ -KScTx are capable of blocking the ERG K<sup>+</sup> channels. Three new sequences, potentially coding for  $\gamma$ -KScTx are reported here (Figure 4). CliKtxGam01 codes for a peptide identical to CIIerg1 (Q86QV0) isolated from the venom of *C. limpidus*, though never tested on ion channels. It is also closely related to CeErgTx5, from *Centruroides elegans*, differing in just one amino acid. CeErgTx5 is active on mammalian ERG K<sup>+</sup> channels [46]. CliKtxGam02 is similar to CnErg1 (Q86QT3) identified in the venom of *C. noxius*, a peptide with activity on mammalian ERG K<sup>+</sup> channels [47]. CliKtxGam03 is closest in sequence to “potassium channel toxin gamma-KTx 1.1-like” (XP\_023241648), derived from the genome sequences of *C. sculpturatus* (only the predicted mature sequence was considered in the alignment in Figure 4C). The three putative sequences, here reported, contained the functionally relevant K13, involved in the interaction with the outer vestibule of the channel [48]. Two other residues, demonstrated to be relevant for the interaction of  $\gamma$ -KScTx with the channels, are Q18 and M35 [49]. Only CliKtxGam01

has those residues conserved (Figure 4C), which might have implications in the functionality of the peptides derived from CliKTxGam02 and CliKTxGam03 if they happen to be indeed expressed in the venom.

The  $\delta$ -KScTx are peptides with 59–70 amino acid residues. They are structurally characterized by the presence of a CS $\alpha$ / $\beta$  motif of the Kunitz type, stabilized by three disulfide bonds [50]. They display a dual activity: as serine protease inhibitors and as blockers of the Kv channels, mainly the Kv1.3, though other channels can also be weakly inhibited. Three different transcripts are here described, which are shown aligned to reference sequences in Figure 4D. CliKTxDel01 is similar to “Kunitz-type serine protease inhibitor BmKTT-2” (P0DJ50), a peptide from *M. martensii* which completely inhibits trypsin and blocks the murine Kv1.3. CliKTxDel02 shares sequence similarity with a genome sequence from *C. sculpturatus* labeled as “iso-inhibitor K-like” (XP\_023217495). CliKTxDel03 was found to be similar to a genome sequence from the spider *Paraestotoda tepidariorum* annotated as “hemolymph trypsin inhibitor B-like isoform X2” (XP\_015905918). The genome-derived sequences used in the alignment in Figure 4 are limited to the predicted mature sequences.

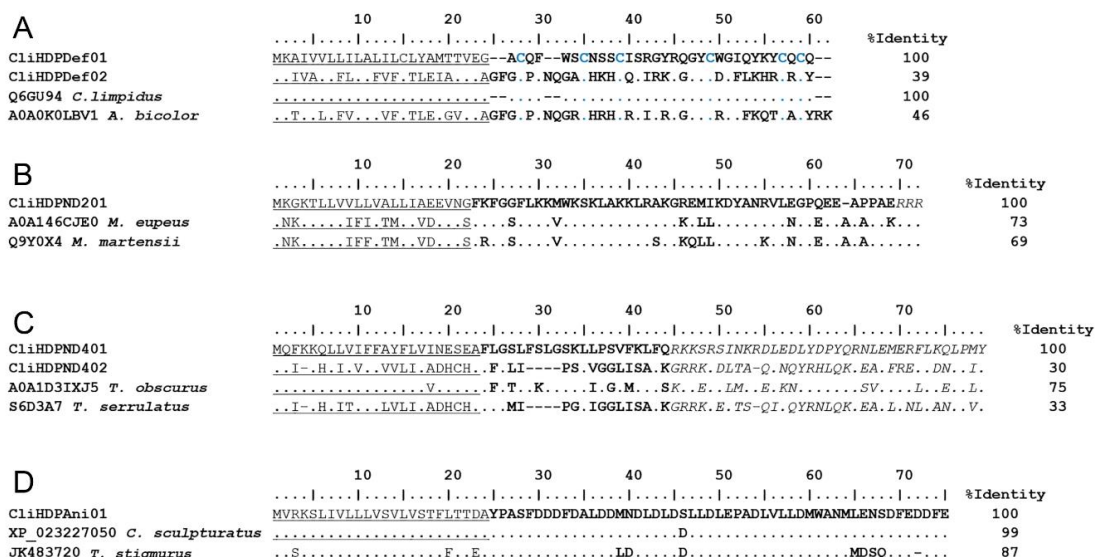
It is very relevant that no transcripts coding for calcium channel-specific toxins of the calcin family (active on the ryanodine receptors of mammalian cardiac or skeletal muscle cells) were found in the analysis, while two transcripts are here reported for the phylogenetically related, insect-specific  $\lambda$ -KScTx family. The  $\lambda$ -KScTx are short (approximately 40 amino acids) peptides which adopt an inhibitor cystine knot (ICK) fold. Both CliKTxLam01 and CliKTxLam02 code for the first  $\lambda$ -KScTx ever described for the genus *Centruroides*. Their finding in a buthid scorpion confirms the proposition that calcins and  $\lambda$ -KScTx are specific for non-buthid families and the buthid family, respectively, being mutually exclusive in those venoms [51]. In order to confirm that CliKTxLam01 and CliKTxLam02 are indeed  $\lambda$ -KScTx and not the structurally related calcins, these two sequences were incorporated into a phylogenetic analysis with those in Reference [51] by Carlos Santibáñez-López (see Acknowledgements) and shown to group with other buthid  $\lambda$ -KScTx and not the non-buthid calcins (Supplementary Materials Figure S2). CliKTxLam01 has sequence similarity with “phi-buthitoxin-Hj1a” (F1CIZ6), derived from a transcript from *Hottentotta judaicus*. CliKTxLam02 matched by sequence similarity the “potassium channel blocker pMeKTx30-1” (A0A088DAF5), derived from a transcript from *Mesobuthus eupeus*. These sequences are shown aligned in Figure 4E. Only two tested K<sup>+</sup> channel blockers belonging to the  $\lambda$ -KScTx family have been reported thus far: ImKTx1 from *Isometrus maculatus* (P0DJL0) [52] and Neurotoxin lambda-MeuTx from *M. eupeus* (P86399) [53]. They are also included in the alignment for reference.

### 2.2.2. Host Defense Peptides (HDPs)

Arachnid venoms are rich sources of host defense peptides (HDPs) [54]. They are characterized by having a broad spectrum of biological activities, including antimicrobial [55–57], insecticidal [58], bradykinin-potentiating [59], antitumoral [60,61], and hemolytic [62], among others. Host defense peptides have been demonstrated to be abundant and highly diverse in the venoms of scorpions belonging to non-Buthidae families [7,24,25], though a previous study with high-throughput sequencing techniques also identified HDPs in the buthid *C. hentzi* [8]. Host defense peptides are divided into two categories: non-disulfide-bridged-peptides (NDBPs) and cysteine-stabilized  $\beta$ -sheet-rich peptides, which include the defensins and scorpines [54]. Though just one HDP (a defensin whose expression is induced in the hemolymph in response to septic injury) had been reported from *C. limpidus* [63], a mass fingerprint of the venom revealed components with lower molecular weights than KScTx, indicating the potential presence of HDPs [26]. This analysis confirms that conclusion. Ten transcripts coding for HDPs were found: six defensins, one NDBP-2, two NDBP-4, and one anionic peptide. Nevertheless, compared to the above referenced studies by massive RNA sequencing of non-buthid scorpions, the diversity here described for the HDPs was significantly lower, confirming the previous empiric observation that, contrary to the ion channel-acting toxins, HDP are more diverse in non-buthid scorpions than in buthids.



CliHDPDef01 is actually the transcript coding for Cll-dlp (Q6GU94), the previously reported hemolymph defensin. It is notable that the same defense peptide is expressed in both tissues. The remaining five defensin transcripts, CliHDPDef02–CliHDPDef06, have similarity to Defensin-1 (A0A0K0LBV1), a transcript from *Androctonus bicolor* [64]. Both CliHDPDef01 and CliHDPDef02 are shown in Figure 5A, aligned to the reference sequences. Since the complete precursors for all the sequences are available, they were used in the alignment and included in the calculated percentage of identity.



**Figure 5.** Examples of host defense peptides (HDPs). The translated sequences from representative transcripts coding for HDPs found in the *C. limpidus* transcriptome are shown aligned to matching sequences from databases. The complete precursor sequences are shown and the complete precursor was considered in the calculation of the percentage of identity. (A) Defensins. (B) The unique NDBP-2 precursor found. (C) Two putative NDBP-4 precursors. (D) The unique precursor for the putative anionic peptide. Accession numbers and species' names of the references were taken from UniProt or GenBank. Conserved cysteine residues are highlighted in blue. The predicted mature sequences are indicated in bold typeface, the predicted signal peptides are underlined and the propeptides are indicated in italics. Dots indicate identical residues and dashes indicate gaps.

Transcripts with sequence similarity to members of families NDBP-2 and NDBP-4 were also found. CliHDPND201 had similarity to “venom toxin meuTx20” (A0A146CJE0) deduced from a transcript from *M. eupeus*, and with BmKbpp (Q9Y0X4), a peptide from *M. martensii* with antimicrobial and bradykinin-potentiating activities against bacteria and fungi [65]. The precursor sequence of these three molecules are shown aligned in Figure 5B. CliHDPND401 has similarity to ToAP2 (A0A1D3IXJ5), a transcript isolated from the venom of *Tityus obscurus*. Synthetic ToAP2 displayed antimicrobial activity, both in vitro and in vivo [66]. CliHDPND402 is similar to peptide TsAP2 (S6D3A7), an antibacterial, antifungal, anticancer, and hemolytic peptide isolated from *T. serrulatus* [61]. The precursor sequences of these NDBP-4 are shown aligned in Figure 5C.

A transcript for a putative anionic HDP without disulfide bridges is also reported. CliHDPAni01 shows similarities to the genome-derived sequence XP\_023227050 from *C. sculptratus*, and with TanP (GeneBank ID of the transcript: JK483720), a peptide isolated from the venom of *Tityus stigmurus* [67]. The predicted precursors translated from these sequences, are shown aligned in Figure 5D.

2.2.3. Enzymes

Enzymes are essential components of many animal venoms. Thought more abundant in snakes [68,69], enzymes have been discovered in venoms from many taxa, including ants [70],

jellyfish [71], wasps [72], spiders [73], and scorpions [74]. Scorpions, in particular, were shown to contain proteases [75], phospholipases [76], and hyaluronidases [77] in their venoms by classical studies. More recently, high-throughput transcriptomic analyses with venom glands have reported the existence of transcripts putatively coding for other enzymes never before reported in scorpion venoms, e.g., 5′nucleotidases in non-buthid scorpions [10,12,24].

The venom gland of *C. limpidus* was found in this study to be rich in terms of transcripts putatively coding for enzymes. Forty-nine different sequences potentially coding for enzymes are here reported. Thirty-eight corresponded to proteases, seven to phospholipases, three to 5′nucleotidases, and one to a putative hyaluronidase. All of them were annotated by sequence similarity to genome-derived sequences from *C. sculpturatus*. They are all reported in detail in Supplementary Materials Tables S1 and S2, together with the reference genomic sequences.

The large number of protease-encoding sequences in this scorpion is remarkable. It is particularly rich in transcripts for metalloproteases, with 24 out of the 38 transcripts potentially coding for proteases, being for metalloproteases (the remaining 14 are for serine proteases). It is also noticeable that the most diverse metalloprotease transcripts are those coding for antarease-type Zn-metalloproteases, 14 out of 24. Antarease was first described in the venom of *T. serrulatus* [78]. The envenomation by *Tityus* species can lead to the development of acute pancreatitis [79–83]. Venoms from *Tityus* scorpions are potent secretagogues that can elicit the release of secretory proteins from the pancreas [80,84–86]. Antarease was shown to specifically cleave the soluble N-ethylmaleimide-sensitive factor attachment protein receptors (SNAREs) involved in pancreatic secretion, disrupting the normal vesicular traffic [78]. Antarease could, therefore, be responsible for the acute pancreatitis induced by the *T. serrulatus* venom. Antarease-like enzymes were shown thereafter to be ubiquitous in the venoms of different scorpion genera [74]. The venom from *C. limpidus* was also shown to elicit manifestations associated with pancreatitis and to act as a secretagogue for amylase in the mouse pancreas [87]. The abundance of transcripts coding for antareases in *C. limpidus* could be an indicative of the prominent role of these enzymes in the manifestation of pancreatitis after scorpion envenomation by this species. The presence of antareases in the venom was here confirmed by the proteomic analysis. The proteins encoded by CliEnzMtp15, CliEnzMtp19, CliEnzMtp20, and CliEnzMtp21 were found by LC-MS/MS.

The other found transcripts, potentially coding for disintegrins, reprotolysins, an astacin, and an angiotensin-converting enzyme complete the picture of the metalloproteases putatively expressed in this venom (Supplementary Materials Table S2). The astacin (CliEnzMtp23) and the angiotensin-converting enzyme (CliEnzMtp24) were also confirmed to be present in the venom by LC-MS/MS.

It is intriguing that, even though transcripts potentially coding for phospholipases A2 and D2 were found in the transcriptomic analysis (Supplementary Materials Table S2), none were confirmed by the proteomic analysis. The inability to detect them in the proteome seems to correlate with the absence of phospholipase activity in the *C. limpidus* crude venom [26] as assayed by the egg-yolk method of Haberman and Hard, which detects phospholipases A, B, and C [88]. Nevertheless, it should be kept in mind that this transcriptomic analysis is descriptive and not quantitative. Therefore, no inferences can be derived from the transcript detection about the transcript and protein levels in the gland and the venom, respectively.

An interesting family of enzymes, just recently identified in scorpion venoms are the 5′nucleotidases. They were initially found in snake venoms where they seem to be ubiquitous [89]. They are known to endogenously liberate purines, which potentiate venom-induced hypotension and paralysis via purine receptors [89]. In addition, they can synergistically interact with other venom proteins, contributing to the overall effects of venoms [89]. The presence of 5′nucleotidases in non-buthid scorpion venoms was demonstrated for *Paravaejovis schwenkmeyeri*, *Thorellius atrox*, and *Megacormus gertschi* [10,12,24]. This study confirms that this enzyme could also be present in a buthid scorpion venom—at least its mRNA is transcribed in the gland—suggesting that these enzymes could also be ubiquitous in scorpion venoms. Three transcripts for *C. limpidus* 5′nucleotidases are reported in Supplementary Materials Table S2.

It was previously shown that the *C. limpidus* crude venom contains active hyaluronidases [26]. A single transcript encoding for a hyaluronidase, CliEnzHya01 (Supplementary Materials Table S2), was found in the present study. Noticeably, its presence in the venom was confirmed by the proteomic analysis. By degrading hyaluronic acid from the extracellular matrix [90], this enzyme could facilitate venom spreading [91].

#### 2.2.4. Other Venom Components

Besides the ones described above, other components frequently found in scorpion venoms were also detected in this analysis (Supplementary Materials Table S1).

Protease inhibitors, which, in principle, protect the venom components from autogenous degradation by venom proteases [69], though might also inhibit enzymes in the targets' tissues with neurological consequences [92], were very diverse in terms of transcripts in *C. limpidus*. Nineteen precursors encoding potential protease inhibitors were found, 13 of the Ascaris-type [93], two of the Kunitz-type [94], and four serpins [95].

Ascaris-type protease inhibitors are serine protease inhibitors, which have a Trypsin-Inhibitor-Like (TIL) domain. They are long-chain peptides with approximately 60 amino acid residues, stabilized by five disulfide bonds [96]. Recombinant SjAPI was the first functionally characterized Ascaris-type protease inhibitor from animal venoms [93]. It inhibits chymotrypsin and elastase while being inactive on trypsin. Three transcripts described here (CliPinTIL01–CliPinTIL03) have sequence similarity with classical Ascaris-type protease inhibitors from scorpion venoms or transcriptomes. Ten others, (CliPinTIL04–CliPinTIL13), with conserved TIL domain, but which are larger (ca 14 kDa instead of the classical ca 6 kDa) and have more disulfide bridges than the classical inhibitors (up to five extra disulfides), are also reported. A fourteenth sequence with a TIL domain (annotated here as CliPinTIL14) was detected in the proteomic analysis but had no matching transcript. It was annotated within this group due to the presence of the TIL domain, but it is a much larger protein (ca 210 kDa).

Kunitz-type protease inhibitors have been described in scorpion venoms [94]. The transcripts here reported (CliPinKun01 and CliPinKun02) were similar to genome sequences from *C. sculpturatus*. Members of the superfamily of serine proteinase inhibitors (serpins) from venomous animals have been poorly studied [95,97]. Four transcripts encoding serpins, CliPinSrp01–CliPinSrp04 are here reported, which also have sequence similarities with genome sequences from *C. sculpturatus*.

The CAP superfamily proteins are well distributed in all organisms, where they display a variety of functions [98]. This superfamily includes three major groups of proteins: cysteine-rich secretory proteins (CRISP), antigen or allergen proteins from arthropod venoms, and pathogenesis-related proteins from plants (PR) [99]. The CRISP proteins and antigen/allergen proteins have been described in venoms [100]. There are reports showing that CRISP proteins can interact with ion channels [101,102]. Six precursors coding for CAP-like proteins are here reported, CliOthCAP01–CliOthCAP06 (Supplementary Materials Table S1). They all code for putative proteins with a Pfam domain corresponding to the CAP superfamily and the transcripts are similar to genome sequences from *C. sculpturatus*. Both CliOthCAP02 and CliOthCAP05 were confirmed to be expressed in the venom by the proteomic analysis.

Transcripts for Insuline Growth Factor Binding Proteins (IGFBPs) have been found in scorpion venom glands [7,8,12,33,64,103,104], though the function of IGFBPs in venoms has not been established. Ten putative IGFBP transcripts are reported for *C. limpidus* (Supplementary Materials Table S1). All of them matched genomic sequences from *C. sculpturatus*, except for CliOthIGF10, which was similar to a peptide annotated as "Venom toxin" (A0A1L4BJ69) from *Hemiscorpius lepturus*.

The first described La1 peptide was found in the venom of *Liocheles australiase* [105]. Peptides with similar amino acid sequences were thereafter annotated as "La1-like" peptides. They are usually long-chain peptides containing 73–100 amino acids, stabilized by four disulfide linkages, with a conserved single domain von Willebrand factor-type C (SVWC) structural motive. Transcripts for La1-like peptides have been identified in buthid [8] and non-buthid scorpions [7,10,12,24,104]. Six transcripts are here reported with those features, assumed to code for La1-like peptides

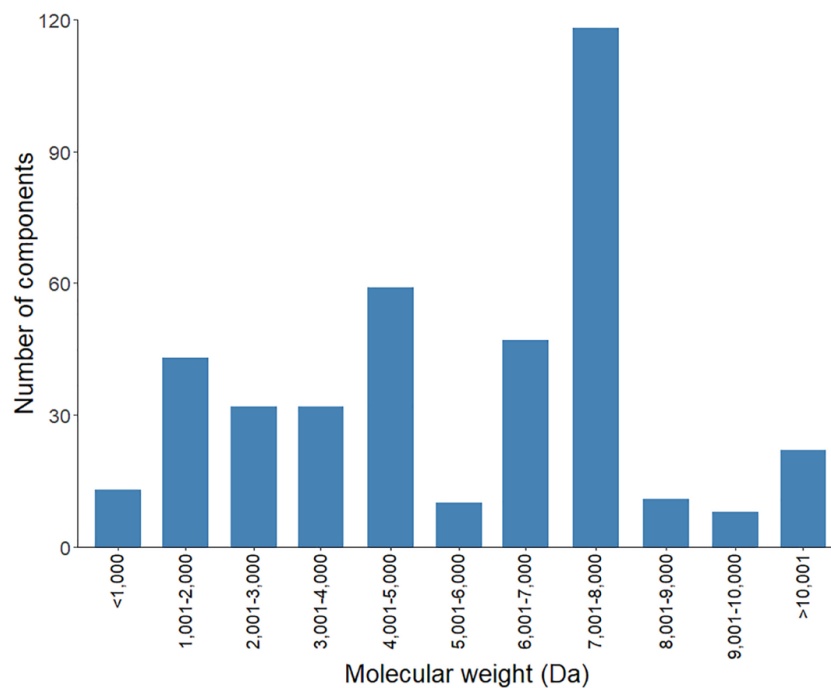
(Supplementary Materials Table S1). Of them, only CliOthLa106 was confirmed to be expressed as protein in the venom. Noticeably, CliOthLa106 is the only transcript among the six, which codes for a long-chain-type La1-like [24] (Supplementary Materials Table S1).

Other orphan transcripts with ORF coding for putative peptides with no conserved structural domains, nor a known function, were also found. Since similar sequences had been described in other scorpion transcriptomic analyses or were directly identified in the *C. limpidus* venom, they were grouped as “undefined peptides” and annotated as CliOthUnd01–CliOthUnd05. Not much information can be provided on them, except for the reference sequences they are similar with (Supplementary Materials Table S1).

### 2.3. Proteomic Exploration of the Venom Components of *C. limpidus*

#### 2.3.1. Mass Fingerprint of the Soluble Venom

The soluble fraction of the whole venom from sixty scorpions of mixed gender was used for the proteomic analysis. An aliquot of the soluble venom was applied to an HPLC coupled to a mass spectrometer, as previously reported [26]. The mass fingerprint detected 395 individual masses, ranging from 800 to 19,000 Da (Supplementary Materials Table S3 and Figure 6). The mass range with the largest number of individual masses detected was the one which spans 7001 to 8000 Da, which is within the expected range for the Na<sup>+</sup> channel-acting scorpion toxins. This mass group was followed, in the number of independent masses, by the 4001 to 5000 Da range, which are the characteristic masses of K<sup>+</sup> channel-acting scorpion toxins. This mass distribution corroborates the findings by the transcriptomic analysis pointing to the ion channel-acting toxins as the most diverse components in the venom of the highly toxic *C. limpidus*. It is relevant to note that, although transcripts potentially coding for enzymes were highly diverse in accordance to the transcriptomic analysis, the used setup for the mass fingerprint cannot precisely detect the mass of high molecular weight components (approximately above 10,000 Da) [24], so all those components are shown grouped in Figure 6.



**Figure 6.** Distribution of the *C. limpidus* venom components detected by mass fingerprint with respect to their molecular masses.

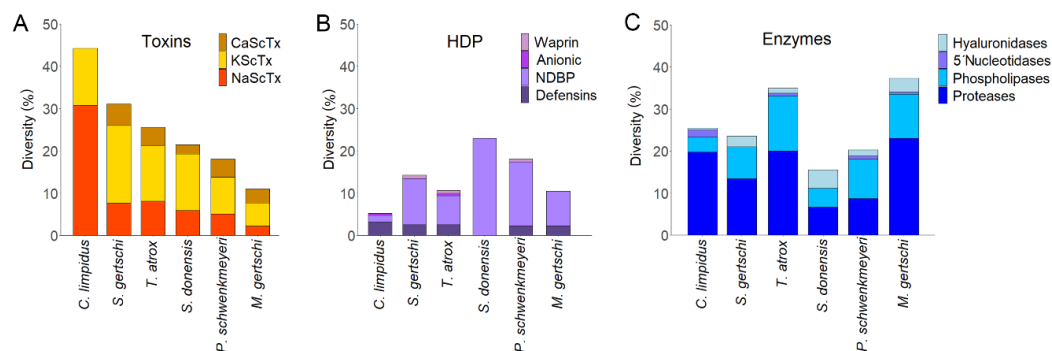
### 2.3.2. Identification of Peptides by LC-MS/MS

The peptides identified by LC-MS/MS, which also matched the sequences discovered by the transcriptomic analysis, were mentioned above in the description of the corresponding transcripts. This section summarizes those findings (Supplementary Materials Table S4).

From a total of 52 identified sequences, 46 corresponded to molecular entities annotated as venom-related peptides/proteins. The remaining sequences corresponded to enzymes or proteins related to cellular processes, or without an identifiable conserved domain or structural motif. In correlation with the results of the transcriptomic and mass fingerprint analyses, the largest class of MS/MS-recovered sequences corresponded to toxins that affect ion channels, of which 26 were Na<sup>+</sup> channel-acting toxins (7  $\alpha$ -type and 19  $\beta$ -type, by sequence similarity) and three K<sup>+</sup> channel-acting toxins. Tryptic peptides from eight enzymes were also recovered. Other venom components were also detected, but to a lesser extent. This includes two HDPs, one Ascaris-type protease inhibitor (though this is a protein much larger than the typical sequences of this kind, as discussed in the transcriptomic section), two proteins of the CAP superfamily, and one La1-like peptide of the long-chain type.

### 2.4. The Venom of the Highly Toxic *C. limpidus* versus the Venoms of Non-buthid Scorpions

Two other transcriptomic/proteomic analyses from scorpions of the genus *Centruroides* have already been published [6,8]. It would be interesting to compare these venoms in terms of their composition to try to define a possible pattern that would differentiate toxic from non-toxic species within this genus (*C. limpidus* and *C. noxius* are highly toxic to mammals while *C. hentzi* is not). However, even minor differences in sample preparation, cDNA library construction, cDNA sequencing protocol, MS/MS protocol or the bioinformatics analysis could lead to inaccurate conclusions. Nevertheless, the results here reported are comparable with those generated for other non-buthid scorpions under the exact same experimental protocol. That is the case for the analyses performed with *Serradigitus gertschi*, *Superstitionia donensis*, *T. atrox*, *P. schwenkmeyeri*, and *M. gertschi* [7,10,12,24,25]. As Figure 7A illustrates, the venom of *C. limpidus* is characterized by the highest diversity of toxins, in general. This is remarkable, considering also that no Ca<sup>2+</sup> channel-acting toxins were recovered for the buthid species, so only two super-families, the NaScTx and the KScTx, are present in the venom. The fraction of recovered transcripts related to Na<sup>+</sup> channel-acting toxins equals or surpasses the fraction of all toxins in the non-buthids, taken together. It is also relevant that the fraction of NaScTx in *C. limpidus* is larger than the fraction of KScTx, while for the non-buthids it is the other way around. It is well established that the toxicity of the scorpion venoms is primordially related to the neurotoxic action of the NaScTx [16,18]. It is well known that the neutralization of the main NaScTx in scorpion venoms results in the neutralization of the whole venom toxicity [106,107]. It is remarkable that, in the highly toxic *C. noxius*, the NaScTx are also more abundant and diverse than the KScTx [6,108], while in the non-lethal *C. hentzi*, the transcripts for NaScTx and KScTx are more or less equally diverse [8] (although, in absolute numbers, not comparable with the abundances here reported, as indicated above).



**Figure 7.** Comparative diversity of the transcripts coding for toxins (A), HDPs (B), and enzymes (C) found in *C. limpidus* versus those from non-buthid species reported in other transcriptomic analyses.

Figure 7B confirms that, on the contrary, the venoms from the non-toxic, non-buthid scorpions seem to be more diverse in terms of HDPs, with respect to the buthids, as proposed above in Section 2.2.2. Additionally, no direct link seems to exist between enzyme diversity and venom toxicity, according to the results charted in Figure 7C.

### 3. Conclusions

A total of 192 transcripts were identified in the present transcriptomic analysis. These sequences are assumed to code for Na<sup>+</sup> and K<sup>+</sup> channel-acting toxins, enzymes, HDPs, protease inhibitors, CAP-super-family proteins, IGFBP, La-1-like peptides, and other orphan venom components. Mass fingerprint of the venom resulted in the detection of 395 individual components, the most abundant of which were peptides with molecular weights in the range of 7000 to 8000 Da, which are known to correspond to Na<sup>+</sup> channel-acting toxins. The LC-MS/MS of the tryptically-digested venom confirmed that at least 46 of the venom-related, transcript-encoded proteins are indeed expressed in the venom.

The molecular dissection of the venom components from the highly toxic buthid scorpion *C. limpidus* revealed that the most abundant (from the mass fingerprint) and most diverse (from the transcriptomic and MS/MS analyses) venom components are the neurotoxic NaScTx. The fraction of toxins (most notably, the NaScTx) is significantly higher in *C. limpidus* than in other non-toxic, non-buthid scorpions from different genera. These findings correlate with classical biochemical and physiological observations on the relevance of the neurotoxic NaScTx in the toxicity of the scorpion venoms for mammals. It also reveals that the efforts directed at generating neutralizing antivenoms from toxin-specific human antibodies or their fragments might require a larger number of antibodies in the cocktail, depending on their effective cross-reactivity.

Molecular support for the observed pancreatitis after envenomation by this species is also provided. The relative abundance and diversity of the antarease-like Zn-metalloproteases seem to confirm their relevant role in this pathology.

Two families of toxins are described here for the first time in a scorpion of the genus *Centruroides*. Transcripts coding for  $\beta$ -KScTx were found, and the presence of the expressed toxins in the venom was confirmed by the proteomic analysis. Transcripts putatively coding for members of the  $\lambda$ -KScTx family are here also reported, though their presence in the venom remains to be demonstrated.

### 4. Materials and Methods

#### 4.1. Biological Material

Adult scorpions of the species *C. limpidus* were collected in Morelos and Guerrero States, with official permit from the Secretaría de Medio Ambiente, Recursos Naturales y Pesca (SEMARNAT, numbers SGPA/DGVS/07805/16 and 004474/18). The collected specimens were maintained in plastic boxes with water ad libitum. Sixty adult scorpions were milked by electrical stimulation for the proteomic analyses. The venom was immediately suspended in deionized water and centrifuged at 15,000 g for 15 min. The protein content of the soluble venom was estimated with a Nanodrop 1000 (Thermo Fisher Scientific, Waltham, MA, USA) based on absorbance at  $\lambda = 280$  nm, assuming that one absorbance unit equaled 1 mg/mL of protein. The soluble fraction of the whole venom was lyophilized and kept at  $-20$  °C until used.

Seven days after milking, 5 male and 5 female scorpions were processed for telson dissection to extract the total RNA. These specimens were thereafter euthanized and preserved in ethanol as vouchers. The remaining scorpions were kept alive and released thereafter in the same locations where they were collected.

#### 4.2. RNA Isolation, Sequencing, and Assembly

Total RNA was isolated as previously described [7,10,12]. The SV Total RNA Isolation System Kit (Promega, Madison, WI, USA) was used for this purpose. The telsons from 5 males and 5 females

were dissected separately, under RNase-free conditions, in microcentrifuge tubes containing the RNA lysis buffer. The optional 70 °C heating step of the protocol was followed before column purification. Total RNA was quantitated with a NanoDrop 1000 (Thermo Fisher Scientific, Waltham, MA, USA) and its quality was assessed with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

Two cDNA libraries were prepared from 1 µg of total RNA for each gender, using the TruSeq Stranded mRNA Sample Preparation Kit (Illumina, San Diego, CA, USA), according to the manufacturer's directions. DNA sequencing was performed at the Massive DNA Sequencing Facility of the Instituto de Biotecnología (Cuernavaca, Mexico) with a Genome Analyzer IIx (Illumina, San Diego, CA, USA) using a 72-bp paired-end sequencing scheme over cDNA fragments ranging 200–400 bp in size. The quality of the reads was verified with the FastQC program [109] after clipping off the adaptors.

The reads resulting from the sequencing of both male and female cDNA libraries were joined together for a de novo assembly into contigs using the Trinity software version 2.0.3 [110], employing the standard protocol. Basic statistics, such as the number of transcripts and contigs, were determined with the TrinityStats.pl script.

#### 4.3. Bioinformatics

The assembled contigs were annotated as previously described [24], using the Uniprot/Uniref90 protein database for BLASTx and BLASTp. The prediction of ORFs were done with TransDecoderLongORFs. Putative signal peptides and propeptides were predicted with the ProP 1.0 server [111] and SpiderP from Arachnoserver [112]. Multiple sequence alignments were performed using MAFFT 7.0 online [113]. Alignments were edited in Bioedit [114] and in Adobe Illustrator CS6. All figures were generated with Rstudio [115]. The GO terms were quantified in WEGO [116].

#### 4.4. Molecular Mass Fingerprint by LC-MS of the Venom

Eight micrograms of soluble venom was applied in an LC-MS system composed of an HPLC UltiMate 3000; Dionex, RSLCnano System (Thermo Fisher Scientific, San Jose, CA, USA) coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA). Venom was fractionated through a 10-cm reversed-phase C18 in-house-made column (filled with Jupiter® 4 µm Proteo 90 Å resin, Phenomenex, Torrance, CA, USA), using a linear gradient of 5% to 90% of solvent B (0.1% formic acid in acetonitrile) in 240 min, with a flow rate of 300 nL/min. The resolved peptides were ionized by a nano-electrospray ion source. A full scan MS was used (400–2000 m/z) with a resolution of 60,000. The monoisotopic molecular mass is reported for components below 3000 Da, and the average molecular masses for those above 3000 Da.

#### 4.5. Identification of Venom Components by LC-MS/MS

Three hundred micrograms of the venom soluble fraction were reduced with dithiothreitol (DTT) 10 mM (Sigma–Aldrich, Saint Louis, MO USA) and alkylated with iodoacetamide (IAA) 55 mM (Sigma–Aldrich, Saint Louis, MO USA). After that, the samples were digested with trypsin (Promega Sequencing Grade Modified Trypsin; Madison, WI, USA) using a 1:50 enzyme:protein ratio by weight, in 50 mM ammonium bicarbonate buffer (ABC). The samples were acidified with 10 µL of 10% formic acid (FA). The tryptic peptides were desalted with Sep-Pak tC18 cartridges (Waters, Milford, MA, USA) following the manufacturer's protocol, and dried in a SpeedVac (Savant SPD1010, Thermo Scientific, San Jose, CA, USA).

Samples containing 4 µg of tryptically digested venom in 10 µL of solution A (0.1% formic acid) were analyzed. The proteins were separated through a 15-cm in-house-made column (filled with the same C18, Jupiter 4 µm Proteo 90 Å resin) using a linear gradient of 5% to 75% of solvent B (0.1% formic acid in acetonitrile) in 270 min. Mass spectra were registered in a full scan of 350 m/z to 1400 m/z with a resolution of 60,000. The MS/MS spectra were analyzed with the Proteome Discoverer 1.4.1.14 suite (Thermo Fisher Scientific, San Jose, CA, USA), employing the Sequest HT search engine with the following parameters: two missed cleavages, dynamic modifications (methionine oxidation, glutamine

and asparagine deamidation), static modifications (cysteine carbamidomethylation), precursor mass tolerance of 20 ppm, fragment mass tolerance 0.6 Da, and 1% false discovery rate. The database of the translated transcripts from Transdecoder was used for protein identification. An identification was considered positive when a minimum of two matching peptides were identified, and Sequest HT gave a global score (sum of all peptides' XCorr) of at least 20.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2072-6651/11/5/247/s1>, Supplementary Figure S1: Distribution of the annotated transcripts in accordance to Gene Ontology (GO) terms, Supplementary Figure S2: Phylogenetic tree of scorpion toxins with the ICK fold. Supplementary Table S1: Translated sequences of the 192 venom-related transcripts obtained from the transcriptomic analysis of the *C. limpidus* venom gland, Supplementary Table S2: The 192 venom-related transcripts obtained from the *C. limpidus* venom gland and their reference proteins, Supplementary Table S3: Mass fingerprint results. The 395 masses identified in the *C. limpidus* venom by LC-MS, sorted by C18-RT-HPLC Retention Times (RT). Supplementary Table S4: LC-MS/MS results. The 52 transcripts that were identified in venom by LC-MS/MS. The parameters generated by the Proteome Discoverer software are shown for each identified fragment.

**Author Contributions:** Conceptualization, J.I.C.-U., E.O. and L.D.P.; methodology, J.I.C.-U., E.P.M. and E.O.; software, J.I.C.-U.; validation, J.I.C.-U., E.P.M., C.V.F.B., E.O. and L.D.P.; formal analysis, J.I.C.-U. and E.P.M.; investigation, J.I.C.-U.; resources, J.I.C.-U. and L.D.P.; data curation, J.I.C.-U. and E.P.M.; writing—original draft preparation, J.I.C.-U. and E.P.M.; writing—review and editing, C.V.F.B., E.O. and L.D.P.; visualization, J.I.C.-U.; supervision, E.O. and L.D.P.; project administration, L.D.P.; funding acquisition, L.D.P.

**Funding:** This research was funded by grants IN203416 and IN202619, from Dirección General de Personal Académico, UNAM and grant SEP-CONACyT 237864 from Consejo Nacional de Ciencia y Tecnología (CONACyT) awarded to Lourival D. Possani and Jimena I. Cid-Urbe was supported by scholarship No. 404460 from CONACyT.

**Acknowledgments:** Specimens were classified by Edmundo González-Santillán based on the available literature [14]. We are indebted to Carlos Santibáñez-López from the University of Wisconsin-Madison, for performing the phylogenetic analysis that allowed the classification of the  $\lambda$ -KScTx. We are grateful to Gloria T. Vázquez-Castro and Ricardo A. Grande-Cano from the Massive Sequencing Unit of the Instituto de Biotecnología-UNAM for their technical support. The computer analysis was performed using the cluster of the Instituto de Biotecnología-UNAM maintained by Jérôme Verleyen. The technical assistance of David S. Castañeda-Carreón, Roberto P. Rodríguez-Bahena, Jesús O. Arriaga-Pérez, Juan M. Hurtado-Ramírez, and Servando Aguirre-Cruz is also greatly acknowledged.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Possani, L.D.; Becerril, B.; Delepierre, M.; Tytgat, J. Scorpion toxins specific for Na<sup>+</sup>-channels. *Eur. J. Biochem.* **1999**, *264*, 287–300. [[CrossRef](#)]
2. The Scorpion Files. Available online: <https://www.ntnu.no/ub/scorpion-files/> (accessed on 29 April 2019).
3. Chippaux, J.-P.; Goyffon, M. Epidemiology of scorpionism: A global appraisal. *Acta Trop.* **2008**, *107*, 71–79. [[CrossRef](#)] [[PubMed](#)]
4. Santibáñez-López, C.E.; Francke, O.F.; Ureta, C.; Possani, L.D.; Lai, R. Scorpions from Mexico: From species diversity to venom complexity. *Toxins* **2016**, *8*, 2. [[CrossRef](#)]
5. González-Santillán, E.; Possani, L.D. North American scorpion species of public health importance with a reappraisal of historical epidemiology. *Acta Trop.* **2018**, *187*, 264–274. [[CrossRef](#)]
6. Rendón-Anaya, M.; Delaye, L.; Possani, L.D.; Herrera-Estrella, A. Global transcriptome analysis of the scorpion *Centruroides noxius*: New toxin families and evolutionary insights from an ancestral scorpion species. *PLoS ONE* **2012**, *7*, e43331. [[CrossRef](#)]
7. Santibáñez-López, C.E.; Cid-Urbe, J.I.; Batista, C.V.F.; Ortiz, E.; Possani, L.D. Venom gland transcriptomic and proteomic analyses of the enigmatic scorpion *Superstitionia donensis* (Scorpiones: *Superstitioniidae*), with insights on the evolution of its venom components. *Toxins* **2016**, *8*, 367. [[CrossRef](#)]
8. Ward, M.J.; Ellsworth, S.A.; Rokyta, D.R. Venom-gland transcriptomics and venom proteomics of the Hentz striped scorpion (*Centruroides hentzi*; *Buthidae*) reveal high toxin diversity in a harmless member of a lethal family. *Toxicon* **2018**, *142*, 14–29. [[CrossRef](#)] [[PubMed](#)]
9. Quintero-Hernández, V.; Ramírez-Carretero, S.; Romero-Gutiérrez, M.T.; Valdez-Velázquez, L.L.; Becerril, B.; Possani, L.D.; Ortiz, E. Transcriptome analysis of scorpion species belonging to the *Vaejovis* Genus. *PLoS ONE* **2015**, *10*, e0117188. [[CrossRef](#)] [[PubMed](#)]



10. Romero-Gutierrez, T.; Peguero-Sanchez, E.; Cevallos, M.A.; Batista, C.V.F.; Ortiz, E.; Possani, L.D. A deeper examination of *Thorellius atrox* Scorpion venom components with omic technologies. *Toxins* **2017**, *9*, 399. [[CrossRef](#)]
11. Zhong, J.; Zeng, X.-C.; Zeng, X.; Nie, Y.; Zhang, L.; Wu, S.; Bao, A. Transcriptomic analysis of the venom glands from the scorpion *Hadogenes troglodytes* revealed unique and extremely high diversity of the venom peptides. *J. Proteom.* **2017**, *150*, 40–62. [[CrossRef](#)]
12. Santibáñez-López, C.E.; Cid-Uribe, J.I.; Zamudio, F.Z.; Batista, C.V.F.; Ortiz, E.; Possani, L.D. Venom gland transcriptomic and venom proteomic analyses of the scorpion *Megacormus gertschi* Díaz-Najera, 1966 (Scorpiones: Euscorpidae: Megacorminae). *Toxicon* **2017**, *133*, 95–109. [[CrossRef](#)] [[PubMed](#)]
13. Luna-Ramírez, K.; Quintero-Hernández, V.; Juárez-González, V.R.; Possani, L.D. Whole transcriptome of the venom gland from *Urodacus yaschenkoi* Scorpion. *PLoS ONE* **2015**, *10*, e0127883. [[CrossRef](#)]
14. Ponce-Saavedra, J.; Francke, O.F.; Cano-Camacho, H.; Hernández-Calderón, E. Morphological and molecular evidence supporting specific status for *Centruroides tecomanus* (Scorpiones, Buthidae). *Rev. Mex. Biodivers.* **2009**, *80*, 71–84.
15. Chavez-Haro, A.L.; Ortiz, E. Scorpionism and Dangerous Species of Mexico. In *Scorpion Venoms*; Springer: Dordrecht, The Netherlands, 2014; pp. 201–213.
16. Riaño-Umbarila, L.; Rodríguez-Rodríguez, E.R.; Santibáñez-López, C.E.; Güereca, L.; Uribe-Romero, S.J.; Gómez-Ramírez, I.V.; Cárcamo-Noriega, E.N.; Possani, L.D.; Becerril, B. Updating knowledge on new medically important scorpion species in Mexico. *Toxicon* **2017**, *138*, 130–137. [[CrossRef](#)]
17. Available online: <https://www.gob.mx/cms/uploads/attachment/file/425972/sem52.pdf> (accessed on 29 April 2019).
18. Riaño-Umbarila, L.; Olamendi-Portugal, T.; Morelos-Juárez, C.; Gurrola, G.B.; Possani, L.D.; Becerril, B. A novel human recombinant antibody fragment capable of neutralizing Mexican scorpion toxins. *Toxicon* **2013**, *76*, 370–376. [[CrossRef](#)]
19. Dehesa-Dávila, M.; Ramfrez, A.N.; Zamudio, F.Z.; Gurrola-Briones, G.; Liévano, A.; Darszon, A.; Possani, L.D. Structural and functional comparison of toxins from the venom of the scorpions *Centruroides infamatus infamatus*, *Centruroides limpidus limpidus* and *Centruroides noxius*. *Comp. Biochem. Physiol. Part B: Biochem. Mol. Boil.* **1996**, *113*, 331–339. [[CrossRef](#)]
20. Ramírez, A.N.; Martín, B.M.; Gurrola, G.B.; Possani, L.D. Isolation and characterization of a novel toxin from the venom of the scorpion *Centruroides limpidus limpidus* Karsch. *Toxicon* **1994**, *32*, 479–490. [[CrossRef](#)]
21. Alagon, A.; Guzmán, H.; Martín, B.; Ramírez, A.; Carbone, E.; Possani, L. Isolation and characterization of two toxins from the mexican scorpion *Centruroides limpidus limpidus* Karsch. *Comp. Biochem. Physiol. Part B: Comp. Biochem.* **1988**, *89*, 153–161. [[CrossRef](#)]
22. Lebreton, F.; Delepierre, M.; Ramírez, A.N.; Balderas, C.; Possani, L.D. Primary and NMR three-dimensional structure determination of a novel crustacean toxin from the venom of the scorpion *Centruroides limpidus limpidus* Karsch. *Biochemistry* **1994**, *33*, 11135–11149. [[CrossRef](#)]
23. Olamendi-Portugal, T.; Restano-Cassulini, R.; Riaño-Umbarila, L.; Becerril, B.; Possani, L.D. Functional and immuno-reactive characterization of a previously undescribed peptide from the venom of the scorpion *Centruroides limpidus*. *Peptides* **2017**, *87*, 34–40. [[CrossRef](#)]
24. Cid-Uribe, J.I.; Santibáñez-López, C.E.; Meneses, E.P.; Batista, C.V.; Jiménez-Vargas, J.M.; Ortiz, E.; Possani, L.D. The diversity of venom components of the scorpion species *Paravaejovis schwenkmeyeri* (Scorpiones: Vaejovidae) revealed by transcriptome and proteome analyses. *Toxicon* **2018**, *151*, 47–62. [[CrossRef](#)] [[PubMed](#)]
25. Romero-Gutiérrez, M.T.; Santibáñez-López, C.E.; Jiménez-Vargas, J.M.; Batista, C.V.F.; Ortiz, E.; Possani, L.D. Transcriptomic and proteomic analyses reveal the diversity of venom components from the vaejovid scorpion *Serradigitus gertschi*. *Toxins* **2018**, *10*, 359. [[CrossRef](#)] [[PubMed](#)]
26. Cid-Uribe, J.I.; Jiménez-Vargas, J.M.; Batista, C.V.F.; Zuñiga, F.Z.; Possani, L.D. Comparative proteomic analysis of female and male venoms from the Mexican scorpion *Centruroides limpidus*: Novel components found. *Toxicon* **2017**, *125*, 91–98. [[CrossRef](#)]
27. Pfam. Available online: <https://pfam.xfam.org/> (accessed on 29 April 2019).
28. Venom Zone. Available online: <https://venomzone.expasy.org/> (accessed on 29 April 2019).
29. Gene Ontology. Available online: <http://geneontology.org/> (accessed on 29 April 2019).
30. Quintero-Hernandez, V.; Jimenez-Vargas, J.; Gurrola, G.; Valdivia, H.; Possani, L. Scorpion venom components that affect ion-channels function. *Toxicon* **2013**, *76*, 328–342. [[CrossRef](#)] [[PubMed](#)]

31. Gurevitz, M.; Froy, O.; Zilberberg, N.; Turkov, M.; Strugatsky, D.; Gershburg, E.; Lee, D.; Adams, M.E.; Tugarinov, V.; Anglister, J.; et al. Sodium channel modifiers from scorpion venom: Structure–activity relationship, mode of action and application. *Toxicon* **1998**, *36*, 1671–1682. [[CrossRef](#)]
32. Escalona, M.P.; Possani, L.D. Scorpion beta-toxins and voltage-gated sodium channels: Interactions and effects. *Front Biosci.* **2013**, *18*, 572–587. [[CrossRef](#)]
33. De Oliveira, U.C.; Candido, D.M.; Dorce, V.A.C.; Junqueira-De-Azevedo, I.D.L.M. The transcriptome recipe for the venom cocktail of *Tityus bahiensis* scorpion. *Toxicon* **2015**, *95*, 52–61. [[CrossRef](#)] [[PubMed](#)]
34. Jablonsky, M.J.; Jackson, P.L.; Krishna, N.R. Solution structure of an insect-specific neurotoxin from the new world scorpion *Centruroides sculpturatus* Ewing<sup>†,‡</sup>. *Biochemistry* **2001**, *40*, 8273–8282. [[CrossRef](#)] [[PubMed](#)]
35. Jablonsky, M.J.; Watt, D.D.; Krishna, N.R. Solution structure of an old world-like neurotoxin from the venom of the new world scorpion *Centruroides sculpturatus* Ewing. *J. Mol. Boil.* **1995**, *248*, 449–458. [[CrossRef](#)]
36. Kuang, Q.; Purhonen, P.; Hebert, H. Structure of potassium channels. *Cell. Mol. Life Sci.* **2015**, *72*, 3677–3693. [[CrossRef](#)] [[PubMed](#)]
37. Jiménez-Vargas, J.M.; Possani, L.D.; Luna-Ramírez, K. Arthropod toxins acting on neuronal potassium channels. *Neuropharmacology* **2017**, *127*, 139–160. [[CrossRef](#)] [[PubMed](#)]
38. Kim, D.M.; Nimigean, C.M. Voltage-gated potassium channels: A structural examination of selectivity and gating. *Cold Spring Harbor Perspect. Boil.* **2016**, *8*, a029231. [[CrossRef](#)]
39. Tytgat, J.; Chandy, K.; Garcia, M.L.; A Gutman, G.; Martin-Eauclaire, M.-F.; Van Der Walt, J.J.; Possani, L.D. A unified nomenclature for short-chain peptides isolated from scorpion venoms:  $\alpha$ -KTx molecular subfamilies. *Trends Pharmacol. Sci.* **1999**, *20*, 444–447. [[CrossRef](#)]
40. Rodríguez De La Vega, R.C.; Possani, L.D. Current views on scorpion toxins specific for K<sup>+</sup>-channels. *Toxicon* **2004**, *43*, 865–875. [[CrossRef](#)]
41. Cremones, C.M.; Maiti, M.; Peigneur, S.; Cassoli, J.S.; Dutra, A.A.A.; Waelkens, E.; Lescrier, E.; Herdewijn, P.; de Lima, M.E.; Pimenta, A.M.C.; et al. Structural and functional elucidation of peptide TS11 shows evidence of a novel subfamily of scorpion venom toxins. *Toxins* **2016**, *8*, 288. [[CrossRef](#)] [[PubMed](#)]
42. Correnti, C.E.; Gewe, M.M.; Mehlin, C.; Bandaranayake, A.D.; Johnsen, W.A.; Rupert, P.B.; Brusniak, M.-Y.; Clarke, M.; Burke, S.E.; De Van Der Schueren, W.; et al. Screening, large-scale production, and structure-based classification for cystine-dense peptides. *Nat. Struct. Mol. Boil.* **2018**, *25*, 270–278. [[CrossRef](#)] [[PubMed](#)]
43. Nieto, A.R.; Gurrola, G.B.; Vaca, L.; Possani, L.D. Noxiustoxin 2, a novel K<sup>+</sup> channel blocking peptide from the venom of the scorpion *Centruroides noxius* Hoffmann. *Toxicon* **1996**, *34*, 913–922. [[CrossRef](#)]
44. Valdivia, H.H.; Smith, J.S.; Martin, B.M.; Coronado, R.; Possani, L.D. Charybdotoxin and noxiustoxin, two homologous peptide inhibitors of the K<sup>+</sup>(Ca<sup>2+</sup>) channel. *FEBS Lett.* **1988**, *226*, 280–284. [[CrossRef](#)]
45. Kuzmenkov, A.I.; Krylov, N.A.; Chugunov, A.O.; Grishin, E.V.; Vassilevski, A.A. Kalium: A database of potassium channel toxins from scorpion venom. *Database* **2016**, *2016*. [[CrossRef](#)]
46. Restano-Cassulini, R.; Olamendi-Portugal, T.; Zamudio, F.; Becerril, B.; Possani, L.D. Two novel ergtoxins, blockers of K<sup>+</sup>-channels, purified from the Mexican scorpion *Centruroides elegans elegans*. *Neurochem Res.* **2008**, *33*, 1525–1533. [[CrossRef](#)]
47. Hill, A.P.; Sunde, M.; Campbell, T.J.; Vandenberg, J.I. Mechanism of block of the hERG K<sup>+</sup> channel by the scorpion toxin CnErg1. *Biophys. J.* **2007**, *92*, 3915–3929. [[CrossRef](#)] [[PubMed](#)]
48. Torres, A.M.; Bansal, P.; Alewood, P.F.; A Bursill, J.; Kuchel, P.W.; I Vandenberg, J.; Vandenberg, J. Solution structure of CnErg1 (Ergtoxin), a HERG specific scorpion toxin. *FEBS Lett.* **2003**, *539*, 138–142. [[CrossRef](#)]
49. Jimenez-Vargas, J.; Restano-Cassulini, R.; Quintero-Hernandez, V.; Gurrola, G.; Possani, L. Recombinant expression of the toxic peptide ErgTx1 and role of Met35 on its stability and function. *Peptides* **2011**, *32*, 560–567. [[CrossRef](#)] [[PubMed](#)]
50. Chen, Z.-Y.; Hu, Y.-T.; Yang, W.-S.; He, Y.-W.; Feng, J.; Wang, B.; Zhao, R.-M.; Ding, J.-P.; Cao, Z.-J.; Li, W.-X.; et al. Hg1, Novel peptide inhibitor specific for Kv1.3 channels from first scorpion kunitz-type potassium channel toxin family\*. *J. Boil. Chem.* **2012**, *287*, 13813–13821. [[CrossRef](#)]
51. Santibáñez-López, C.E.; Kriebel, R.; Ballesteros, J.A.; Rush, N.; Witter, Z.; Williams, J.; Janies, D.A.; Sharma, P.P.; Crandall, K. Integration of phylogenomics and molecular modeling reveals lineage-specific diversification of toxins in scorpions. *PeerJ* **2018**, *6*, e5902. [[CrossRef](#)]
52. Chen, Z.; Hu, Y.; Han, S.; Yin, S.; He, Y.; Wu, Y.; Cao, Z.; Li, W. ImKTx1, a new Kv1.3 channel blocker with a unique primary structure. *J. Biochem. Mol. Toxicol.* **2011**, *25*, 244–251. [[CrossRef](#)]

53. Gao, B.; Harvey, P.J.; Craik, D.J.; Ronjat, M.; De Waard, M.; Zhu, S. Functional evolution of scorpion venom peptides with an inhibitor cystine knot fold. *Biosci. Rep.* **2013**, *33*, 513–527. [[CrossRef](#)] [[PubMed](#)]
54. Wang, X.; Wang, G. Insights into Antimicrobial Peptides from Spiders and Scorpions. *Protein Pept. Lett.* **2016**, *23*, 707–721. [[CrossRef](#)]
55. Yan, L.; Adams, M.E. Lycotoxins, antimicrobial peptides from venom of the wolf spider *Lycosa carolinensis*. *J. Boil. Chem.* **1998**, *273*, 2059–2066. [[CrossRef](#)]
56. Torres-Larios, A.; Gurrola, G.B.; Zamudio, F.Z.; Possani, L.D.; Torres-Larios, A. Hadrurin, a new antimicrobial peptide from the venom of the scorpion *Hadrurus aztecus*. *JBIC J. Boil. Inorg. Chem.* **2000**, *267*, 5023–5031. [[CrossRef](#)]
57. Lorenzini, D.M.; Da Silva, P.I.; Fogaça, A.C.; Bulet, P.; Daffre, S. Acanthoscurrin: A novel glycine-rich antimicrobial peptide constitutively expressed in the hemocytes of the spider *Acanthoscurria gomesiana*. *Dev. Comp. Immunol.* **2003**, *27*, 781–791. [[CrossRef](#)]
58. Luna-Ramirez, K.; Skaljic, M.; Grotmann, J.; Kirfel, P.; Vilcinskas, A.; Possani, L.D. Orally delivered scorpion antimicrobial peptides exhibit activity against pea aphid (*Acyrtosiphon pisum*) and Its Bacterial Symbionts. *Toxins* **2017**, *9*, 261. [[CrossRef](#)]
59. Ferreira, L.; Alves, W.; Lucas, M.; Habermehl, G. Isolation and characterization of a bradykinin potentiating peptide (BPP-S) isolated from *Scaptocosa raptorica* venom. *Toxicon* **1996**, *34*, 599–603. [[CrossRef](#)]
60. Liu, Z.; Deng, M.; Xiang, J.; Ma, H.; Hu, W.; Zhao, Y.; Li, D.-C.; Liang, S. A novel spider peptide toxin suppresses tumor growth through dual signaling pathways. *Mol. Med.* **2012**, *12*, 1350–1360. [[CrossRef](#)]
61. Guo, X.; Ma, C.; Du, Q.; Wei, R.; Wang, L.; Zhou, M.; Chen, T.; Shaw, C. Two peptides, TsAP-1 and TsAP-2, from the venom of the Brazilian yellow scorpion, *Tityus serrulatus*: Evaluation of their antimicrobial and anticancer activities. *Biochimie* **2013**, *95*, 1784–1794. [[CrossRef](#)]
62. Zeng, X.-C.; Zhou, L.; Shi, W.; Luo, X.; Zhang, L.; Nie, Y.; Wang, J.; Wu, S.; Cao, B.; Cao, H. Three new antimicrobial peptides from the scorpion *Pandinus imperator*. *Peptides* **2013**, *45*, 28–34. [[CrossRef](#)]
63. De La Vega, R.R.; García, B.I.; D'Ambrosio, C.; Diego-García, E.; Scaloni, A.; Possani, L.D. Antimicrobial peptide induction in the haemolymph of the Mexican scorpion *Centruroides limpidus limpidus* in response to septic injury. *Cell. Mol. Life Sci.* **2004**, *61*, 1507–1519. [[CrossRef](#)] [[PubMed](#)]
64. Zhang, L.; Shi, W.; Zeng, X.-C.; Ge, F.; Yang, M.; Nie, Y.; Bao, A.; Wu, S.; E, G. Unique diversity of the venom peptides from the scorpion *Androctonus bicolor* revealed by transcriptomic and proteomic analysis. *J. Proteom.* **2015**, *128*, 231–250. [[CrossRef](#)] [[PubMed](#)]
65. Zeng, X.-C.; Wang, S.; Nie, Y.; Zhang, L.; Luo, X. Characterization of BmKbpp, a multifunctional peptide from the Chinese scorpion *Mesobuthus martensii* Karsch: Gaining insight into a new mechanism for the functional diversification of scorpion venom peptides. *Peptides* **2012**, *33*, 44–51. [[CrossRef](#)]
66. Marques-Neto, L.M.; Trentini, M.M.; Das Neves, R.C.; Resende, D.P.; Procopio, V.O.; Da Costa, A.C.; Kipnis, A.; Mortari, M.R.; Schwartz, E.F.; Junqueira-Kipnis, A.P. Antimicrobial and chemotactic activity of scorpion-derived peptide, ToAP2, against *Mycobacterium massiliensis*. *Toxins* **2018**, *10*, 219. [[CrossRef](#)]
67. Melo, M.M.; Daniele-Silva, A.; Teixeira, D.G.; Estrela, A.B.; Melo, K.R.; Oliveira, V.S.; Rocha, H.A.; Ferreira, L.D.S.; Pontes, D.L.; Lima, J.P.; et al. Structure and in vitro activities of a copper II-chelating anionic peptide from the venom of the scorpion *Tityus stigmurus*. *Peptides* **2017**, *94*, 91–98. [[CrossRef](#)]
68. Ramos, O.; Selistre-De-Araújo, H. Snake venom metalloproteases—structure and function of catalytic and disintegrin domains. *Comp. Biochem. Physiol. Part C: Toxicol. Pharmacol.* **2006**, *142*, 328–346. [[CrossRef](#)]
69. Lu, J.; Yang, H.; Yu, H.; Gao, W.; Lai, R.; Liu, J.; Liang, X. A novel serine protease inhibitor from *Bungarus fasciatus* venom. *Peptides* **2008**, *29*, 369–374. [[CrossRef](#)]
70. Touchard, A.; Aili, S.R.; Fox, E.G.P.; Escoubas, P.; Orivel, J.; Nicholson, G.M.; Dejean, A.; King, G.F. The biochemical toxin arsenal from ant venoms. *Toxins* **2016**, *8*, 30. [[CrossRef](#)] [[PubMed](#)]
71. Toom, P.M.; Chan, D.S. Enzymatic activities of venom from the jellyfish *Stomolophus meleagris*. *Comp. Biochem. Physiol. Part B: Comp. Biochem.* **1972**, *43*, 435–441. [[CrossRef](#)]
72. De Graaf, D.C.; Aerts, M.; Brunain, M.; Desjardins, C.A.; Jacobs, F.J.; Werren, J.H.; Devreese, B. Insights into the venom composition of the ectoparasitoid wasp *Nasonia vitripennis* from bioinformatic and proteomic studies. *Insect Mol. Boil.* **2010**, *19*, 11–26.
73. Trevisan-Silva, D.; Bednaski, A.V.; Gremski, L.H.; Chaim, O.M.; Veiga, S.S.; Senff-Ribeiro, A. Differential metalloprotease content and activity of three *Loxosceles* spider venoms revealed using two-dimensional electrophoresis approaches. *Toxicon* **2013**, *76*, 11–22. [[CrossRef](#)]

74. Ortiz, E.; Rendón-Anaya, M.; Rego, S.C.; Schwartz, E.F.; Possani, L.D. Antarease-like Zn-metalloproteases are ubiquitous in the venom of different scorpion genera. *Biochim. Biophys. Acta (BBA)—General Subj.* **2014**, *1840*, 1738–1746. [[CrossRef](#)]
75. Almeida, F.; Pimenta, A.; De Figueiredo, S.; Santoro, M.; Martin-Eauclaire, M.; Diniz, C.; De Lima, M.; Pimenta, A.M.D.C. Enzymes with gelatinolytic activity can be found in *Tityus bahiensis* and *Tityus serrulatus* venoms. *Toxicon* **2002**, *40*, 1041–1045. [[CrossRef](#)]
76. Conde, R.; Zamudio, F.Z.; Becerril, B.; Possani, L.D. Phospholipin, a novel heterodimeric phospholipase A2 from *Pandinus imperator* scorpion venom. *FEBS Lett.* **1999**, *460*, 447–450. [[CrossRef](#)]
77. Morey, S.S.; Kiran, K.; Gadag, J. Purification and properties of hyaluronidase from *Palamneus gravimanus* (Indian black scorpion) venom. *Toxicon* **2006**, *47*, 188–195. [[CrossRef](#)] [[PubMed](#)]
78. Fletcher, P.L.; Fletcher, M.D.; Weninger, K.; Anderson, T.E.; Martin, B.M. Vesicle-associated membrane protein (VAMP) cleavage by a new metalloprotease from the Brazilian scorpion *Tityus serrulatus*. *J. Biol. Chem.* **2010**, *285*, 7405–7416. [[CrossRef](#)] [[PubMed](#)]
79. Otero, R.; Navío, E.; Céspedes, F.; Núñez, M.; Lozano, L.; Moscoso, E.; Matallana, C.; Arsuza, N.; García, J.; Fernandez, D.; et al. Scorpion envenoming in two regions of Colombia: Clinical, epidemiological and therapeutic aspects. *Trans. Soc. Trop. Med. Hyg.* **2004**, *98*, 742–750. [[CrossRef](#)] [[PubMed](#)]
80. Fletcher, P.L., Jr.; Fletcher, M.D.; Possani, L.D. Characteristics of pancreatic exocrine secretion produced by venom from the brazilian scorpion, *Tityus serrulatus*. *Eur. J. Cell Biol.* **1992**, *58*, 259–270. [[PubMed](#)]
81. D'Suze, G.; Sevcik, C.; Ramos, M. Presence of curarizing polypeptides and a pancreatitis-inducing fraction without muscarinic effects in the venom of the Venezuelan scorpion *Tityus discrepans* (Karsch). *Toxicon* **1995**, *33*, 295. [[CrossRef](#)]
82. Borges, A.; Trejo, E.; Vargas, A.M.; Céspedes, G.; Hernández, A.; Alfonzo, M.J. Pancreatic toxicity in mice elicited by *Tityus zulianus* and *Tityus discrepans* scorpion venoms. *Investig. Clin.* **2004**, *45*, 269–276.
83. Bartholomew, C. Acute Scorpion Pancreatitis in Trinidad. *BMJ* **1970**, *1*, 666–668. [[CrossRef](#)]
84. Possani, L.D.; Martin, B.M.; Fletcher, M.D.; Fletcher, P.L. Discharge effect on pancreatic exocrine secretion produced by toxins purified from *Tityus serrulatus* scorpion venom. *J. Boil. Chem.* **1991**, *266*, 3178–3185.
85. Fletcher, M.D.; Possani, L.D.; Fletcher, P.L., Jr. Morphological studies by light and electron microscopy of pancreatic acinar cells under the effect of *Tityus serrulatus* venom. *Cell Tissue Res.* **1994**, *278*, 255–264. [[CrossRef](#)]
86. Fletcher, P.L.; Fletcher, M.D.; Fainter, L.K.; Terrian, D.M. Action of new world scorpion venom and its neurotoxins in secretion. *Toxicon* **1996**, *34*, 1399–1411. [[CrossRef](#)]
87. Jiménez-Ferrer, E.; Reynosa-Zapata, I.; Pérez-Torres, Y.; Tortoriello, J. The secretagogue effect of the poison from *Centruroides limpidus limpidus* on the pancreas of mice and the antagonistic action of the *Bouvardia ternifolia* extract. *Phytomedicine* **2005**, *12*, 65–71. [[CrossRef](#)]
88. Habermann, E.; Hardt, K. A sensitive and specific plate test for the quantitation of phospholipases. *Anal. Biochem.* **1972**, *50*, 163–173. [[CrossRef](#)]
89. Dhananjaya, B.L.; D'Souza, C.J.M. The pharmacological role of nucleotidases in snake venoms. *Cell Biochem.* **2010**, *28*, 171–177. [[CrossRef](#)]
90. Khan, N.; Niazi, Z.R.; Rehman, F.U.; Akhtar, A.; Khan, M.M.; Khan, S.; Baloch, N.; Khan, S. Hyaluronidases: A Therapeutic Enzyme. *Protein Pept. Lett.* **2018**, *25*, 663–676. [[CrossRef](#)] [[PubMed](#)]
91. Bordon, K.C.F.; Wiezel, G.A.; Amorim, F.G.; Arantes, E.C. Arthropod venom *Hyaluronidases*: Biochemical properties and potential applications in medicine and biotechnology. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **2015**, *21*, 165. [[CrossRef](#)]
92. Almonte, A.G.; Sweatt, J.D. Serine proteases, serine protease inhibitors, and protease-activated receptors: Roles in synaptic function and behavior. *Brain Res.* **2011**, *1407*, 107–122. [[CrossRef](#)] [[PubMed](#)]
93. Chen, Z.; Wang, B.; Hu, J.; Yang, W.; Cao, Z.; Zhuo, R.; Li, W.; Wu, Y. SjAPI, the first functionally characterized ascaris-type protease inhibitor from animal venoms. *PLoS ONE* **2013**, *8*, e57529. [[CrossRef](#)] [[PubMed](#)]
94. Zhao, R.; Dai, H.; Qiu, S.; Li, T.; He, Y.; Ma, Y.; Chen, Z.; Wu, Y.; Li, W.; Cao, Z. SdPI, the first functionally characterized kunitz-type trypsin inhibitor from scorpion venom. *PLoS ONE* **2011**, *6*, e27548. [[CrossRef](#)] [[PubMed](#)]
95. Law, R.H.P.; Zhang, Q.; McGowan, S.; Buckle, A.M.; Silverman, G.; Wong, W.; Rosado, C.J.; Langendorf, C.G.; Pike, R.N.; I Bird, P.; et al. An overview of the serpin superfamily. *Genome Biol.* **2006**, *7*, 216.

96. Gronenborn, A.M.; Nilges, M.; Peanasky, R.J.; Clore, G.M. Sequential resonance assignment and secondary structure determination of the ascaris trypsin inhibitor, a member of a novel class of proteinase inhibitors. *Biochemistry* **1990**, *29*, 183–189. [[CrossRef](#)]
97. Meekins, D.A.; Kanost, M.R.; Michel, K. Serpins in arthropod biology. *Semin. Cell Dev. Boil.* **2017**, *62*, 105–119. [[CrossRef](#)] [[PubMed](#)]
98. Gibbs, G.M.; Roelants, K.; O'Bryan, M.K. The CAP Superfamily: Cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins—Roles in reproduction, cancer, and immune defense. *Endocr. Rev.* **2008**, *29*, 865–897. [[CrossRef](#)]
99. Abraham, A.; Chandler, D.E. Tracing the evolutionary history of the CAP superfamily of proteins using amino acid sequence homology and conservation of splice sites. *J. Mol. Evol.* **2017**, *85*, 137–157. [[CrossRef](#)]
100. Fry, B.G.; Roelants, K.; Champagne, D.E.; Scheib, H.; Tyndall, J.D.; King, G.F.; Nevalainen, T.J.; Norman, J.A.; Lewis, R.J.; Norton, R.S.; et al. The toxicogenomic multiverse: Convergent recruitment of proteins into animal venoms. *Annu. Genom. Hum. Genet.* **2009**, *10*, 483–511. [[CrossRef](#)]
101. Wang, J.; Duan, Y.; Guo, M.; Huang, Q.; Liu, Q.; Niu, L.; Teng, M.; Hao, Q.; Shen, B.; Cheng, X.P.; et al. Blocking effect and crystal structure of natrin toxin, a cysteine-rich secretory protein from *Naja atra* Venom that Targets the BK Ca Channel<sup>†,‡</sup>. *Biochemistry* **2005**, *44*, 10145–10152. [[CrossRef](#)] [[PubMed](#)]
102. Gibbs, G.M.; Orta, G.; Reddy, T.; Koppers, A.J.; Martínez-López, P.; De La Vega-Beltrán, J.L.; Lo, J.C.Y.; Veldhuis, N.; Jamsai, D.; McIntyre, P.; et al. Cysteine-rich secretory protein 4 is an inhibitor of transient receptor potential M8 with a role in establishing sperm function. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 7034–7039. [[CrossRef](#)] [[PubMed](#)]
103. Ruiming, Z.; Yibao, M.; Yawen, H.; Zhiyong, D.; Yingliang, W.; Zhijian, C.; Wenxin, L. Comparative venom gland transcriptome analysis of the scorpion *Lychas mucronatus* reveals intraspecific toxic gene diversity and new venomous components. *BMC Genom.* **2010**, *11*, 452. [[CrossRef](#)]
104. Rokyta, D.R.; Ward, M.J. Venom-gland transcriptomics and venom proteomics of the black-back scorpion (*Hadrurus spadix*) reveal detectability challenges and an unexplored realm of animal toxin diversity. *Toxicon* **2017**, *128*, 23–37. [[CrossRef](#)] [[PubMed](#)]
105. Miyashita, M.; Otsuki, J.; Hanai, Y.; Nakagawa, Y.; Miyagawa, H. Characterization of peptide components in the venom of the scorpion *Liocheles australasiae* (Hemiscorpiidae). *Toxicon* **2007**, *50*, 428–437. [[CrossRef](#)]
106. Possani, L.D.; Becerril, B.; Riaño-Umbarila, L.; Juárez-González, V.R.; Olamendi-Portugal, T.; Ortíz-León, M.; Riaño-Umbarila, L.; Juárez-González, V.R.; Olamendi-Portugal, T.; Ortiz-León, M. A strategy for the generation of specific human antibodies by directed evolution and phage display. *FEBS J.* **2005**, *272*, 2591–2601.
107. Riaño-Umbarila, L.; Rudiño-Piñera, E.; Becerril, B.; Torres-Larios, A.; Canul-Tec, J.C.; Possani, L.D. Structural basis of neutralization of the major toxic component from the scorpion *Centruroides noxius Hoffmann* by a human-derived single-chain antibody fragment. *J. Boil. Chem.* **2011**, *286*, 20892–20900.
108. Zamudio, F.; Saavedra, R.; Martin, B.M.; Gurrola-Briones, G.; Herion, P.; Possani, L.D. Amino acid sequence and immunological characterization with monoclonal antibodies of two toxins from the venom of the scorpion *Centruroides noxius Hoffmann*. *JBIC J. Boil. Inorg. Chem.* **1992**, *204*, 281–292. [[CrossRef](#)]
109. Fastqc. Available online: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed on 29 April 2019).
110. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)] [[PubMed](#)]
111. ProP. Available online: <http://www.cbs.dtu.dk/services/ProP/> (accessed on 29 April 2019).
112. SpiderP. Available online: <http://www.arachnoserver.org/spiderP.html> (accessed on 29 April 2019).
113. Mafft. Available online: <https://www.ebi.ac.uk/Tools/msa/mafft/> (accessed on 29 April 2019).
114. Hall, T.A. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **1999**, *41*, 95–98.
115. Rstudio. Available online: <https://www.rstudio.com/> (accessed on 29 April 2019).
116. Wego. Available online: <http://wego.genomics.org.cn/> (accessed on 29 April 2019).

