

GOPEN ACCESS

Citation: Quintero-Hernández V, Ramírez-Carreto S, Romero-Gutiérrez MT, Valdez-Velázquez LL, Becerril B, Possani LD, et al. (2015) Transcriptome Analysis of Scorpion Species Belonging to the *Vaejovis* Genus. PLoS ONE 10(2): e0117188. doi:10.1371/ journal.pone.0117188

Academic Editor: Claude Wicker-Thomas, CNRS, FRANCE

Received: October 16, 2014

Accepted: December 22, 2014

Published: February 6, 2015

Copyright: © 2015 Quintero-Hernández et al. This is an open access article distributed under the terms of the <u>Creative Commons Attribution License</u>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: The work described here was partially financed by grants from the Consejo Nacional de Ciencia y Tecnología (CONACyT) of the Mexican Government (SEP-CONACyT #153496) and Dirección General de Asuntos del Personal Académico (DGAPA) of the National Autonomous University of Mexico (grant #200113) to LDP. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. RESEARCH ARTICLE

Transcriptome Analysis of Scorpion Species Belonging to the *Vaejovis* Genus

Verónica Quintero-Hernández^{1‡}, Santos Ramírez-Carreto^{1‡}, María Teresa Romero-Gutiérrez¹, Laura L. Valdez-Velázquez², Baltazar Becerril¹, Lourival D. Possani^{1*}, Ernesto Ortiz^{1*}

1 Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología, Universidad Nacional Autonóma de México, Cuernavaca, Morelos, México, 2 Facultad de Ciencias Químicas and Facultad de Medicina, Universidad de Colima, Colima, México

‡ These authors contributed equally to this work.
* erne@ibt.unam.mx (EO); possani@ibt.unam.mx (LDP)

Abstract

Scorpions belonging to the Buthidae family have traditionally drawn much of the biochemist's attention due to the strong toxicity of their venoms. Scorpions not toxic to mammals, however, also have complex venoms. They have been shown to be an important source of bioactive peptides, some of them identified as potential drug candidates for the treatment of several emerging diseases and conditions. It is therefore important to characterize the large diversity of components found in the non-Buthidae venoms. As a contribution to this goal, this manuscript reports the construction and characterization of cDNA libraries from four scorpion species belonging to the Vaejovis genus of the Vaejovidae family: Vaejovis mexicanus, V. intrepidus, V. subcristatus and V. punctatus. Some sequences coding for channelacting toxins were found, as expected, but the main transcribed genes in the glands actively producing venom were those coding for non disulfide-bridged peptides. The ESTs coding for putative channel-acting toxins, corresponded to sodium channel β toxins, to members of the potassium channel-acting α or κ families, and to calcium channel-acting toxins of the calcin family. Transcripts for scorpine-like peptides of two different lengths were found, with some of the species coding for the two kinds. One sequence coding for La1-like peptides, of yet unknown function, was found for each species. Finally, the most abundant transcripts corresponded to peptides belonging to the long chain multifunctional NDBP-2 family and to the short antimicrobials of the NDBP-4 family. This apparent venom composition is in correspondence with the data obtained to date for other non-Buthidae species. Our study constitutes the first approach to the characterization of the venom gland transcriptome for scorpion species belonging to the Vaejovidae family.

Introduction

Scorpions are one of the oldest groups of living organisms on the earth with an estimated apparition over 400 million of years ago $[\underline{1,2}]$. These arthropods gather around 1700 species, which



Competing Interests: The authors have declared that no competing interests exist.

are distributed in almost any place in the world [3,4]. Scorpions protect themselves against predators and capture preys by means of sophisticated venoms that contain a molecular arsenal of different toxins and other peptides, which are synthesized in and secreted from specialized glands. Scorpion venoms can be considered as true combinatorial libraries of peptides [5]. They contain toxins with lethal effects on humans [6] and potential therapeutic molecules, which have evolved under selective pressure [7]. As early as the 1960s venoms were already known as complex mixtures of dozens of other (i.e. not neurotoxic) peptide components. Furthermore, enzymatic and cytolytic activities of whole or fractionated venoms have been known at least since the 1920s [7]. Recently, scorpion venom peptides have been classified into two main groups: the disulfide-bridged peptides (DBPs), which affect membrane-bound ion chan $nels[\underline{8-10}]$ and the non-disulfide-bridged peptides (NDBPs), which show multiple activities [11–13]. DBPs contain three to four disulfide bridges which can be further sub-classified, into four different families depending on the type of membrane channels that they affect. These families of toxins in order of medical relevance are the ones that affect: Na⁺, K⁺, Ca²⁺ and Cl⁻ channels respectively [4]. These channels perform fundamental activities geared at regulating the normal cellular physiology in mammals. The perturbation of the regular activity of these channels caused by the binding of scorpion venom peptides can result in significant alterations leading to the well-known symptoms that appear during envenomation [14,15].

The fortuitous medical relevance of scorpion neurotoxins (the "killing component" with lethal effects on humans [16]), and their associated importance as pharmacological tools for the evaluation of ion channel's function, has biased research priorities towards the characterization of these particular components of scorpion venoms [7]. However, the non-toxic (NDBPs) component of the venoms with potential therapeutic properties has just started to be characterized. Several NDBPs have been identified and functionally characterized only in the last decade. Some of the biological functions determined in NDBPs include: antimicrobial, antiviral, antimalarial, anticancer, hemolytic, anti-inflammatory, immune-modulatory and bradykinin potentiating activities [17]. Those activities, though clinically relevant, were frequently accompanied by low selectivity and high cytotoxicity in many of the reported NDBPs. This hindered their potential therapeutic applications [18,19]. For example, hadrurin, the first reported scorpion NDBP with antibacterial activity, inhibited the growth of Gram-positive and Gram-negative bacteria with a Minimal Inhibitory Concentration (MIC) of 10–50 µM, but it also resulted highly hemolytic to human erythrocytes at 30 μ M [20]. A decade of research in the field of scorpion NDBPs has yielded several peptides with relatively mild cytotoxic effects and improved therapeutic indexes (the ratio between the toxic and the therapeutic doses), opening a window to their use as drug candidates [21].

Early methods for the identification of NDBPs were based on chromatographic separations combined with mass spectrometry. Further characterization included biological assays dedicated to determine the pharmacological properties of the potential therapeutic peptides [22]. These early biochemical characterizations were later complemented by Edman degradation peptide sequencing, although with technical limitations when sequencing lengthy peptides and/or with N-terminal modifications [20,23]. One of the major breakthroughs in scorpion venom peptidomics was the combination of molecular cloning with HPLC fractionation and mass spectrometry. This integration of technical tools even with the limitations imposed by limited N-terminal sequence information was circumvented by designing degenerate primers allowing the cloning of full-length cDNA sequences of the toxin precursor peptides [24].

The results obtained by several groups using cDNA libraries, revealed the existence of a rich biodiversity of components in the species studied. Apart from the peptides that modify ion-channel permeability, initially isolated and characterized by classical biochemical methods from species belonging to Buthidae family, many other components were identified, such as:

factors that activate lipolysis, phospholipase A2, serine-proteases, metalloproteinases, protein homologs of tick salivary glands, precursors of cytolytic peptides, cysteine-rich peptides with no homology with common scorpion toxins and a great number of proteins and peptides deduced from the ESTs for which the function is still unknown [7,25–27]. Confirmation of these findings and new information about novel potential therapeutic peptides in Buthidae species has been recently reported by our group [28]. Furthermore, construction of cDNA libraries allowed the screening of several random cDNA clones, proving to be a successful strategy for the identification of several putative NDBPs reported elsewhere[13].

The determination of the transcriptomic profile of NDBPs provides additional information about the post-translational processing and the evolutionary characteristics of such peptides. These approaches are proving to be important tools for taxonomy as well [7]. Taken together, recent proteomic and transcriptomic analysis of scorpion venom and venom glands, have provided new data that have critically increased our knowledge about scorpion venom biology and on the cellular processes taking part on the highly specialized venom glands. Particular emphasis is made on novel components whose function, whether in the context of venom effects in susceptible animals or within the venom gland itself, has just started to be known [7].

Mass fingerprinting studies have revealed that NDBPs constitute more than 30% of all peptides present within scorpion venoms [7,29]. In spite of constituting a significant fraction of the total venom, this group of peptides has been poorly characterized when compared with DBPs that represent the majority of functionally characterized scorpion venom peptides reported in the literature. The importance of characterizing NDBPs is sustained by their biological and structural diversity associated to their multiple potential effects on cognate targets.

The Buthidae family, to which poisonous scorpions belong, makes up more than 50% of the approximately 1700 scorpion species currently known [30]. From approximately 800 Buthidae species only 34, are potentially dangerous to humans [6]. Non-Buthidae species have been largely neglected by toxicological research. Recently it has been demonstrated that even neglected lineages of scorpions are a rich source of novel biochemical components, which have evolved over millions of years to target specific ion channels in prey animals, but also new NDBPS with significant implications in therapeutics [<u>31,32</u>]. Non-Buthidae scorpions produce venoms with low or no toxic effects to mammals. The Vejovidae family is one of the non-Buthidae scorpion families that does not pose any risk for human health and which members can be localized in the North American subcontinent, mostly in Mexico [3]. The Vaejovidae family is composed of 17 genera, including Smeringurus, Paruroctonus, Pseudouroctonus, Serradigitus, and Vaejovis, among others. This last genus, Vaejovis (corr. Vejovis): 1836, C.L. Koch, is distributed from the southwestern United States to Guatemala, with the large majority of the 70 species inhabiting the Mexican territory [33-35]. In accordance with the classification proposed by Sissom in 2010, the Vaejovis genus is subdivided in 5 groups: eusthenura, intrepidus, mexicanus, nitidulus and punctipalpi, plus a group classified as incertae sedis. Examples of species belonging to some of the groups are Vaejovis punctatus (eusthenura group), Vaejovis intrepidus and Vaejovis subcristatus (intrepidus group), Vaejovis mexicanus (mexicanus group), among others.

In scorpions of the *Vaejovis* genus, the venom contains lesser amounts of ion channel toxins and higher amounts of NDBPs [36]. From two Mexican species belonging to the Vejovidae family (*V. mexicanus* and *V. subcristatus*), some NDBPs with therapeutic properties have been identified: Vejovine, with antibiotic activity against a broad spectrum of clinical isolates of bacteria from different genera [37]; Vm24, an immune-suppressive peptide selective to Kv1.3 potassium channels of human lymphocyte T cells [38,39]; VmCT1 and VmCT2, two antimicrobial peptides with a broad-spectrum activity against Gram negative and Gram positive bacteria [40]; and VsCT1 and VsCT2, two hemolytic peptides [40].

All this information as a whole is changing the paradigm of the study of scorpion venom toward a vision beyond their toxic properties. Every day new peptides with therapeutic properties are discovered. These results encourage to further study the venoms of non-Buthidae scorpions, in which NDBPs constitute a rich mine to be exploited. The detailed transcriptomic profile of the venom glands of species belonging to *Vaejovis* genus will retrieve valuable information that would help to identify new potential candidates to be used as therapeutic agents against different diseases. In this contribution, we report the generation and characterization of the cDNA libraries corresponding to four species of scorpions from the *Vaejovis* genus: *V. mexicanus*, *V. subcristatus*, *V. punctatus* and *V. intrepidus*.

Materials and Methods

2.1. Biological Material

The scorpion specimens used in this work were collected in several locations in Mexico, under the official permit SGPA/DGVS/10638/11 by the Mexican Federal Government (issued by *Secretaría de Medio Ambiente y Recursos Naturales*, SEMARNAT). *V. mexicanus* and *V. punctatus* were collected in Cuernavaca, Morelos; *V. intrepidus* in Colima, Colima; and *V. punctatus* in Totalco, Veracruz.

The specimens were identified and classified in accordance with the available taxonomic keys [<u>34,41</u>]. The animals were kept in captivity, under standard conditions of temperature, humidity, and light and dark periods. They were periodically fed with crickets and supplied with water *ad libitum*.

2.2. cDNA Library Construction

Scorpions produce and store their venom only in the venom glands. These are a pair of well-delimited structures within the telson, which is the last segment of the metasome. For library construction, total RNA was extracted from telson macerates. For V. punctatus and V. subcristatus one specimen was used, whereas for V. mexicanus and V. intrepidus two telsons were dissected. Total RNA was extracted with the "SV Total RNA Isolation" Kit (Promega, Madison, WI) in accordance to the instructions from the manufacturer. Full-length cDNA libraries were prepared using the Creator SMART cDNA Library Construction Kit (CLONTECH Lab., Palo Alto, CA), designed to preferentially enrich for full-length cDNAs. For the first-strand cDNA synthesis, the provided oligonucleotides SMART IV (5'-AAGCAGTGGTATCAACGCAGAG TGGCCATTACGGCCGGG-3') and CDS III/3' PCR (5'-ATTCTAGAGGCCGAGGCGGCCG ACATG-d(T)₃₀N₋₁N-3'; with N = A, G, C, or T; $N_{-1} = A$, G or C) were used as primers. For cDNA amplification, the oligonucleotides 5' PCR Primer (5'-AAGCAGTGGTATCAACG CAGAGT-3') and CDS III/3' PCR Primer were used. For both reactions, the conditions suggested by the manufacturer were followed. After the amplification, asymmetrical SfiI restriction sites (named SfiIA and SfiIB sites) flank the resulting cDNA molecules, allowing directional cloning. Following SfiI digestion, the cDNA was fractionated on CHROMA SPIN-400 columns to enrich the medium-sized fragments that correspond to the expected size of peptide-coding mRNAs. The selected fractions were pooled and concentrated by ethanol precipitation. The cDNA was ligated to the pDNR-LIB plasmid through the SfiIA and SfiIB sites, and the ligation was electroporated into electrocompetent Escherichia coli DH5a cells. Four full-length cDNA libraries were produced with this protocol. The titer of the non-amplified libraries was in the order of 1x10⁶ cfu/mL for V. mexicanus, V. intrepidus and V. punctatus and 1x10⁵ cfu/mL for V. subcristatus, all the libraries with 99% recombinant clones.

2.3. DNA Sequencing and Bioinformatic Analysis

Plasmids from selected colonies were isolated according to the standard alkaline lysis protocol [42]. The cloned inserts were sequenced on one direction with the T7 primer (5'-GTAATAC GACTCACTATAGGG-3') using an automatic machine (Model 3100, Applied Biosystems, Foster city, CA) according to the manufacturer's instructions.

The nucleotide sequences obtained in this work were deposited in the GenBank (EST database: dbEST JZ818318-JZ818449). Phred scores were used to assess the sequence quality and to remove low-quality regions (end clipping) [43,44], with the window length set to 100 and the standard quality to 20. The cross_match utility was used to remove irrelevant vector and E. coli DNA sequences (http://www.phrap.com). ESTs that shared an identity of >95 out of 100 nucleotides were assembled in contiguous sequences with the CAP3 program (http://pbil.univlyon1.fr/cap3.php) [45]. All the bioinformatic analyses were performed online through the server of the Laboratorio de Biologia Molecular, Universidade de Brasilia (http://www.biomol. unb.br), using the default settings. The Vaejovis mexicanus, V. intrepidus, V. punctatus and V. subcristatus venom gland ESTs (clusters and singlets) were searched against the non-redundant nucleotide and protein databases using the blastn and blastx algorithms, respectively (http:// <u>www.ncbi.nlm.nih.gov/blast</u>) with an e-value cutoff set to $<10^{-5}$ to identify matches for the new ESTs. Additional search was performed with ORF Finder (Open Reading Frame Finder; http://www.ncbi.nlm.nih.gov/projects/gorf/), PROSITE (http://prosite.expasy.org/) and Pfam (http://pfam.sanger.ac.uk). Signal peptides for secreted proteins were predicted with the SignalP 4.0 program (http://www.cbs.dtu.dk/services/SignalP/). The possible propetide regions were predicted by using the Prop 1.0 server (http://www.cbs.dtu.dk/services/ProP/) and the SpiderP server (http://www.arachnoserver.org/spiderP). Multiple sequence alignments were performed with ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/), ClustalX 2.0 and T-Coffee (http://www.ebi.ac.uk/Tools/msa/tcoffee/). The consensus sequences were determined with CLC Sequence Viewer (v 6.8.1). The percentages of sequence identity were determined with LALIGN (http://embnet.vital-it.ch/software/LALIGN form.html). The peptides chemical parameters, net charge and molecular weight were predicted with ProtParam (http://web. expasy.org/protparam/).

Results and Discussion

3.1. Transcriptomic Analysis of the Vaejovis cDNA Libraries

For the four species used in this work, cDNA libraries were constructed and screened for sequences encoding venom components. Their titers ranged from hundreds of thousands (*V. subcristatus*) to millions (*V. mexicanus*, *V. intrepidus* y *V. punctatus*) of colony-forming units (cfu).

From the *V. punctatus* library 71 clones were sequenced, of which 28 ESTs were clustered in 11 contigs (corresponding to two or more ESTs) and 43 singlets (unique sequences). The ESTs were putatively classified in accordance with their similarity to reported sequences as 12 (16.9%) toxin genes, 12 (16.9%) antimicrobial peptides, 6 (8.5%) proteins involved in cellular processes, 6 (8.5%) proteins with unknown function, 1 (1.4%) La1-like peptide, 1 (1.4%) corresponded to other venom components, and 33 (46.5%) had no match with any sequence in the databases.

For *V. subcristatus*, 46 clones were sequenced. Twenty-four ESTs were grouped in 8 contigs and the remaning 22 were singlets. By sequence similarity, the ESTs were identified as possibly coding for 2 (4.3%) toxins, 6 (13%) antimicrobial peptides, 4 (8.7%) proteins involved in

cellular processes, 12 (26.1%) proteins with undefined function, 1 (2.2%) La1-like peptide, 1 (2.2%) matched with other venom components, and 20 (43.5%) found no match in the databases.

Two hundred twenty-five clones were sequenced from the *V. mexicanus* library, resulting in 197 ESTs, clustered in 36 contigs and 61 singlets. Considering their possible function by homology, the ESTs were classified as 23 (11.7%) toxins, 119 (60.4%) antimicrobial peptides, 17 (8.6%) proteins involved in cellular processes, 17 (8.6%) proteins with undefined function, 1 (0.5%) La1-like peptide, 2 (1%) showed similarity with other venom components and 18 (9.1%) had no match with the sequences in the databases.

From the *V. intrepidus* library 103 clones were sequenced. Fifty-five ESTs were grouped into 17 contigs and the remaining 48 were singlets. Eleven (10.7%) matched with toxin sequences, 23 (22.3%) with antimicrobial peptides, 20 (19.4%) with proteins involved in cellular processes, 14 (13.6%) with proteins with undefined function, 4 (3.9%) with La1-like peptides, 2 (1.9%) with other venom components, and the remaining 29 (28.1%) found no homologs in the databases.

<u>Fig. 1</u> shows a comparison of the found relative abundance of the different peptide classes in the sequenced clones of the *Vaejovis* libraries. It is clear that, with respect to the remaining venom components, transcripts coding for ion channel-acting toxins are relatively scarce in the *Vaejovis* libraries. This is in accordance with previous findings [36].

3.2. Disulfide-bridged peptides found in the Vaejovis species

The venoms of the non-Buthidae scorpions contain non-disulfide bridged peptides as their main components [10,46-52], whereas some disulfide-bridged peptides can also be present. The sequence analysis performed with the cDNA libraries of the four *Vaejovis* species revealed the presence of precursor sequences coding for sodium, potassium and calcium channel toxins,



Fig 1. Relative distribution of the sequences found in the V. intrepidus, V. mexicanus, V. subcristatus and V. punctatus libraries with respect to the putative function of their encoded peptides.

doi:10.1371/journal.pone.0117188.g001

plus some sequences corresponding to La1-like and scorpine-like peptides, all of which are rich in cysteines.

3.3. Sodium channels toxins (NaScTx)

The scorpion toxins that alter the functionality of sodium ion channels are 58–76 amino acidslong peptides with a molecular weight of 6,500–8,500 Da [53]. These toxins are abundant in the venoms of the scorpions belonging to the Buthidae family and are responsible for the toxicity of several of those species on mammals, including humans [53]. In the non-Buthidae scorpions they are relatively less represented, and those found are generally active against insect sodium channels [46]. In the *V. intrepidus*, *V. punctatus*, *V. subcristatus*, and *V. mexicanus* libraries, sequences coding for putative NaScTx's were found (ViNaTx1, VpNaTx1 and VpNaTx2, VsNaTx1 and VsNaTx2, and VmNaTx1, respectively). The mature peptides derived from the translated sequences contain six conserved cysteines and share sequence similarity with known scorpion sodium channel-acting β-toxins (Fig. 2A). The mature *Vaejovis* NaScTxlike peptides display high sequence identity with each other, from 52.5% and up.

To date, the transcriptomic analyses performed with scorpions have revealed more than 140 sequences coding for putative NaScTx's [16]. Some Buthidae scorpions apparently produce a large diversity of these toxins. For example, 39 NaScTx-like coding sequences were reported from a cDNA library of *Buthus occitanus israelis* [54] and 27 from a massive RNA-seq with *Centruroides noxius* [55]. In sharp contrast with this, the transcriptome analyses of non-Buthidae scorpions have yielded no putative NaScTx-coding sequences or just a few. The trascriptomes of *Hadrurus gertschi* [10], *Opisthacanthus cayaporum* [52], *Scorpiops jendeki* [49], *Heterometrus petersii* [50], *Pandinus cavimanus* [47] and *Scorpiops margerisonae* [31] revelaed no putative NaScTx-coding sequences, whereas those of *Scorpio maurus palmatus* [46] and *U. yaschenkoi* [48] returned only one each. Our analyses are in line with these reports, with one putative NaScTx found for *V. intrepidus* and *V. mexicanus*, and two found for *V. punctatus* and *V. subcristatus*. Since the NaScTx's are known to be the main venom components responsible for the severe envenomation symptoms produced by the sting of the dangerous scorpion species [53], their low representativeness in the venoms of the Vaejovidae explain why none of the four species here studied are considered dangerous to humans [56].

3.4. Potassium channel toxins (KScTx)

Scorpion venoms are rich in toxins that affect potassium channels. They are classified in accordance to their primary sequence and their cysteine arrangement in four families: α -, β -, γ - and κ -KTx's [57]. The α -KTx family is the most abundant in the scorpion venoms. These peptides are composed of 23–42 amino acids and adopt a typical cysteine-stabilized α/β fold with three or four disulfide bridges [58]. We found precursor sequences for putative α -KTx's of the two kinds. First, those with three disulfide bridges: ViKTx1 from *V. intrepidus* and VpKTx4 from *V. punctatus*. An alignment of these possible toxins with other members of the α -KTx subfamily, is shown in Fig. 2B. Second, those with four disulfide bridges: VmKTx1 from *V. mexicanus* and VpKTx3 from *V. punctatus*. These are shown in Fig. 2B, aligned to two previously reported toxins of the same kind for comparison.

A precursor sequence coding for a putative κ -KTx was also found in the *V. mexicanus* library: VmKTx2. κ -KTx's are toxins conformed by two parallel α -helixes connected by two disulfide bonds [16]. The precursor sequence for VmKTx2 contains a signal peptide, a propeptide and a predicted 27 amino acids-long mature peptide (Fig. 2C).

Our results are in correspondence with previous reports for non-Buthidae species. In the transcriptome of *H. gertschi* two α -KTx- and one β -KTx-coding precursors were found [10].

58 100

aa %I 78 **100** 78 98.7

78 93.6

A) β-NaTx (3 C-C):

VsNaTx1

VpNaTx1	MKCFAFLTVCLAVVLFEFTEVEGKDG-YPLDANGNVFRC-VPGKLEVCKSACSERGSDQVSCLAFTCWCDGMNDDAKTW	/KHS 5/	94.8
ViNaTx1	MKCFAFLTVCLAVVLFEFTEVEG-KKDG-YPVDANGNVFRC-VPGKLEVCKSACGERGSDQISCLAFTCWCDGMNDDAETV	KHD 58	87.9
VsNaTx2	MKCFTFLTVCLAVVLFEFTEVEG-KIYG-YPLDANGKVFKC-VPGKLEVCKSACSERGSDOVSCLAFTCWCDGMNDDAKTV	HTOLKLI 62	87.1
VmNaTx1	MKCCAILILCLVVFLFEVSDVQGS-KSG-YPLGANGKVTDC-VPHKLAKCKSTCGAHGSNKASCLAFTCWCDGMNDDVKVV	/KRE 58	67.2
VpNaTx2	MKCYAFLTFCLIVLFFESSDGKKDG-YPVDKYGRVSNC-VVGMLGDNTFCESICRSRGGS-GYCYVFACWCDGMNDDVKIV	KKG 60	52.5
Birtoxin	ECLNVCKLHGVOYGYCYASKCWCEYLEDDKDGNTYKCFLLGGNEECLNVCKLHGVOYGYCYASKCWCEYLEDDKDS	/ 58	31.1
Ikitoxin	ECLNVCKLHGVOYGYCYASKCWCEYLEDDKDGNTYKCFLLGENEECLNVCKLHGVOYGYCYASKCWCEYLEDDKDSV	/ 58	31.1
Altitoxin	nKDCQKVCKLHGVQYGYCYAFFCWCKEL-DDKDGNTYTCLELGENKDCQKVCKLHGVQYGYCYAFFCWCKEL-DDKDVS	SV 58	27.9
B)			
a-KTy (3 C	C-C):		
winin (o c	22 ST		
Vikmy1	MEVERAULTELEVCS_MITCINAUDIDUSCERESCOTEDERD_ACMPECETINDOCOCEPER 37 100		
VIKIXI VDKTXA	MSTYPRINT WSSVITTEDIAWALTYSCHESSGETRE AGAALAGAALAAGAAG		
VPRIX4			
mdr1			
Iuni			
α-KTx (4 C	C-C):		
	aa %I		
VmKTx1	MKTSCLLTILLLSFLVAVAVAEGERSARAAAISCVGSPECPPKC-RAQGCKNGKCMNRKCECYYCG- 37 100		
VpKTx3	MGTKFVFIILLATTVILTFGALTVEGGKKCSGSKECYDYCYKTTKCPRAKCINRMCKCYGCGG 43 39.5		
Vm24	RAQGCKNGKCMNRKCKCYYC 36 94.6		
HgeTx1	TGTSCISPKQCTEPC-RAKGCKHGKCMNRKCHCMLCL- 36 51.4		
C)			
к-КТх (2 С	C-C):		
	aa %I		
VmKtx2	MKFLALTFVFLLVCVMLPNQASGTPLGEENEYYLEAARARKYLAKACSRLC-DKVPLIKFCIGMCP 27 100		
k-HfTx2	GHACYRNCWREGNDEETCKERCG 23 21.4		
k-HfTx1	GHACYRNCWREGNDEETCKERC 22 21.4		
HelaTx1	MKLLPLLFVILIVCAILPDEASCDQSELERKEENFKDESREIVKRSCKKEC-SGSRRTKKCMQKCNREHGHGR 27 14.7		
D)			
Scornine-lil	ike (3 C-C):		
Scor pine-in	ike (3 C-C):		
ViScplp1	MNAKITVLFFLVAITIASCAWISEKKIQDAIDKRLPNGAVKSAIKAVVHKAAKNKHG <mark>C</mark> LADFDVGGG <mark>C</mark> EQH <mark>C</mark> RKTEI	KEGVCHGTKCI	GIPRAYRI
ViScplp2	MNAKITVLFFLVAITIASCAWISEKKIQDAIDKRLPNGAVKSAIKAVVHKAAKNKHGCLADFDVGGGCEQHCRKTEI	KEGVCHGTKCK	GIPRAYRI
VpScplp1	MNAKITVLFFLVAITIASCAWISEKKIQDAIDKKLPNGMVKNAIKAVVHKAAKNKHGCLADFDVGGGCEQHCRKTEL	KEGVCHGTKCK	GIPRAYKI
VmScplp1	MNAKITVLIFLVAITIASCGWISEKKAQDFVDSKLPNGMVKNAIKAVIHKASKNKHGCFADIDVGGGCTSHCKK-EI	KEGICHGTKCKC	GIPRAYK
17mCanln2	TYPETTITITITE COCAT TERVELYCUPATED TETEVCCTUMMUAROWTHETCHTOCT CAEDEDWACT CEVCOVER	VYCUCHCTYCY	CYDIDYY

MKCFTFLTVCLAVVLFEFTEVEG-KKDG-YPLDANGKVFKC-VPGKL---EVCKSACSERGSDQVSCLAFTCWCDGMNDDAKTWKHS-

VmScplp1	MNAKITVLIFLVAITIASCGWISEKKAQDFVDSKLPNGMVKNAIKAVIHKASKNKHGCFADIDVGGGCTSHCKK-EDKEGICHGTKCKCGIPRAYKK	77	75.6	
VmScplp2	IKFTILLLLVISTFCSCALIREKYFHKGVDALTPLIPIPVVGGIVNNVAKQMIHKIGKIQSLCAFDKDMAGLCEKKCQKTEKKKGVCHGTKCKCGKPLDYK-	84	42.4	
VmScplp3	MQKFFAVLCILGLATFCTCGLIREKHIQKGVDALTNLIPAPVIGGIINKVAKQMVHKLGKVQEL <mark>C</mark> TFNKDMMGWCEASCQEKEQAKGYCHGTKCKCGKPLGYK-	84	38.8	
ViScplp3	MQIKFTILILLVISTFCSCGLIREKYFHQGVDALTPLIPVPVVGGVVNKVAKQMIHKIGKIQSLCAFNKDMASLCEKKCQETEKVKGYCHGTKCKCGKPLSYK-	84	35.3	
VsScplpl	MVKNAIKAVVHKAAKNKHGCLADFDVGGGCEQHCKKTEDKEGVCHGTKCKCGIPRAYKK	59	nd	
VsScplp2	NKDMMGWCEKACQEKEQTNGFCHGTKCKCGKPLSYK-	36	nd	
HgeScplp1	MNTKLTVLCFLGIVTIVSCGWMSEKKVQGILDKKLPEGIIRNAAKAIVHKMAKNQFGCFANVDVKGDCKRHCKA-EDKEGICHGTKCKCGVPISYL-	76	57.7	
Scorpine	MNSKLTALIFLGLIAIAYCGWINEEKIQKKIDERMGNTVLGGMAKAIVHKMAKNEFQCMANMDMLGNCEKHCQTSGEK-GYCHGTKCKCGTPLSY	75	48.7	
HgeScplp2	MKLTILILLVITSFCSCGILREKYAHKAIDVLTPMIGVPVVSKIVNNAAKOLVHKIAKNOOLCMFNKDVAGWCEKSCOOSAHOKGYCHGTKCKCGIPLNYK-	84	41.2	

Fig 2. Putative toxin sequences derived from the precursors found in the Vaejovis libraries. The peptide length (aa) always refers to the confirmed (when the peptide has been isolated from the venom) or software-predicted mature peptides. The identity (%I) is always relative to the first sequence of the alignment, and considers only the mature peptide regions. When present, the signal peptides are shown underlined and the propertides are in italics and bold. The conserved cysteine arrangement typical of each family is highlighted. A) Sequence alignment of the putative sodium channel toxins. The precursor sequences of ViNaTx1 from V. intrepidus, VpNaTx1 and VpNaTx2 from V. punctatus, VsNaTx1 and VsNaTx2 from V. subcristatus, and VmNaTx1 from V. mexicanus are compared to the known sodium channel-specific ß-toxins Birtoxin (UniProt:P58752), Ikitoxin (UniProt:P0C1B8) and Altitoxin (UniProt: P0C1B5) from Parabuthus transvaalicus. B) Comparative alignments of the sequences found that belong to the two families of potassium channel-specific αtoxins. First, the precursor sequences of the six-cysteines α-toxins ViKTx1 form V. intrepidus and VpKTx4 from V. punctatus aligned to KTX-2 (Kaliotoxin-2, UniProt:P45696) from Androctonus australis and TdK1 (UniProt:P59925) from Tityus discrepans as references. Second, the precursor sequences coding for the toxins belonging to the family of the eight-cysteines a-toxins VmKTx1 from V. mexicanus and VpKTx3 from V. punctatus, aligned to Vm24 (UniProt: P0DJ31) from V. mexicanus and HgeTx1 (UniProt:P84864) from Hadrurus gertschi as references. C) The precursor sequence of VmKTx2 from V. mexicanus compared to the precursor sequence of HeIaTx1 (UniProt:P0DJ41) from Heterometrus laoticus and κ-HfTx1 (UniProt:P82850) plus κ-HfTx2 (UniProt: P82851) from Heterometrus fulvipes, all belonging to the family of potassium channel-specific κ-toxins. D) The scorpine-like sequences ViScplp1-3 from V. intrepidus, VmScplp1-3 from V. mexicanus, VpScplp1 from V. punctatus, and the partial sequences of VsScplp1-2 from V. subcristatus are aligned. For comparison, the sequences of scorpine (UniProt:P56972) from P. imperator, HgeScplp1 (UniProt:Q0GY40) and HgeScplp2 (UniProt:P0C8W5) from H. gertschi are also included in the alignment.

doi:10.1371/journal.pone.0117188.g002

For *O. cayaporum* only one putative α -KTx was reported [52]. *S. jendeki* and *H. petersii* were more diverse, with eight [49] and four [50] transcripts respectively, all coding for putative α -KTx's. *P. cavimanus* produced three precursors, coding for one possible toxin of each of the α -, β -, and κ -KTx families [47]. *S. margerisonae* was shown to contain transcripts for two putative α -KTX's [<u>31</u>]. More recently, for *S. maurus palmatus* three transcripts were found, one for a putative α -KTx and two for possible κ -KTx's [<u>46</u>].

Selective blockers of potassium ion channels have a remarkable therapeutic potential in the treatment of several pathological conditions, including autoimmunity. For example, the proliferation of auto-reactive T lymphocyte lineages can lead to autoimmune diseases, including multiple sclerosis or type I diabetes [59]. High affinity blockers of the Kv1.3 potassium channels can specifically inhibit the proliferation of effector memory T (T_{EM}) cells [60]. Vm23 and Vm24, two α -KTx's purified from the venom of *V. mexicanus* are potent blockers of the human Kv1.3 channel [38]. The fact that we were able to identify cDNA sequences coding for yet undescribed putative KTx's from the *Vaejovis* libraries reflects the value that this genus has a source of bioactive compounds with medical importance.

3.5. Scorpine-like peptides (Scplp)

Scorpine is an interesting peptide isolated from the *Pandinus imperator* venom. Its carboxylterminal region shows sequence similarity with the β -KTx family, with the typical cysteine-stabilized α/β fold, with 3 disulfide bridges. Its amino-terminal region, on the other side, is similar to peptides of the cecropin family. Scorpine was shown to display antibacterial activity and potent inhibitory effects on the development of *Plasmodium berghei* at the ookinete and gamete stages [61]. Scorpine homologs were thereafter discovered for *Opistophthalmus carinatus* [62], *Heterometrus laoticus* [63] and *H. gertschi* [10]. From the last species, HgScplp1 showed antibacterial activity against *Bacillus subtilis* and cytolitic activity on *Xenopus laevis* oocytes. When the N-terminal region of HgScplp1 was cleaved with cyanogen bromide, the remaining peptide displayed itself as an effective blocker of the Kv1.1, Kv1.2 and Kv1.3 channels [64]. This dual activity is unique to scorpine-like peptides, which can therefore be considered as modular peptides.

In the *Vaejovis* transcriptomes several scorpine-like sequences were found: ViScplp1, ViScplp2 and ViScplp3 in *V. intrepidus*; VmScplp1, VmScplp2 and VmScplp3 in *V. mexicanus*; VpScplp1 in *V. punctatus*; and the partial sequences of VsScplp1 and VsScplp2 in *V. subcristatus*. The last two sequences were not considered for sequence analysis, since they are incomplete, but the remaining precursor sequences share some degree of sequence identity, mainly driven by the C-terminal CS- α/β motif (see the %I in Fig. 2D). An interesting pattern emerges when only the N-terminal region is considered. In this case, some of the here reported sequences group with Scorpine and HgScplp1 (ViScplp1, ViScplp2, VmScplp1 and VpScplp1) to conform a short-chain clade, whereas others (ViScplp3, VmScplp2 and VmScplp3) share Nterminal similarity with HgScplp2 and conform a long-chain clade. It is remarkable that (as in the case of *H. gertschi*), both short- and long-chain scorpine-like peptides can be found in the same *Vaejovis* species. Whether this has any biological implications remains to be determined.

3.6. Calcins

The scorpion toxins that affect calcium channels are referred to as calcins. They are structurally characterized by the presence of an inhibitor cystine knot (ICK) motif, a folding that puts them structurally apart from the sodium, potassium and chloride channel toxins [65]. Calcins can penetrate the cell membranes and affect the ryanodine receptors (RyR) of the endoplasmic and sarcoplasmic reticula of the cardiac and skeletal muscle [66,67]. The alterations in the calcium potentials result in muscular paralysis, which contributes to the immobilization of preys [66].

The first RyR-specific activity recorded for scorpion venom components was discovered in a gel-filtration fraction obtained from the venom of *Buthotus hottentota* [68]. After that, calcins with high affinity for the RyR were purified from the *P. imperator* venom: imperatoxins IpTxa and

LaIT1

U1-LITX-Lwla

ViCaTx1	MRQNTMTIIFIVFIVTFASLTIYGAEA S-EANFLERR ADCLAHLKLCKKNKDCCSKKCSRRGTNPEQRCR	33	100	
VpCaTx1	MKQNTMTIMFIVFFITFASLAIYGAEASEEANFLERRADCLPHLKLCKKNKDCCSKECSRRGTNPEQRCR	33	93.9	
Maurocalcin	GDCLPHLKLCKENKDCCSKKCKRRGTNIEKRCR	33	81.8	
Opicalcin-1	MKPSLIIVTFIVVFMAISCVAADDEQETWIEKRGDCLPHLKRCKENNDCCSKKCKRRGTNPEKRCR	33	78.8	
Opicalcin-2	MKPSLIIVTFIVVFMTISCVAA DDEQETWIEKR GDCLPHLKRCKENNDCCSKKCKRRGANPEKRCR	33	75.8	
Hadrucalcin	MKTSSLTIIFIAVIITIICLNIHDIEA REIEFNAGRVVR SEKD <mark>C</mark> IKHLQR <mark>C</mark> RENKD <mark>CC</mark> SKK <mark>C</mark> SRRGTNPEKR <mark>C</mark> R	35	71.4	
UyCaTx20	MKASTLVVIFIVIFITISSFSIHDVQA SGVEKR EQKDCLKKLKLCKENKDCCSKSCKRRGTNIEKRCR	35	71.4	
Imperatoxin	GDCLPHLKRCKADNDCCGKKCKRRGTNAEKRCR	33	69.7	
B)				
Calcins (2 C-C):			
			aa	8I
VmCaTx1	MGFATKISILFLMVAVISTITG REILGSLEEFGNSEIER DLPLSGEYETCVRPRKCKPNLHCNAAQICV	DPKK	G- 35	100
VmCaTx2	MGFATKISILFLMVAVISTITG REILGSLEEFGNSEIER GLPLSGEYET <mark>C</mark> VRPRK <mark>C</mark> KPNLHCNAAQIC	DPKK	G- 35	97.1
VpCaTx2	MNSVAKISFLLLLVAVISTVTG RDVVNVFESEETEESVETDTER DLPLSNEYET <mark>C</mark> VRPRK <mark>C</mark> KPHLKCNAAQICV	DPKK	GW 36	88.9

MNFATKVSLLLLAIAVIVIVEGGEGDSWFEEHEESDTER----DFPLSKEYESCVRPRKCKPPLKCNKAQICVDPNKGW

Fig 3. Putative calcium channel toxins found in Vaejovis. The peptide length (aa) always refers to the confirmed (when the peptide has been isolated from the venom) or software-predicted mature peptides. The given percentages of identity (%I) correspond to the mature peptides only, and refer to the first sequence. The predicted signal peptides are shown underlined, the propeptides in italics and bold. A) Sequence alignment of the three-disulfide-bridged putative calcinsVpCaTx1 from V. punctatus and ViCaTx1 from V. intrepidus with imperatoxin (UniProt: P59868) from P. imperator, maurocalcin (UniProt: P60254) from S. maurus palmatus, opicalcin-1 (UniProt:P60252) and opicalcin-2 (UniProt:P60253) from Opistophthalmus carinatus, UyCaTx20 (Calciumchannel toxin-like 20, UniProt:AGA82762) from U. yaschenkoi and hadrucalcin (UniProt:B8QG00) from H. gertschi. B) The precursors of the two-disulfidebridged putative calcins VmCaTx1 and VmCaTx2 from V. mexicanus and VpCaTx2 from V. punctatus, are compared to previously reported sequences of their kind: LaIT1 (UniProt:P0C5F2) from L. australasiae, and U1-LITX-Lw1a (UniProt:P0DJ08) from L. waigiensis.

---DFPLSKEYETCVRPRKCOPPLKCNKAOICVDPKKGW

doi:10.1371/journal.pone.0117188.g003

IpTxi [67]. Other peptides were also characterized from different venoms, such as ryanotoxin from Buthotus judaicus [69] and the toxins Bmk-AS and Bmk-AS-1 from Buthus martensii Karsch [70].

%I aa

80.6

77.8

36

36

In the Vaejovis libraries we found two precursor sequences coding for putative calcins with three disulfide bridges: VpCaTx1 from V. punctatus and ViCaTx1 from V. intrepidus. Both toxins are composed of 33 amino acids, with 6 cysteines. They show high similarity with previously reported calcins, such as imperatoxin A from *P. imperator* [67], maurocalcin from *S. maurus palmatus* [71], hadrucalcin from *H. gertschi* [66], among others (Fig. 3A). The two Vaejovis calcins differ only by two residues, and share 70% identity with imperatoxin A, the most studied calcin to date.

Three precursors coding for possible calcins with two disulfide bridges were also found: VmCaTx1 and VmCaTx2 from V. mexicanus, and VpCaTx2 from V. punctatus. They display high sequence identity with the insect toxin LaIT1 from *Liocheles australasiae* [72] and toxin U_1 -LITX-Lw1a (ϕ -liotoxin-Lw1a) from *Liocheles waigiensis* [73] (Fig. 3B). The *Liocheles* toxins bind to and modulate the activity of RyR1 and RyR2 ryanodine receptors. They are reported to adopt the structural fold known as disulfide-directed hairpin (DDH) motif, a proposed evolutionary precursor of the three-disulfide ICK [73].

The putative calcins here reported are yet to be purified from the parental scorpions venoms. If they are indeed active on the RyR, they could become valuable tools for the study of these calcium channels.

3.7. La1-like peptides (La1lp)

La1-like peptides are cysteine-rich peptides that have thus far been found most frequently in the venoms of scorpions that are not toxic to mammals. La1, the first described peptide of this family, was reported as the most abundant component of the venom of the scorpion L. australasiae [74]. The La1 peptide was tested with crickets but showed no toxicity. Lately, La1-like peptides have been found in other scorpion venoms, as O. cayaporum [52], U. yashenkoi [48], *P. cavimanus* [47], *S. margerisonae* [31], *H. petersii* [50] and *S. maurus palmatus* [46]. Six

different transcripts coding for putative La1-like peptides were found for *S. jendeki* [49]. Interestingly, similar peptides have also been found in scorpions belonging to the Buthidae family: *Lychas mucronatus* and *Isometrus maculatus* [31].

The La1-like peptides are presumed to possess a single Von Willebrand factor type C domain (VWC), a structural feature shared by several eukaryotic proteins with cell biology functions [75,76], although this has not been proven [77].

Four precursors of La1-like peptides were found in the *Vaejovis* libraries: VmLa1lp1 from *V. mexicanus*, VsLa1lp1 from *V. subcristatus* and VpLa1lp1 from *V. punctatus* and ViLa1lp1 from *V. intrepidus*. They share over 70% sequence identity, and present high similarity with the other previously described La1-like peptides, as shown in Fig. 4. They all present the VWC domain with conserved 8 cysteines. The purified mature La1 was confirmed to be amidated [74]. In the sequences encoded by all the precursors, right after the positions equivalent to the last amino acid of La1, a potential amidation substrate (GK or GR) is present. We therefore assume that all La1-like peptides are amidated and that the precursors contain a propeptide (post-peptide) after the mature region, as indicated in Fig. 4.

The activity of the La1-like peptides remains to be determined. This is a relevant subject due to their seemingly wide distribution in the Scorpiones order.

3.8. Non-disulfide-bridged peptides (NDBPs)

The interest for the scorpion venom non-disulfide-bridged peptides (NDBPs) has increased in the last decade. This is due to their structural diversity and the multiple functions that they can perform, which highlights them as potential candidates for many applications, including drug development. They are 13–56 amino acids-long cationic peptides with random coil conformation in water, but in cell membrane-mimicking environments, they adopt an amphipathic α -helical structure. More than 40 scorpion NDBPs have been functionally characterized to date [17]. Among them, more than 30 display antimicrobial and cytolytic properties. They can inhibit the growth of Gram-negative and Gram-positive strains of bacteria, and are therefore considered to be the organism's first line of defense against pathogens. Some of these peptides have been isolated from scorpion venoms by biochemical techniques, but an increasing number of NDBP sequences have been derived from the precursor sequences found in cDNA libraries [78].

Since the first report of a scorpion antimicrobial peptide, hadrurin from *Hadrurus aztecus* (now known as *H. gertschi*) [20], we have found several other scorpion NDBPs, from both Buthidae and non-Buthidae species [37,40,48,79]. We have already reported two NDBPs from the *Vaejovis* genus, vejovine a peptide with 47 amino acids isolated from the venom of *V. mexicanus* and VmCT1 a short peptide derived from a cDNA library of *V. mexicanus* [37,40]. The

La1-like (4 C-C):

									aa	16 L
VsLallpl	MKHFCVVVLFGCLSLCTLFSFSKGVAEI	QVGSMSINVGRKKQDPKN	CVLYE	CLEQNNRILLSKMS	CSPQVPKRG	RNVPGPVDAPFPD	CCPTSL	CRGKQWDDE	73	100
ViLallpl	MRHLCVVALFGCLSLCTLFSFSKGVGEI	QVGSMSINVGKKMQDPKS	CVIYE	CVEQNYRILLSKMS	CSPQVPKRG	RNVPGPVDAPFPD	CCPTSL	CRGKQWDE	73	90.4
VpLa11p1	MRHLCIVVLFGCLSFCTLFSFSKGIGEI	QVGSMSINVGNKMKDPKS	CVLYQ	CLEQNNRILLSKMS	CSPQVPKRG	RNVPGPADAPFPE	CCTTSL	CRGKQWDEN	73	87.7
VmLa11p1	MRFLCVVILFGFLSLCTLFSLSKGAGEA	QVGGMSLPVGNKKQDPNS	CIVYN	CIVQNNRILLSKLS	CSPQLLKRG	KNVPGPTGAPFPG	CCPTSL	CRGKQWDK	73	71.2
Lallp-15	MKHLSDAVFFVCLSICALFSLTLCDGEI	CQVGNMAIPVGKEQPDPKG	CAKYE	CLSQSNRVLIKKVT	CASQALKRG	KSVPGPAGKRFPE	CCPTTL	CRGKQWGQ	73	57.5
HsTx1	MSHLRIAVTFLCTLFALTAGAGES	QVGGSAIPVGKTKQEL	CNLYE	CATDSNRVVLKKLT	CAEQAVKTG	RSVPGDANAPFPL	CCPTTL	CRGRQWDH	71	54.8
VenPepPc	MSHLRIAVIFLCTLFALTAGAEES	QVGGLTIPVGKTQQDR	CTLYE	CTMESNRVVLKSMI	CAEQSLRRG	KRVPAQATAPFPI	CCPTTL	CRGRQWDRPRTL	71	52.1
Lal	FGES	LIAGRFIVPLGOOVTDORD	CALYK	CVNYNKKFALETKR	CATVNLKSG	KTVPGGAGAAFPS	CCPMVT	СК	73	34.2

Fig 4. Sequence alignment of the La1-like peptides found in the Vaejovis libraries. The precursor sequences of VmLa1lp1 from *V. mexicanus*, VsLa1lp1 from *V. subcristatus* and VpLa1lp1 from *V. punctatus* are aligned to the previously reported precursors of La1lp-15 (UniProt:AGA82761) from *U. yaschenkoi*, HsTx1 (UniProt:K7WMX6) from *H. spinifer*, VenPepPc (UniProt:H2CYP1) from *P. cavimanus*, and the mature La1 (UniProt:POC5F3) from *L. australasiae*. The peptide length (aa) always refers to the confirmed (when the peptide has been isolated from the venom) or software-predicted mature peptides. The given percentages of identity (%I) correspond to the mature peptides only, and are relative to the first sequence. The predicted signal peptides are shown underlined, the propeptides in italics and bold.

doi:10.1371/journal.pone.0117188.g004

analysis of the cDNA libraries constructed for the four *Vaejovis* species revealed the presence of several new peptides. They belong to the subfamilies NDBP-2 and NDBP-4, according to the classification recently proposed by Almaaytah and Albas [17].

3.9. Short antimicrobial peptides (NDBP-4)

The members of the NDBP-4 subfamily are short antimicrobial peptides. They are derived from their much longer peptidic precursors by posttranslational processing of a signal peptide and a carboxy-terminal propeptide region (post-peptide) that flank the mature peptide. The signal peptide has 23-24 amino acids and is the most conserved region of the precursor sequence. The conserved GKR sequence delimits the mature region. It is a putative protease cleavage site, probably for a peptidylglycine alpha-amidating monooxygenase, the kind of enzyme that renders the mature peptide amidated in the carboxyl-terminus [80]. The propeptide region has 32-42 amino acids and is rich in Glu and Asp residues. The mature peptide contains only 13 amino acids, most of them hydrophobic. Though there is some sequence conservation among the mature members of the NDBP-4 family, the region of the mature sequence is the least conserved throughout the precursor sequences. The secondary structure prediction analysis shows that these peptides can fold into an amphipathic α -helix. Several mature peptides from this family have been tested for their functionality and shown to display antimicrobial and hemolytic activities. For example, IsCT (CT stands for cytotoxic), a peptide isolated from the venom of Opisthacanthus madagascariensis, has been shown to display antibacterial activity against Gram-negative and Gram-positive strains of bacteria with minimum inhibitory concentrations (MIC) in the range of 0.7-17.0 µM [81]. To date, about 24 different peptides with sequence similarities to IsCTs have been tested for activity. They have been found in all scorpion families thus far studied [17]. We have previously reported the characterization of two homologs of IsCT derived from a cDNA library of V. mexicanus: VmCT1 and VmCT2 [40]. Both peptides were able to inhibit the growth of Gram-negative and Gram-positive strains of bacteria with MICs of $5-25 \,\mu$ M. VmCT2 was mildly hemolytic at its MIC range, but showed strong hemolysis at higher concentrations. VmCT1 produced only mild (12%) hemolysis even at $50 \,\mu\text{M}$, far beyond the MIC range.

The analysis of the cDNA libraries made with four Vaejovis species revealed the presence of several precursors for peptides belonging to the NDBP-4 subfamily. From the V. mexicanus library, 12 clusters were found (9 contigs and 3 singlets); for V. intrepidus, 7 clusters (3 contigs and 4 singlets); for V. punctatus, 5 clusters (2 contigs and 3 singlets); and for V. subcristatus, 3 clusters (3 contigs). Fig. 5 shows the alignment of the sequences found in the four Vaejovis libraries with the precursors of some members of the NDBP-4 subfamily, which mature peptides have been functionally characterized. The alignment shows that the region of highest identity corresponds to the signal peptide, while the region of the mature peptide is the one with the lowest identity. All the Vaejovis propeptides show a high content of acidic amino acids, and present the conserved GKR sequence at their amino-terminus. Some sequence patterns are noticeable in the Vaejovis mature peptides that allow distinguishing three types. First, some sequences present the GIIDTV motif, which is found in at least one sequence per species. Second, some others present a tryptophan residue at position 6, which has been proposed to play an important role in the structural properties of these peptides as well as in their antibacterial activities [40]. Finally, the sequences VpCT1, ViCT1 and VsCT3, present a pair of leucines at positions 5 and 6, plus another pair at positions 12 and 13. The sequence of VpCT4 is remarkable since it has a mature peptide of 16 amino acids, being the only peptide of this subfamily with more than 13 amino acids found in the Vaejovis libraries. Note that in Fig. 5 the sequences here reported are not ordered in accordance to the percentage of identity determined



Short antimicrobial peptides (NDBP-4):

		aa	%I	aa	a %I	MW(Da)	CHG		aa	8I
VmCT1	MKTQFVILIVAVVLLQLISHSE-A	23	100	FLGALWNVAKSVF 13	3 100	1450.7	7 +1	GKRGLRNFDDLDD-TFEPEMSEADLKYLQDLLR	32	100
VmCT2	MKTQFVILIVAVVLLQLIANSE-A	23	91.3	FLSTLWNAAKSIF 13	3 69.	2 1497.7	7 +1	GKRGLRNLDNLDDDIFEPEMSEADLRYLQDLLR	33	84.8
VmCT3	MKTQFVILMMAVVLLQLIANSE-A	23	82.6	FLSTLWNAAKSIF 13	3 69.	2 1497.7	7 +1	GKRGLRNLDNLDDDIFEPEMSEADLRYLQDLLR	33	84.8
VmCT4	MKTQFVILMMAVVLLQLIANSE-A	23	82.6	FLSTLWNAAKSIF 1:	3 69.	2 1497.7	+1	GKRGLRNLDNLDD-IFEPEMSEADLRHLQDLLR	32	84.4
VmCT5	MKTQIVILVVAVLFLQLVSQSD-A	23	69.6	FLSGIIDAVKGLF 13	3 30.	8 1379.6	5 0	GKRGLKNLDQYNDLFDGEISDADVKFLQDLMR	32	59.4
VmCT6	MKTQIVILVVAVLFLQLVSQSD-A	23	69.6	FLSGIIDTVKKIF 13	3 30.	8 1480.8	3 +1	GKRGLKNLDQYNDLFDGEISDADVKFLQDLMR	32	59.4
VmCT7	MKTOIVILVVAVLFLOLISOSD-A	23	73.9	FLSGIVDAVKSIF 13	3 38.	5 1395.6	5 0	GKRGLKNLDQYNDLFDGEISDADVKFLQDLMR	32	59.4
VpCT1	MKNQFAVLLVALVLLQLFSQSE-A	23	69.6	FWSTLLSIGKSLL 13	3 30.	8 1464.7	7 +1	GKRGLRNFD-LEQMDDTYEPELSEADLRYLQDLLR	34	79.4
VpCT2	MKTQLVVLIVALVFMQLFAQSE-A	23	65.2	FWSTIWNAAKSLI 13	3 46.	2 1536.7	7 +1	GKRGLRNLDDLDQFDDSFEPELSEADLKYLEDLLR	35	80.0
VpCT3	MKTQIVILIVAVLVLQLVSQSD-A	23	73.9	FLOGIIDTVGKWL 1:	3 15.	4 1489.7	1 0	GKRGLKNLDQYNDLFDGEISDADIKFLQDLMR	32	59.4
VpCT4	MKTOFVILIVAIVILOLISOSE-A	23	87.0	LWGALLGLGSTLLSKL 10	5 18.	8 1642.0) +1	GKRGVONMDOFDDIFEPELSEADLRYLODLLR	32	75.0
VpCT5	MKTOIVILIVAVLFLOLVSOSD-A	23	73.9	FLOGIIDTVGKWL 13	3 15.	4 1489.7	7 0	GKRGLKNLDQYNDLFDGEISDADIKFLQDLMR	32	59.4
ViCT1	MKNOFAVLLVALVLLOLFSOSEGS	24	62.5	FWNTLLSVGKSLL 1:	3 38.	5 1477.7	+1	GKRGLRNFD-FDOLDDTYEPELSEADLRYLODLLR	34	82.4
ViCT2	MKTOLVVLIVALVFMOLFAOSH-A	23	60.9	FWGAVWNAAKSIL 13	3 61.	5 1462.7	7 +1	GKRGLRNLDDLDOFDDSFEPELSEADLKYLEDLLR	35	80.0
ViCT3	MKTOIVILIVAVLFLOLVSOSD-A	23	73.9	FLKGIIDTVSNWL 1:	3 15.	4 1505.7	7 0	GKRSLKNLDQYNDLFDGEISDADINFLRNLMR	32	46.9
ViCT4	MKTQLVVLIVALVFMQLFAQSH-A	23	60.9	FWGAVWDAAKSIL 13	3 53.	8 1463.7	0 1	GKRGLRNLDDLDQFDDSFEPELSEADLKYLEDLLRWI	37	75.7
ViCT5	MKTQIVILIVAVLFLQLVSQSD-A	23	73.9	FLKGIIDTVGKWL 13	3 15.	4 1489.8	3 +1	GKRGLKNLDQYNDLFDGEISDADIKYLQDLMR	32	62.5
ViCT6	MKTOIVILIVAMLFLOLVSOSD-A	23	69.6	FLKGIIDTVGKWF 13	3 23.	1 1523.8	3 +1	GKRGLKNLDOYNDLFDGEISDADIKYLODLMR	32	62.5
VsCT1	MKTOIVILIVAVLFLOLVSOSD-A	23	73.9	FLKGIIDTVSNWL 13	3 15.	4 1505.7	7 0	GKRGLKNLDQYNDLFDGEISDADIKFLKDLMR	32	56.2
VsCT2	MKTQIVILIVAVLFLQLVSQSD-A	23	73.9	FLKGIIDTVSKLF 13	3 23.	1 1480.8	3 +1	GKRGLKNLDQYNDLFDGEISDADIKYLQDLMR	32	62.5
VsCT3	MKNQFAVLLVALVLLQLFSQSE-A	23	69.6	FWSTLLSIGKSLL 13	3 30.	8 1464.7	7 +1	GKRGLRNMD-LDQLDDTYEPELSEADLRYLQDLLR	34	79.4
IsCT	MKTOFAILLVALVLFOMFAOSD-A	23	60.9	ILGKIWEGIKSLF 13	3 46.	2 1503.8	3 +1	GKRGLSDLDGLDELFDGEISKADRDFLRELMR	32	50.0
ISCT2	MKTQFAILLVALVLFQMFAQSE-A	23	65.2	IFGAIWNGIKSLF 1:	3 53.	8 1465.7	1 +1	GRRALNNDLDLDGLDELFDGEISQADVDFLKELMR	35	42.9
Hp1090	MKTQFAIFLITLVLFQMFSQSD-A	23	52.2	IFKAIWSGIKSLF 13	3 38.	5 1509.8	3 +2	GKRGLSDLDDLDESFDGEVSQADIDFLKELMQ	32	50.0
pantinin1	MKTOFVILMITVILMOMLVOTEGG	24	50.0	ILGKLWEGFKSIV 13	3 46.	2 1489.8	3 +1	GKRGLNDRDQLDDLFDSDLSDADIKLLKEMFK	32	46.9
pantinin2	MKAQFAILLITLVLFQMFSQSE-A	23	56.5	IFGAIWKGISSLL 13	3 30.	8 1404.7	7 +1	GKRGLNNLNDFDELFDGEITKADLDFMREIMK	32	40.6
pantinin3	MKTQFAILLIALVLFQLLSQSD-A	23	65.2	FLSTIWNGIKSLL 13	3 46.	2 1491.7	7 +1	GRRGLNELDNLDELFDGEISQADIDFLKELMS	32	43.8
UyCT1	MKTQLAFLAITVILMQLFAQTEAG	24	45.8	FWGKLWEGVKNAI 13	3 38.	5 1547.8	3 +1	GKRGLRNVDQIADLFDSGLSDADDLFDSGLSDADAKFMKMFM	42	37.2
UyCT3	MKNQFVLLLLAIVFLQMIFQSD-A	23	56.5	ILSAIWSGIKSLF 13	3 46.	2 1434.7	7 +1	GKRGLENMDKFDELFDGDLSEADLDFLKELMR	32	53.1
UyCT5	MKNQFAILLLAVVFLQLISQSD-A	23	69.6	IWSAIWSGIKGLL 13	3 23.	1 1443.7	7 +1	GKRGLKNADRLDELFDGDISDADLDFLRELMR	32	53.1
StCT2	MKTQFAVLIISMILMQMLVQTEAG	24	45.8	FWGKLWEGVKSAI 1:	3 46.	2 1520.7	7 +1	GKRSLRNQDQFDNMFDSDLSDADLKLLDDLFD	32	53.1
BmKn2	MKSQTFFLLFLVVLLLAISQSE-A	23	56.5	FIGAIANLLSKIF 1	3 38.	5 1406.7	7 +1	GKRSMRDMDTMKYLYDPSLSAADLKTLQKLMENY	34	41.2
Meucin-13	MKSQTFFLLFLVVFLLAITQSE-A	23	47.8	IFGAIAGLLKNIF 13	3 30.	8 1376.7	+1	GKRSLRDMDTMKYLYDPSLSAADLKTLQKLMENY	34	44.1
Consensus	MKTQFVILIVAVVLLQLFSQSE-A	23	91.3	FLXAIWDGVKSLF 1	3 53.	8		GKRGLRNLDQLDDLFDGEISDADLKFLQDLMR	32	71.9
	Signal peptides			Mature peptides				Propeptides		
Fig 5 Mu	Itiple alignment of the se	an	ences	corresponding to	the n	entidic r	recu	rsors belonging to the NDBP-4 subfamily foun	nd ir	n tha l

Fig 5. Multiple alignment of the sequences corresponding to the peptidic precursors belonging to the NDBP-4 subfamily found in the Vaejovis libraries. Separate alignments are shown for the signal peptides, the mature peptides and the propeptides, with the number of residues (aa) and the percentage of identity (%I) indicated for each segment. For the mature peptide the prominent sequence features are highlighted with colors and in bold typeface (see text) and the theoretical molecular weight (MW) and net charge (CHG) are also given. In the propeptide region, the post-translational cleavage and amidation substrate sequence is underlined. For comparison, the sequences of IsCT (UniProt:Q8MMJ7.1), IsCT2 (UniProt:Q8MTX2.1), UyCT1 (UniProt: L0GCV8.1), UyCT3 (UniProt:L0GCI6.1), UyCT5 (UniProt:L0GAZ8.1), Hp1090 (UniProt:P0DJ02.1), Pantinin 1 (UniProt:R4JNJ5), Pantinin 2 (UniProt: R4JQZ0), Pantinin 3 (UniProt:R4JJN6), Meucin-13 (UniProt:E4VP07.1), BmKn2(UniProt:Q6QN2.1) and StCT2 (UniProt:P0DJ04.1) were used.

doi:10.1371/journal.pone.0117188.g005

for any particular region of the precursors. They are instead sequentially ordered by species, so the diversity of the above mentioned sequence landmarks within each species is more evident.

3.10. Long-chain multifunctional peptides (NDBP-2)

The NDBP-2 subfamily groups long-chain multifunctional peptides. Members of this subfamily have been shown to display antimicrobial, bradykinin-potentiating, insecticidal and anticancer activities. To the NDBP-2 subfamily belong peptides like hadrurin, parabutoporin, pandinin-1 and the opistoporins [17]. They are also produced as larger precursors that suffer posttranslational processing. The structure of these precursors is similar to that of the abovedescribed NDBP-4 subfamily. Sequentially, the amino-terminal signal peptide is followed by the mature peptide and then, a propeptide. Signal peptides typically contain 22 amino acids, mature peptides 39-50, and propeptides 13-17. Mature peptides are rich in Lys and Arg residues, so they are usually highly charged cationic peptides. It was recently reported that vejovine, from V. mexicanus, is able to inhibit the growth of multiresistant Gram-negative bacterial strains from clinical isolates (Pseudomonas aeruginosa, Klebsiella pneumoniae, Escherichia coli, Enterobacter cloacae y Acinetobacter baumanii) with MICs of 4.4–50.0 µM. Here we report four new peptides belonging to the NDBP-2 subfamily that were derived from sequences found in the Vaejovis cDNA libraries. Fig. 6 shows the alignment of the precursors of these peptides with vejovine and other members of the NDBP-2 subfamily. Two distinct sequence types can be distinguished in the Vaejovis peptides. On one side, vejovine and VpVlp1 (Vlp stands for vejovine-like peptide) are larger peptides that show a high percentage of identity, with only 5



Long-chain multifunctional peptides (NDBP-2)

		aa	¥1	WM(Da)	CHO
VpVlp1	MNAKTIFAIFLIGMLVTEQAEAG-IWGTIKKYASKAWNSDIGKSLRNKAAGAINKFVADKIGVTPSQAASMTLDQIVDAMYYD	47	100	4920.6	+6
VpVlp2	MNGKALLVVFLVTMLVMEPTEAG-LFSWVKKAWNSGVGKSLRKQAVKVAKNYVANKLGGTPE EAGAMPFDEFMDVLHY-	39	40.4	4217.9	+7
ViVlp1	MNGKALLVVFLVTMLVMEPTEAG-IFSWVKKAWNSGVGKSLRKQAVKAAKNYVANKLGGTPEEAGAMPFDEFMDVLHYN	39	44.7	4189.8	+7
VmVlp1	MNRKVLLVVFMVTMLVMEPTEAG-LFSWAKKAWNSNIGKSLRKQAVKAAKSYVANKLGGTPEEAGAMPFDEFMDVLHYN	39	44.7	4205.8	+7
Vejovine	MNAKTLFVVFLIGMLVTEQVEAG-IWSSIKNLASKAWNSDIGQSLRNKAAGAINKFVADKIGVTPSQAASMTLDEIVDAMYYD	47	89.4	4872.5	+4
Opistoporin2	GKVWDWIKSTAKKLWNSEPVKELKNTALNAAKNFVAEKIGATPSGKVWDWIKSTAKKLWNSEPVKELKNTALNAAKNFVAEKIGATPS	44	47.9	4870.6	+4
Opistoporin1	MNRKLLFVTLMVTMLVMQPSEGGKVWDWIKSTAKKLWNSEPVKELKNTALNAAKNLVAEKIGATPSEAGQMPFDEFMDILYE-	44	45.8	4836.6	+4
Hadrurin	IN-WVANKLGVSP-QAA	41	42.6	4436.1	+5
Pandinin1	GKVWDWIKSAAKKIWSSEPVSQLKGQVLNAAKNYVAEKIGATPTGKVWDWIKSAAKKIWSSEPVSQLKGQVLNAAKNYVAEKIGATPT	44	35.4	4799.5	+4
Parabutoporin	FKLGSFLKKAWKSKLAKKLRAKGKEMLKDYAKGLLEGGSEEVPGQ	45	22.9	4994.9	+7
BmKbpp	MNKKTLLVIFFVTMLIVDEVNSFRFGSFLKKVWKSKLAKKLRSKGKOLLKDYANKVLNGPEEEAAAPAERRR	47	22.0	5321.2	+7

Fig 6. Multiple alignment of the sequences corresponding to the precursors of the peptides belonging to the NDBP-2 subfamily found in the *Vaejovis* libraries. The peptide length (aa) refers to the confirmed (when the peptide has been isolated from the venom) or software-predicted mature peptides. The identity (%I) is relative to the first sequence of the alignment, and considers only the mature peptide region. When present, the signal peptides are shown underlined and the propeptides are in italics and bold typeface. The theoretical molecular weight (MW) and net charge (CHG) are also given for each mature peptide. Fully conserved residues are highlighted with color and bold. For comparison purposes, the sequences of vejovine (UniProt: ADZ24463.1), hadrurin (UniProt:P82656.1), opistoporin-1 (UniProt:P83313.2), opistoporin-2 (UniProt:P83314.1), pandinin1 (UniProt:P83239.1), parabutoporin (UniProt:P83312.1), and BmKbpp (UniProt:Q9Y0X4.1) were used.

doi:10.1371/journal.pone.0117188.g006

differences in the 47-amino acids sequence. On the other, VpVlp2, ViVlp1 and VmVlp1 are shorter, 39-amino acids peptides, that are less related to vejovine, but display a high percentage of identity (89.7–88.5%, not shown in Fig. 6) with each other. It is remarkable that these last peptides are the shortest members of the NDBP-2 family, while at the same time they have a large positive net charge of +7, found only in larger peptides, like parabutoporin and BmKbpp. They thus represent the NDBPs with the largest known density of positive charges. Peptides with increased positive charge have been shown to be more effective as antimicrobials (but also more hemolytic) [82].

Concluding Remarks

This communication reports the first approach to the transcriptome analysis of scorpion species belonging to the Vaejovidae family. Many new unknown sequences were identified, but in general, these results confirm the previously observed predominance of non-disulfide bound peptides in species of scorpions distinct from those classical cysteine-rich channel-acting toxins found in the venoms of Buthidae species. In addition, it also shows that the non-Buthidae species have a rich diversity of components present in their venoms. Some of the new sequences are currently being assayed to assess their biological activity. Some of these peptides belong to families of previously characterized scorpion venom components that have shown potential application in fields of medical relevance, e.g. the development of new antibiotics and antimalarial drugs.

Acknowledgments

The authors acknowledge the technical assistance provided by MSc. Timoteo Olamendi Portugal with DNA sequencing, and the computer and bioinformatics support received from MSc. Juan Manuel Hurtado Ramírez and Eng. Roberto Rodríguez Bahena.

Author Contributions

Conceived and designed the experiments: VQH BB LDP EO. Performed the experiments: VQH SRC MTRG LLVV. Analyzed the data: VQH BB LDP EO. Contributed reagents/materials/analysis tools: LDP. Wrote the paper: VQH SRC MTRG LLVV BB LDP EO.

References

- 1. Dunlop JA, Selden PA (2009) Calibrating the chelicerate clock: a paleontological reply to Jeyaprakash and Hoy. Exp Appl Acarol 48: 183–197. doi: 10.1007/s10493-009-9247-1 PMID: 19199056
- Jeyaprakash A, Hoy MA (2009) First divergence time estimate of spiders, scorpions, mites and ticks (subphylum: Chelicerata) inferred from mitochondrial phylogeny. Exp Appl Acarol 47: 1–18. doi: <u>10.</u> <u>1007/s10493-008-9203-5</u> PMID: <u>18931924</u>
- Fet V, Sissom WD, Lowe GJ, Braunwalder ME (2000) Catalog of the Scorpions of the World (1758– 1998). New York, NY: New York Entomological Society. PMID: <u>25506959</u>
- Possani LD, Becerril B, Delepierre M, Tytgat J (1999) Scorpion toxins specific for Na+-channels. Eur J Biochem 264: 287–300. PMID: <u>10491073</u>
- Sollod BL, Wilson D, Zhaxybayeva O, Gogarten JP, Drinkwater R, et al. (2005) Were arachnids the first to use combinatorial peptide libraries? Peptides 26: 131–139. PMID: 15626513
- Chippaux JP, Goyffon M (2008) Epidemiology of scorpionism: A global appraisal. Acta Tropica 107: 71–79. doi: <u>10.1016/j.actatropica.2008.05.021</u> PMID: <u>18579104</u>
- Rodriguez de la Vega RC, Schwartz EF, Possani LD (2010) Mining on scorpion venom biodiversity. Toxicon 56: 1155–1161. doi: <u>10.1016/j.toxicon.2009.11.010</u> PMID: <u>19931296</u>
- Chen T, Walker B, Zhou M, Shaw C (2005) Molecular cloning of a novel putative potassium channelblocking neurotoxin from the venom of the North African scorpion, Androctonus amoreuxi. Peptides 26: 731–736. PMID: <u>15808902</u>
- Chen TB, Folan R, Kwok HF, O'Kane EJ, Bjourson AJ, et al. (2003) Isolation of scorpion (Androctonus amoreuxi) putative alpha neurotoxins and parallel cloning of their respective cDNAs from a single sample of venom. Regulatory Peptides 115: 115–121. PMID: 12972326
- Schwartz EF, Diego-Garcia E, Rodriguez de la Vega RC, Possani LD (2007) Transcriptome analysis of the venom gland of the Mexican scorpion Hadrurus gertschi (Arachnida: Scorpiones). BMC Genomics 8: 119. PMID: 17506894
- Zeng XC, Corzo G, Hahin R (2005) Scorpion venom peptides without disulfide bridges. IUBMB Life 57: 13–21. PMID: <u>16036557</u>
- Zeng XC, Li WX, Peng F, Zhu ZH (2000) Cloning and characterization of a novel cDNA sequence encoding the precursor of a novel venom peptide (BmKbpp) related to a bradykinin-potentiating peptide from Chinese scorpion Buthus martensii Karsch. IUBMB Life 49: 207–210. PMID: <u>10868911</u>
- Zeng XC, Wang SX, Zhu Y, Zhu SY, Li WX (2004) Identification and functional characterization of novel scorpion venom peptides with no disulfide bridge from Buthus martensii Karsch. Peptides 25: 143– 150. PMID: 15062994
- Chippaux JP (2012) Emerging options for the management of scorpion stings. Drug Des Devel Ther 6: 165–173. doi: 10.2147/DDDT.S24754 PMID: 22826633
- Tuuri RE, Reynolds S (2011) Scorpion envenomation and antivenom therapy. Pediatr Emerg Care 27: 667–672; quiz 673–665. doi: 10.1097/PEC.0b013e3182228dfa PMID: 21730810
- Quintero-Hernandez V, Jimenez-Vargas JM, Gurrola GB, Valdivia HH, Possani LD (2013) Scorpion venom components that affect ion-channels function. Toxicon 76: 328–342. doi: <u>10.1016/j.toxicon.</u> <u>2013.07.012</u> PMID: <u>23891887</u>
- Almaaytah A, Albalas Q (2014) Scorpion venom peptides with no disulfide bridges: a review. Peptides 51: 35–45. doi: <u>10.1016/j.peptides.2013.10.021</u> PMID: <u>24184590</u>
- Gao B, Sherman P, Luo L, Bowie J, Zhu SY (2009) Structural and functional characterization of two genetically related meucin peptides highlights evolutionary divergence and convergence in antimicrobial peptides. Faseb Journal 23: 1230–1245. doi: 10.1096/fj.08-122317 PMID: 19088182
- Kuhn-Nentwig L (2003) Antimicrobial and cytolytic peptides of venomous arthropods. Cell Mol Life Sci 60: 2651–2668. PMID: <u>14685689</u>
- Torres-Larios A, Gurrola GB, Zamudio FZ, Possani LD (2000) Hadrurin, a new antimicrobial peptide from the venom of the scorpion Hadrurus aztecus. Eur J Biochem 267: 5023–5031. PMID: <u>10931184</u>
- Ortiz E, Gurrola GB, Schwartz EF, Possani LD (2015) Scorpion venom components as potential candidates for drug development. Toxicon 93: 125–135.
- 22. Ferreira LA, Alves EW, Henriques OB (1993) Peptide T, a novel bradykinin potentiator isolated from Tityus serrulatus scorpion venom. Toxicon 31: 941–947. PMID: <u>8212046</u>
- Moerman L, Bosteels S, Noppe W, Willems J, Clynen E, et al. (2002) Antibacterial and antifungal properties of alpha-helical, cationic peptides in the venom of scorpions from southern Africa. Eur J Biochem 269: 4799–4810. PMID: <u>12354111</u>

- Dai C, Ma Y, Zhao Z, Zhao R, Wang Q, et al. (2008) Mucroporin, the first cationic host defense peptide from the venom of Lychas mucronatus. Antimicrob Agents Chemother 52: 3967–3972. doi: <u>10.1128/</u> <u>AAC.00542-08</u> PMID: <u>18779362</u>
- Fletcher PL Jr, Fletcher MD, Weninger K, Anderson TE, Martin BM (2010) Vesicle-associated membrane protein (VAMP) cleavage by a new metalloprotease from the Brazilian scorpion Tityus serrulatus. J Biol Chem 285: 7405–7416. doi: 10.1074/jbc.M109.028365 PMID: 20026600
- Ortiz E, Rendon-Anaya M, Rego SC, Schwartz EF, Possani LD (2014) Antarease-like Zn-metalloproteases are ubiquitous in the venom of different scorpion genera. Biochim Biophys Acta 1840: 1738– 1746. doi: <u>10.1016/j.bbagen.2013.12.012</u> PMID: <u>24361608</u>
- Soudani N, Gharbi-Chihi J, Srairi-Abid N, Kaabi H, Margotat A, et al. (2005) Identification of second lipolysis activating protein from scorpion Buthus occitanus tunetanus. Arch Inst Pasteur Tunis 82: 39–46. PMID: <u>16929753</u>
- Valdez-Velazquez LL, Quintero-Hernandez V, Romero-Gutierrez MT, Coronas FI, Possani LD (2013) Mass fingerprinting of the venom and transcriptome of venom gland of scorpion Centruroides tecomanus. PLoS One 8: e66486. doi: <u>10.1371/journal.pone.0066486</u> PMID: <u>23840487</u>
- 29. Nascimento DG, Rates B, Santos DM, Verano-Braga T, Barbosa-Silva A, et al. (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. PMID: <u>16551474</u>
- **30.** Soleglad ME, Fet V (2003) High-level systematics and phylogeny of the extant scorpions (Scorpiones: Orthosterni). Euscorpius 11: 175.
- Ma Y, He Y, Zhao R, Wu Y, Li W, et al. (2012) Extreme diversity of scorpion venom peptides and proteins revealed by transcriptomic analysis: implication for proteome evolution of scorpion venom arsenal. J Proteomics 75: 1563–1576. doi: <u>10.1016/j.jprot.2011.11.029</u> PMID: <u>22155128</u>
- Sunagar K, Undheim EAB, Chan AHC, Koludarov I, Munoz-Gomez SA, et al. (2013) Evolution Stings: The Origin and Diversification of Scorpion Toxin Peptide Scaffolds. Toxins 5: 2456–2487. doi: <u>10.</u> <u>3390/toxins5122456</u> PMID: 24351712
- 33. Hoffmann CC (1931) Monografías para la entomología médica de México. Monografía núm. 2. Los scorpiones de México. Primera parte. Diplocentridae, Chactidae, Vejovidae. Anal Inst Biol Universidad Nacional Autónoma de México 2: 117.
- Santibañez-Lopez C, Sissom WD (2010) A new species of the Vaejovis eusthenura group in Oaxaca, Mexico (Scorpiones: Vaejovidae). Zootaxa 2493: 49–58. doi: <u>10.1109/IEMBS.2010.5626618</u> PMID: <u>21096168</u>
- Sissom WD (2010) Family Vaejovidae Thorell 1876. In: Fet V, Sissom WD, Lowe GJ, Braunwalder ME, editors. Catalog of the Scorpions of the world (1758–1998). New York, NY: New York Entomological Society. pp. 503–553.
- Smith JJ, Jones A, Alewood PF (2012) Mass landscapes of seven scorpion species: The first analyses of Australian species with 1,5-DAN matrix. J Venom Res 3: 7–14. PMID: 23236582
- Hernandez-Aponte CA, Silva-Sanchez J, Quintero-Hernandez V, Rodriguez-Romero A, Balderas C, et al. (2011) Vejovine, a new antibiotic from the scorpion venom of Vaejovis mexicanus. Toxicon 57: 84–92. doi: 10.1016/j.toxicon.2010.10.008 PMID: 20969885
- Gurrola GB, Hernandez-Lopez RA, Rodriguez de la Vega RC, Varga Z, Batista CV, et al. (2012) Structure, function, and chemical synthesis of Vaejovis mexicanus peptide 24: a novel potent blocker of Kv1.3 potassium channels of human T lymphocytes. Biochemistry 51: 4049–4061. doi: <u>10.1021/bi300060n PMID: 22540187</u>
- Varga Z, Gurrola-Briones G, Papp F, Rodriguez de la Vega RC, Pedraza-Alva G, et al. (2012) Vm24, a natural immunosuppressive peptide, potently and selectively blocks Kv1.3 potassium channels of human T cells. Mol Pharmacol 82: 372–382. doi: 10.1124/mol.112.078006 PMID: 22622363
- Ramirez-Carreto S, Quintero-Hernandez V, Jimenez-Vargas JM, Corzo G, Possani LD, et al. (2012) Gene cloning and functional characterization of four novel antimicrobial-like peptides from scorpions of the family Vaejovidae. Peptides 34: 290–295. doi: 10.1016/j.peptides.2012.02.002 PMID: 22342498
- **41.** Stahnke HL (1974) Revision and keys to the higher categories of Vejovidae. Journal of Arachnology 1: 35.
- Sambrook J, Russell D (2001) Molecular Cloning: A Laboratory Manual: Cold Spring Harbor Laboratory Press. PMID: <u>25506954</u>
- **43.** Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8: 175–185. PMID: <u>9521921</u>
- Lazo GR, Tong J, Miller R, Hsia C, Rausch C, et al. (2001) Software scripts for quality checking of highthroughput nucleic acid sequencers. Biotechniques 30: 1300–1305. PMID: <u>11414222</u>

- Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res 9: 868–877. PMID: 10508846
- Abdel-Rahman MA, Quintero-Hernandez V, Possani LD (2013) Venom proteomic and venomous glands transcriptomic analysis of the Egyptian scorpion Scorpio maurus palmatus (Arachnida: Scorpionidae). Toxicon 74: 193–207. doi: 10.1016/j.toxicon.2013.08.064 PMID: 23998939
- Diego-Garcia E, Peigneur S, Clynen E, Marien T, Czech L, et al. (2012) Molecular diversity of the telson and venom components from Pandinus cavimanus (Scorpionidae Latreille 1802): transcriptome, venomics and function. Proteomics 12: 313–328. doi: <u>10.1002/pmic.201100409</u> PMID: <u>22121013</u>
- Luna-Ramirez K, Quintero-Hernandez V, Vargas-Jaimes L, Batista CV, Winkel KD, et al. (2013) Characterization of the venom from the Australian scorpion Urodacus yaschenkoi: Molecular mass analysis of components, cDNA sequences and peptides with antimicrobial activity. Toxicon 63: 44–54. doi: <u>10.1016/j.toxicon.2012.11.017</u> PMID: <u>23182832</u>
- 49. Ma Y, Zhao R, He Y, Li S, Liu J, et al. (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. doi: 10.1186/1471-2164-10-290 PMID: 19570192
- Ma Y, Zhao Y, Zhao R, Zhang W, He Y, et al. (2010) Molecular diversity of toxic components from the scorpion Heterometrus petersii venom revealed by proteomic and transcriptome analysis. Proteomics 10: 2471–2485. doi: 10.1002/pmic.200900763 PMID: 20443192
- Roeding F, Borner J, Kube M, Klages S, Reinhardt R, et al. (2009) A 454 sequencing approach for large scale phylogenomic analysis of the common emperor scorpion (Pandinus imperator). Mol Phylogenet Evol 53: 826–834. doi: 10.1016/j.ympev.2009.08.014 PMID: 19695333
- 52. Silva EC, Camargos TS, Maranhao AQ, Silva-Pereira I, Silva LP, et al. (2009) Cloning and characterization of cDNA sequences encoding for new venom peptides of the Brazilian scorpion Opisthacanthus cayaporum. Toxicon 54: 252–261. doi: <u>10.1016/j.toxicon.2009.04.010</u> PMID: <u>19379768</u>
- Rodriguez de la Vega RC, Possani LD (2005) Overview of scorpion toxins specific for Na+ channels and related peptides: biodiversity, structure-function relationships and evolution. Toxicon 46: 831–844. PMID: 16274721
- Kozminsky-Atias A, Bar-Shalom A, Mishmar D, Zilberberg N (2008) Assembling an arsenal, the scorpion way. BMC Evol Biol 8: 333. doi: <u>10.1186/1471-2148-8-333</u> PMID: <u>19087317</u>
- 55. Rendon-Anaya M, Delaye L, Possani LD, Herrera-Estrella A (2012) Global transcriptome analysis of the scorpion Centruroides noxius: new toxin families and evolutionary insights from an ancestral scorpion species. PLoS One 7: e43331. doi: 10.1371/journal.pone.0043331 PMID: 22912855
- Dehesa-Davila M, Possani LD (1994) Scorpionism and Serotherapy in Mexico. Toxicon 32: 1015– 1018. PMID: 7801335
- Tytgat J, Chandy KG, Garcia ML, Gutman GA, Martin-Eauclaire MF, et al. (1999) A unified nomenclature for short-chain peptides isolated from scorpion venoms: alpha-KTx molecular subfamilies. Trends Pharmacol Sci 20: 444–447. PMID: <u>10542442</u>
- Rodriguez de la Vega RC, Possani LD (2004) Current views on scorpion toxins specific for K+-channels. Toxicon 43: 865–875. PMID: <u>15208019</u>
- Beeton C, Wulff H, Standifer NE, Azam P, Mullen KM, et al. (2006) Kv1.3 channels are a therapeutic target for T cell-mediated autoimmune diseases. Proc Natl Acad Sci U S A 103: 17414–17419. PMID: 17088564
- Panyi G, Possani LD, Rodriguez de la Vega RC, Gaspar R, Varga Z (2006) K+ channel blockers: novel tools to inhibit T cell activation leading to specific immunosuppression. Curr Pharm Des 12: 2199– 2220. PMID: <u>16787250</u>
- Conde R, Zamudio FZ, Rodriguez MH, Possani LD (2000) Scorpine, an anti-malaria and anti-bacterial agent purified from scorpion venom. Febs Letters 471: 165–168. PMID: <u>10767415</u>
- **62.** Zhu S, Tytgat J (2004) The scorpine family of defensins: gene structure, alternative polyadenylation and fold recognition. Cellular and Molecular Life Sciences 61: 1751–1763. PMID: <u>15241551</u>
- Uawonggul N, Thammasirirak S, Chaveerach A, Arkaravichien T, Bunyatratchata W, et al. (2007) Purification and characterization of Heteroscorpine-1 (HS-1) toxin from Heterometrus laoticus scorpion venom. Toxicon 49: 19–29. PMID: <u>17056081</u>
- 64. Diego-Garcia E, Abdel-Mottaleb Y, Schwartz EF, de la Vega RCR, Tytgat J, et al. (2008) Cytolytic and K+ channel blocking activities of beta-KTx and scorpine-like peptides purified from scorpion venoms. Cellular and Molecular Life Sciences 65: 187–200. PMID: <u>18030427</u>
- 65. Darbon H (1999) Animal toxins and ion channels. J Soc Biol 193: 445–450. PMID: 10783702
- 66. Schwartz EF, Capes EM, Diego-Garcia E, Zamudio FZ, Fuentes O, et al. (2009) Characterization of hadrucalcin, a peptide from Hadrurus gertschi scorpion venom with pharmacological activity on

ryanodine receptors. Br J Pharmacol 157: 392–403. doi: <u>10.1111/j.1476-5381.2009.00147.x</u> PMID: <u>19389159</u>

- Valdivia HH, Kirby MS, Lederer WJ, Coronado R (1992) Scorpion toxins targeted against the sarcoplasmic reticulum Ca(2+)-release channel of skeletal and cardiac muscle. Proc Natl Acad Sci U S A 89: 12185–12189. PMID: <u>1334561</u>
- Valdivia HH, Fuentes O, el-Hayek R, Morrissette J, Coronado R (1991) Activation of the ryanodine receptor Ca2+ release channel of sarcoplasmic reticulum by a novel scorpion venom. J Biol Chem 266: 19135–19138. PMID: <u>1655775</u>
- Morrissette J, Beurg M, Sukhareva M, Coronado R (1996) Purification and characterization of ryanotoxin, a peptide with actions similar to those of ryanodine. Biophys J 71: 707–721. PMID: <u>8842209</u>
- Ji YH, Liu Y, Xu K, Ohishi T, Mochizuki T, et al. (1997) Amino acid sequence of BmK AS, a novel polypeptide activator of ryanodine receptor on skeletal muscle. Chinese Science Bulletin 42: 952–956.
- Fajloun Z, Kharrat R, Chen L, Lecomte C, Di Luccio E, et al. (2000) Chemical synthesis and characterization of maurocalcine, a scorpion toxin that activates Ca2+ release channel/ryanodine receptors. Febs Letters 469: 179–185. PMID: 10713267
- 72. Matsushita N, Miyashita M, Sakai A, Nakagawa Y, Miyagawa H (2007) Purification and characterization of a novel short-chain insecticidal toxin with two disulfide bridges from the venom of the scorpion Lio-cheles australasiae. Toxicon 50: 861–867. PMID: <u>17681581</u>
- 73. Smith JJ, Vetter I, Lewis RJ, Peigneur S, Tytgat J, et al. (2013) Multiple actions of phi-LITX-Lw1a on ryanodine receptors reveal a functional link between scorpion DDH and ICK toxins. Proceedings of the National Academy of Sciences of the United States of America 110: 8906–8911. doi: <u>10.1073/pnas.</u> 1214062110 PMID: 23671114
- Miyashita M, Otsuki J, Hanai Y, Nakagawa Y, Miyagawa H (2007) Characterization of peptide components in the venom of the scorpion Liocheles australasiae (Hemiscorpiidae). Toxicon 50: 428–437. PMID: 17559900
- Sheldon TJ, Miguel-Aliaga I, Gould AP, Taylor WR, Conklin D (2007) A novel family of single VWC-domain proteins in invertebrates. FEBS Lett 581: 5268–5274. PMID: <u>18028914</u>
- 76. Zhang JL, Huang Y, Qiu LY, Nickel J, Sebald W (2007) von Willebrand factor type C domain-containing proteins regulate bone morphogenetic protein signaling through different recognition mechanisms. Journal of Biological Chemistry 282: 20002–20014. PMID: 17483092
- 77. Kamachi S, Nagao J, Miyashita M, Nakagawa Y, Miyagawa H, et al. (2014) Crystallization and preliminary X-ray diffraction studies of La1 from Liocheles australasiae. Acta Crystallogr F Struct Biol Commun 70: 915–917. doi: 10.1107/S2053230X14010589 PMID: 25005088
- Quintero-Hernandez V, Ortiz E, Rendon-Anaya M, Schwartz EF, Becerril B, et al. (2011) Scorpion and spider venom peptides: gene cloning and peptide expression. Toxicon 58: 644–663. doi: <u>10.1016/j.</u> toxicon.2011.09.015 PMID: 21978889
- 79. Garcia F, Villegas E, Espino-Solis GP, Rodriguez A, Paniagua-Solis JF, et al. (2013) Antimicrobial peptides from arachnid venoms and their microbicidal activity in the presence of commercial antibiotics. Journal of Antibiotics 66: 3–10. doi: 10.1038/ja.2012.87 PMID: 23093034
- Becerril B, Vazquez A, Garcia C, Corona M, Bolivar F, et al. (1993) Cloning and characterization of cDNAs that code for Na(+)-channel-blocking toxins of the scorpion Centruroides noxius Hoffmann. Gene 128: 165–171. PMID: <u>8390386</u>
- Dai L, Yasuda A, Naoki H, Corzo G, Andriantsiferana M, et al. (2001) IsCT, a novel cytotoxic linear peptide from scorpion Opisthacanthus madagascariensis. Biochemical and Biophysical Research Communications 286: 820–825. PMID: <u>11520071</u>
- Cao L, Dai C, Li Z, Fan Z, Song Y, et al. (2012) Antibacterial activity and mechanism of a scorpion venom peptide derivative in vitro and in vivo. PLoS One 7: e40135. doi: <u>10.1371/journal.pone.0040135</u> PMID: <u>22792229</u>