

RESEARCH

Open Access



Deep sequencing analysis of toad *Rhinella schneideri* skin glands and partial biochemical characterization of its cutaneous secretion

Priscila Yumi Tanaka Shibao¹, Camila Takeno Cologna¹, Romualdo Morandi-Filho², Gisele Adriano Wiesel¹, Patricia Tiemi Fujimura², Carlos Ueira-Vieira² and Eliane Candiani Arantes^{2,3*} 

Abstract

Background: Animal poisons and venoms are sources of biomolecules naturally selected. *Rhinella schneideri* toads are widespread in the whole Brazilian territory and they have poison glands and mucous gland. Recently, protein from toads' secretion has gaining attention. Frog skin is widely known to present great number of host defense peptides and we hypothesize toads present them as well. In this study, we used a RNA-seq analysis from *R. schneideri* skin and biochemical tests with the gland secretion to unravel its protein molecules.

Methods: Total RNA from the toad skin was extracted using TRizol reagent, sequenced in duplicate using Illumina HiSeq2500 in paired end analysis. The raw reads were trimmed and de novo assembled using Trinity. The resulting sequences were submitted to functional annotation against non-redundant NCBI database and Database of Anuran Defense Peptide. Furthermore, we performed caseinolytic activity test to assess the presence of serine and metalloproteases in skin secretion and it was fractionated by fast liquid protein chromatography using a reverse-phase column. The fractions were partially sequenced by Edman's degradation.

Results: We were able to identify several classes of antimicrobial peptides, such as buforins, peroniins and brevinins, as well as PLA₂, lectins and galectins, combining protein sequencing and RNA-seq analysis for the first time. In addition, we could isolate a PLA₂ from the skin secretion and infer the presence of serine proteases in cutaneous secretion.

Conclusions: We identified novel toxins and proteins from *R. schneideri* mucous glands. Besides, this is a pioneer study that presented the in depth characterization of protein molecules richness from this toad secretion. The results obtained herein showed evidence of novel AMP and enzymes that need to be further explored.

Keywords: RNA-seq, *Rhinella schneideri*, Toad secretion, Transcriptome, Illumina, Cutaneous secretion, Skin secretion, Toad protein

* Correspondence: ecabraga@fcrp.usp.br

²Laboratory of genetics – LABGEN, Institute of Genetics and Biochemistry, Campus Umuarama, Federal University of Uberlândia, Avenida Pará, Uberlândia, MG 1720, Brazil

³Department of Physics and Chemistry, University of São Paulo, School of Pharmaceutical Sciences of Ribeirão Preto, Av. do Café s/n°, Monte Alegre, Ribeirão Preto, SP 14040-903, Brazil

Full list of author information is available at the end of the article



Background

Animal's and microorganisms' secretions, as well as plant extracts, have been used as folk medicine since the dawn of the humanity [1]. Therefore, molecules found in poisons and venoms are interesting, once they were selected by evolution to act in their molecular targets with high specificity [1, 2]. Such molecules can be used for feeding (predation), defense or even to dispose advantage in inter and intra-specific competition [3, 4]. As an example, the glandular secretory product from toads *Bufo melanostictus Schneider* and *Bufo Bufo gargarizans cantor*, known as Chan Su, is used as medicine in treatment for several physiological disturbances [5].

Rhinella schneideri toads are widely found in South America territory: Paraguay, Bolivia, Argentina, Uruguay and Brazil. Regarding the Brazilian territory, they are found especially in cerrado. These toads have shown remarkable adaptations skills and live in urban areas as well as rural [6].

These toads present two types of glands: granulous or parotoid and mucous glands. The first one is responsible for the animal protection against predators and are located in the postorbital region of the animal's body; they can look bigger when the animal feels endangered due to body's inflation, and acts as airbags against predator bites [7, 8]. The secretions are composed mainly by biogenic amines and steroids, as bufodienolides and bufotoxines, but they also produce proteins and glicoconjugate molecules [9, 10]. Although it was previously believed that this poison presented only few or even no protein, lately it was revealed this secretion possess up to 30% of weight in proteins, but there is a lack of data to full assess them [11, 12]. *Rhinella schneideri* parotoid gland poison has shown activity against different human cancer cells proliferation [13], activate human complement system [14] and inhibit chymotrypsin [15]. Protein components present anti-inflammatory, anti-nociceptive and toxic activities in mice [16]. Similarly, protein from *Bufo bufo* parotoid secretions are likely to play a role in cardiotoxic effects [17] and *B. bufo*, *B. verrucosissimus* and *Bufotes variabilis* both parotoid and skin secretion protein are capable of inhibiting different gram-negative bacteria and cytotoxic effects on different types of cells [18].

Anuran cutaneous gland secretions are widely known to contain several classes of antimicrobial peptides (AMP) and function as the first barrier against microorganisms. Although frogs' mucous gland secretions are broadly studied and there are more than 40 classes of AMP reported up to date [19], there is no survey regarding *R. schneideri* putative protein and AMPs in the cutaneous secretions. The "omics" technologies are powerful tools to overcome this issue.

Transcriptomics, one of the "omics techniques", is one powerful approach to unravel peptides and protein in a holistic manner. Currently, RNA-seq is the state-of-art technique used to predict all protein molecules that can be produced by a specific issue with the greatest outcome of information, thus making possible the discovery of minor toxins that could not be detected through traditional techniques due to their low abundance in the final secretion [20]. This approach was used to unravel frogs AMP and adaptations [21, 22] and immune system [23]. However, there is no transcriptomic information regarding toads' skin peptides and proteins.

In order to overcome this lack of information, we constructed a RNA-seq transcriptome from skin of an individual *R. schneideri* toad. The transcriptome was sequenced in duplicate using Illumina HiSeq 2500, the reads were treated and the contigs were de novo assembled with the aid of Trinity. The results were annotated against non-redundant (nr) NCBI database and enriched with database of anuran defense peptides (DADP). Thus, the cutaneous secretion from the same toad, milked prior its death, was used to carry out biochemical analysis, as assessing its protein profile by SDS-PAGE, RP-FPLC fractionation in C18 column, peptide and protein sequencing by Edman's degradation and activity upon casein to better investigate this secretion. To the best of our knowledge, this is the first study to unravel the potential of *R. schneideri* cutaneous gland secretions.

Methods

Ethics statement

Animal's experiments were designed according to the Normative Resolution N. 13, from Brazilian Minister of Science, Technology and Innovation. The experiments were reviewed and approved by the Animals usage Ethic Committee from School of Pharmaceuticals Science of Ribeirão Preto – University of São Paulo (Nº: 15.1.341.60.2).

Sample collection and RNA extraction

One adult *Rhinella schneideri* toad was sacrificed through anesthetic overdose and the skin was dissected for mRNA assessment. According to the animal's body weight, a dose of combined anesthetic ketamine (10 mg/kg) and pentobarbital (150 mg/kg) was applied intraperitoneally after 3 days of cutaneous secretion milking through electric stimulation (5 V, 100 Hz, 140 ms). Toad skin patches were carefully dissected and washed with RNAlater® (Life technologies, USA), immediately thawed with dry ice and ethanol bath and kept at -80 °C until the moment of use. All the instruments and materials used were previously cleaned in Diethylpyrocarbonate (DEPC) solution 0, 1% (v/v).

The total RNA was extracted using liquid nitrogen and TRIzol® reagent (Life technologies) following the manufacturer's instructions. RNA integrity was assessed with 1% agarose gel and quantified with a Qubit® RNA assay Kit with a Qubit® 2.0 Fluorometer (Life technologies). Thus, the RNA integrity was attested using 2100 Bioanalyzer (Agilent, USA) analysis.

Toad's cutaneous secretion (CS) was stored at -20°C until the moment of usage for RP-FPLC and biochemical analysis.

Transcriptome construction and sequencing

The transcriptome was constructed using the TruSeq Stranded mRNA library kit (Illumina, USA) according to the manufacturer's instructions. The library containing 100 bp fragments was paired-end sequenced in duplicate in the Illumina HiSeq 2500 platform (Illumina).

De novo assembling and functional annotation

Raw reads were trimmed using FastQC ($Q < 20$) [24] and the adaptor sequences were discarded. The quality control was confirmed using the FastQC tool and the reads with good quality were submitted to the de novo assembling using the Trinity software with K-mer = 31. The reads were mapped against the constructed transcriptome using the Tophat tool to identify splice junctions between exons. Transcripts per million (TPM) was calculated using Salmon tool. The contigs were assembled against National Center for Biotechnology Information (NCBI) non-redundant (NR) database, with the aid of FunctionAnnotator website available at <http://163.25.92.60/index.php> [25], and specific anuran antimicrobial peptides (AMP) database DADP [26] using the blastx algorithm, to obtain the functional annotation. The resulting annotated sequences were the ones with cut-off value of significance lower than 1×10^{-5} , coverage higher than 70% and protein identity (pident) higher than 60.

Fractionation of cutaneous secretion (CS) by RP-FPLC, SDS-PAGE and N-terminal sequencing

CS was lyophilized and the dried secretion (25 mg) was dispersed in deionized water (5 mL). The insoluble part was separated after centrifugation ($10,000 \times g$, 5 min, room temperature) and the supernatant was filtered in $0.22 \mu\text{m}$ polyvinylidene fluoride (PVDF) membrane. CS solution (1.5 mL) was submitted to fast protein liquid chromatography (FPLC) in a C18 column ($5 \mu\text{m}$, $250 \times 10.0 \text{ mm}$, 300 \AA , Jupiter, Phenomenex) using Äkta Pure system (GE Healthcare) as described by Shibao et al. [15]. C18 column was firstly equilibrated with solution A (TFA 0, 1%) and the fractions were eluted with segmented gradient of acetonitrile until 100% of solution

B (acetonitrile 60% in TFA 0, 1%) under 5 mL/min flow rate and 214 nm monitoring. The resulting fractions were collated and storage at -20°C until the moment of usage. The chromatographic profile was generated using the software Unicorn 5.20 (GE Healthcare).

An aliquot of 100 μL of each fraction was dried and dispersed in 50% acetonitrile (ACN) solution. Each fraction was submitted to sodium dodecyl-sulphate-polyacrylamide gel electrophoresis (SDS-PAGE), according to Schagger and Von Jagow [27]. In addition, different volumes (5, 10 and 20 μL) of the crude secretion used to RP-FPLC was also submitted to SDS-PAGE. Bench marker Amersham low molecular weight calibration kit for SDS electrophoresis (GE Healthcare) was also used to estimate protein molecular weight. The gel was submitted to 90 V, 40 mA and 15 W for 4 h and stained with the PlusOne Silver Staining Kit (GE Healthcare).

Protein fractions identified in the SDS-PAGE were submitted to amino terminal sequencing through Edman degradation [28] by the automatic protein sequenator model PPSQ-334 (Shimadzu).

Peptides and protein alignment

Primary peptides and proteins sequences were deduced from the cDNAs sequences from the transcriptomes with ExPasy translator tool. The deduced sequences and the sequences determined by N-terminal sequencing were aligned using Multalin algorithm [29]. The alignments were formatted using Esript 3.0 [30].

Caseinolytic activity

A chromogenic proteolytic assay with the CS was performed in the presence and absence of ethylenediamine tetraacetic acid (EDTA) and phenylmethylsulfonyl fluoride (PMSF). The assay was conducted following the method described by Wang [31]. For this assay, we used 90 μL azocasein (10 mg/mL) in 50 mM Tris-HCl buffer with 0.15 M NaCl and 0.15 M CaCl_2 (pH 8.0), different volumes (10 μL , 20 μL and 30 μL) of CS (5 mg of dried secretion in 1 mL of deionized water), 100 mM EDTA or 100 mM PMSF and Tris-HCl buffer solution (100 mM) to complete the reactions to 120 μL . Positive control was performed using 10 μL Trypsin (100 mM) and negative control was carried out using the same volume of buffer. The reactions were incubated at 37°C for 90 min and stopped by adding 120 μL of 0.5 M trichloroacetic acid. All the tubes were centrifuged at $1000 \times g$ for 5 min, 150 μL of the supernatant was mixed with the same volume of 0.5 M NaOH and absorbance was determined at 450 nm. This assay was carried out in triplicate. Data were plotted using the software GraphPad Prism 6.0 (GraphPad Software Inc).

Results

Transcriptome sequencing, de novo assembly and functional annotation

The same transcriptome was sequenced in duplicate resulting in 129,467,414 and 131,652,320 raw reads (considering forward and reverse reads) for each duplicate. The data obtained from de novo assembling is summarized in Table 1. The contigs were analyzed according to their functional annotation regarding Gene Ontology (Additional file 1), hits with deposited nucleotide and protein sequences from nr NCBI database and DADP, being the latter very important for results enrichment, once there is not much information regarding toads in NCBI database.

AMP assessment

The functional analysis of the transcriptome data and the AMP Database showed the presence of 43 different peptides and protein classes. Table 2 summarizes the more abundant contigs (TPM > 100) and are clustered in 33 classes of AMP. The five major classes of AMP, considering the TPM values, are kassinin, temporin, peroniin, rugosauperolein and buforin.

Other proteins of interest

The main protein of interest that are not considered AMP are listed in Table 3. Two contigs related to cobatoxin were found in this study. The first one, identified as TRINITY_DN69643_c0_g1_i, is identical to cobatoxin from *Helicoverpa armigera*, identified by access number ADR51150.1 (gi|313,247,974). The second one, identified as TRINITY_DN121110_c0_g1_i1, has matched cobatoxin A from *Spodoptera exigua* (gi|827,029,657).

In the present transcriptome we found 19 full-length sequences with high homology to PLA₂. Interestingly, two contigs are similar to snake PLA₂. Contig TRINITY_DN72825_c0_g1_i1 encodes a PLA₂ highly similar to *Crotalus durissus terrificus* PLA₂ crotoxin basic subunit (gi 48,429,036) (Fig. 1a). This PLA₂ was also identified in the raw skin secretion in the fractions 24A, 25 and 26 (Fig. 2) and confirmed by Edman degradation sequencing of the fractions. Contig TRINITY_DN72825_c0_g2_i1 is

also related to PLA₂ from *C. d. terrificus* (Fig. 1b), but to the acid subunit (gi|129,456).

This study identified one full-length contig related to metalloproteases and 14 contigs related to serine proteases. Contig TRINITY_DN11900_c0_g1_i1 is highly homologous to a metallo-disintegrin from *Xenopus laevis*. Fourteen full-length sequences related to serine proteases were obtained. Six of them showed high similarity to *Crotalus ssp* snakes, being three (TRINITY_DN87578_c1_g1_i2, TRINITY_DN87578_c1_g1_i5, TRINITY_DN87578_c1_g1_i6,) containing the same coding sequence (herein named RsSVSP) highly related to gyroxin (Fig. 3). Contigs TRINITY_DN87578_c1_g1_i3, TRINITY_DN87578_c1_g1_i4 and TRINITY_DN87578_c1_g1_i8 encode a protein (RsSVSP2) very similar to serine protease 6 from *C. adamanteus* (gi|338,855,342).

Thirteen complete open reading frames (ORFs) related to galectins and 6 related to ficolins were found in the transcriptome. From those, 12 are related to predicted galectin from different genomes and transcriptomes. Contig TRINITY_DN92920_c1_g1_i1 is similar to galectin from *Rhinella arenarum* ovary. Four complete ORFs were found matching different galectins from *Xenopus* genome assessment. All contigs related to ficolins were annotated against model organisms' genome (*Xenopus spp.* and *Monodelphis domestica*).

Fractionation of CS by RP-FPLC, SDS-PAGE and N-terminal sequencing

Crude secretion SDS-PAGE profile is shown in Additional file 2. CS was separated in 26 fractions, named CS1 to CS26 (Fig. 2). The fractions were further submitted to SDS-PAGE, gel was stained with silver and the fractions named CS1, CS24, CS25 and CS26 (Fig. 2, insert) were identified containing protein compounds. In addition, probably fractions CS5, CS21 and CS22 also contain protein molecules, but due to its low concentration, they were not investigated in this study.

Both gels show some interference on their top, probably caused by the silver staining. It is possible to see bands with approximately 16 kDa in the fractions CS24, CS25 and CS26, which showed similar diffusion profiles. Therefore, these fractions were submitted to N-terminal sequencing by Edmans' degradation, but it was possible to obtain only CS1 and CS24 – CS26 partial sequences (Table 4).

Caseinolytic activity

The transcriptome functional annotation showed some sequences that can be related to serine and metalloproteases. In order to investigate if the sequences could really deduce these enzymes, we performed a proteolytic test using azocasein as substrate (Fig. 4).

Table 1 Statistical analysis of the transcriptome sequencing and de novo assembling with Trinity

Parameters	Values
Number os trimmed reads	174,308
Average size (bp)	633.58
N50 (bp)	365
Largest sequence (bp)	22,684
Smallest sequence (bp)	201

Table 2 List of main antimicrobial peptides from transcriptome

AMP families	Accession number	Identification	ENA Transcript identification	TPM
Aurein	sp P69021	Aurein-3.1	TRINITY_DN64440_c0_g1_i1	256.721
Bombesin	sp P84211	Bombesin-like peptide	TRINITY_DN48099_c0_g1_i1	164.379
	sp P84212	Bombesin-like peptide	TRINITY_DN27392_c0_g1_i1	192.298
			TRINITY_DN61842_c0_g2_i1	106.099
			TRINITY_DN80913_c0_g1_i2	102.443
	sp P84214	Bombesin	TRINITY_DN71876_c0_g1_i1	117.297
	sp P86026	[Asn3,Lys6,Phe13] 3-14-bombesin	TRINITY_DN76333_c0_g1_i1	135.698
Bradykinin-like peptide	sp P84823	Bradykinin-like peptide	TRINITY_DN71047_c0_g1_i2	123.521
Brevinin	sp P32423	Brevinin-1	TRINITY_DN70354_c0_g3_i1	175.329
	sp P82233	Brevinin-1Ta	TRINITY_DN73322_c2_g2_i1	122.993
Buforin	sp C0HJB7	Buforin-EC	TRINITY_DN54614_c0_g1_i1	766.952
			TRINITY_DN60267_c0_g1_i1	190.875
			TRINITY_DN60267_c1_g1_i1	231.457
			TRINITY_DN60267_c0_g2_i1	228.282
Caerin	sp P69032	Caerin-2.3	TRINITY_DN66726_c0_g1_i1	182.654
			TRINITY_DN66726_c0_g1_i2	121.242
Caeridin	sp P82076	Caeridin-4	TRINITY_DN79589_c4_g3_i1	226.024
Citropin	sp P81840	Citropin-1.2	TRINITY_DN86384_c3_g3_i1	111.715
Cruzioseptin	sp C0HK11	Cruzioseptin-15	TRINITY_DN70275_c0_g1_i1	293.883
Dermaseptin	sp P85523	Dermaseptin-1	TRINITY_DN76697_c0_g1_i2	115.602
Hylambatin	sp P08614	Hylambatin	TRINITY_DN87362_c21_g12_i1	174.475
Hylin	sp P84003	Hylin-b2	TRINITY_DN84946_c0_g1_i2	971.000
Hyposin	sp P84957	Hyposin-HA3	TRINITY_DN72338_c1_g1_i1	616.344
Japonicin	sp P83306	Japonicin-1	TRINITY_DN64014_c0_g1_i1	310.058
Kassinin	sp P08611	Kassinin	TRINITY_DN120600_c0_g1_i1	346.986
			TRINITY_DN127618_c2_g1_i1	5081.92
			TRINITY_DN52216_c0_g1_i1	247.853
	sp P42988	Kassinin-like peptide	TRINITY_DN81065_c1_g3_i1	130.979
			TRINITY_DN87139_c3_g3_i1	164.946
Nigrocin	sp B3A0M7	Nigrocin-2JDb	TRINITY_DN87453_c9_g4_i1	327.271
	sp P0C8U1	Nigrocin-2HSb	TRINITY_DN85182_c2_g10_i1	122.641
			TRINITY_DN85182_c2_g3_i1	221.000
Peptide PGLa	sp C0HKP4	Peptide PGLa-R6	TRINITY_DN80328_c3_g2_i1	229.068
Peroniin	sp P86487	Peroniin-1.1	TRINITY_DN53498_c0_g1_i1	279.654
	sp P86488	Peroniin-1.1a	TRINITY_DN67234_c0_g1_i1	347.803
	sp P86493	Peroniin-1.3a	TRINITY_DN81622_c0_g1_i1	129.528
	sp P86495	Peroniin-1.1b	TRINITY_DN46821_c0_g1_i1	841.554
Phyllocaerulein	sp P86625	[Arg4]-Phyllocaerulein	TRINITY_DN85513_c6_g2_i2	280.537
Phylloseptin	sp P84569	Phylloseptin-4	TRINITY_DN87185_c0_g1_i2	109.815
	sp P84931	Phylloseptin-3	TRINITY_DN64642_c0_g2_i1	168.944
	sp P86283	Phylloseptin Bu-2	TRINITY_DN78083_c0_g4_i1	389.329
		TRINITY_DN78083_c0_g6_i1	334.985	
		TRINITY_DN78083_c0_g8_i1	126.679	

Table 2 List of main antimicrobial peptides from transcriptome (*Continued*)

AMP families	Accession number	Identification	ENA Transcript identification	TPM		
Ranacyclin	sp P83663	Ranacyclin-E	TRINITY_DN46765_c0_g1_i1	685.023		
			TRINITY_DN81581_c5_g1_i4	371.815		
Ranalexin	sp P82876	Ranalexin-1Ca	TRINITY_DN110865_c1_g1_i1	343.086		
Ranatensin	sp P08951	Ranatensin-C	TRINITY_DN70922_c0_g2_i1	116.215		
Ranauterin	sp P86020	Ranaturin-9	TRINITY_DN85516_c7_g13_i1	111.483		
Riparin	sp P86125	Riparin-1.2	TRINITY_DN83324_c7_g3_i1	215.018		
			TRINITY_DN92795_c0_g1_i1	164.545		
			TRINITY_DN85374_c2_g1_i1	330.835		
			TRINITY_DN86546_c0_g1_i2	264.226		
Rothein	sp P86126	Riparin-1.3	TRINITY_DN86547_c3_g6_i1	130.999		
			TRINITY_DN87840_c8_g1_i1	261.426		
			TRINITY_DN87822_c5_g12_i1	826.816		
Rugosin	sp P08616	Rugosauperolein-2	TRINITY_DN82274_c0_g1_i5	196.425		
	sp P84912	Rugosin A-like peptide	TRINITY_DN82274_c0_g1_i6	314.015		
Signiferin	sp P86123	Signiferin-1	TRINITY_DN86818_c0_g1_i1	140.473		
Temporin	sp C0HJB9	Temporin-ECa	TRINITY_DN86407_c15_g1_i1	274.636		
			TRINITY_DN86407_c15_g2_i1	915.026		
			TRINITY_DN86407_c15_g3_i1	478.761		
			TRINITY_DN86407_c15_g4_i1	527.791		
			TRINITY_DN86407_c15_g5_i1	256.836		
			TRINITY_DN86407_c15_g7_i1	176.844		
			sp P0C5X6	Temporin-1DYa	TRINITY_DN29229_c0_g1_i1	133.681
			sp P56917	Temporin-A	TRINITY_DN13166_c0_g1_i1	527.935
					TRINITY_DN13166_c0_g2_i1	158.363
			sp P82832	Temporin-1Lc	TRINITY_DN2695_c0_g1_i1	175.444
					TRINITY_DN72771_c0_g1_i1	105.843
					TRINITY_DN73309_c0_g1_i1	140.262
					TRINITY_DN80955_c1_g1_i1	254.369
					TRINITY_DN80955_c1_g2_i1	387.789
TRINITY_DN84769_c0_g1_i1	373.663					
TRINITY_DN86165_c3_g1_i1	155.493					
sp P82881	Temporin-1Cb	TRINITY_DN56332_c0_g1_i1			105.694	
sp P84858	Temporin-GH	TRINITY_DN77159_c0_g2_i9	102.951			
Tigerinin	sp P82652	Tigerinin-2	TRINITY_DN70574_c0_g1_i1	242.086		
			TRINITY_DN71694_c0_g1_i1	101.217		
			TRINITY_DN85395_c7_g6_i1	130.962		
			sp C0HL42	Tigerinin-2OS	TRINITY_DN85615_c2_g3_i2	125.931
			sp P82651	Tigerinin-1OS	TRINITY_DN17276_c0_g1_i1	245.073
					TRINITY_DN43562_c0_g1_i1	104.073
Uperolein	sp P08612	Uperolein	TRINITY_DN6203_c0_g1_i1	454.682		
			TRINITY_DN73092_c0_g1_i1	248.482		
			TRINITY_DN85479_c6_g29_i1	457.183		
			TRINITY_DN111770_c0_g1_i1	107.555		
			TRINITY_DN83356_c0_g1_i1	424.761		

Table 2 List of main antimicrobial peptides from transcriptome (*Continued*)

AMP families	Accession number	Identification	ENA Transcript identification	TPM
			TRINITY_DN83356_c0_g1_i2	339.468
			TRINITY_DN83356_c0_g1_i3	101.171
			TRINITY_DN81958_c0_g1_i1	277.617
			TRINITY_DN81958_c0_g1_i2	216.205
Uperin	sp P82036	Uperin-5.1	TRINITY_DN87247_c0_g1_i1	332.281
			TRINITY_DN87247_c1_g1_i1	137.665
			TRINITY_DN87247_c0_g1_i3	185.119
			TRINITY_DN77327_c0_g2_i1	238.182
			TRINITY_DN87210_c0_g1_i1	105.306
			TRINITY_DN87210_c0_g2_i1	132.473

CS was capable of degrading azocasein. The tests in presence of EDTA and PMSF did not show differences.

Discussion

Although toads are widely spread in Brazilian and Latin America territory, its cutaneous peptides and proteins have come to the spotlight only recently and the scarce information hinders its assessment [11, 16, 17].

The present study reveals the abundance of the *Rhinella schneideri* mucous gland defense peptides and protein through the RNA-seq transcriptome and some peptides and proteins present in its secretions. We used the cutaneous secretions from one specimen that was milked prior the gland extraction to obtain a more accurate result comparing the biochemical tests with the transcriptome.

Transcriptome survey is snapshots of mRNA obtained in a specific time. Therefore, the first step of this study was milking the toad's skin secretions to maximize the transcripts production. We also used the secretion to carry out biochemical analysis subsequently. Using the secretions from the same toad to perform the experiments carried out in this paper allows us to discard any variation that may occur and obtain a more accurate analysis. RNA-seq analysis revealed the presence of distinct proteins and peptides. Most of the transcripts are related to housekeeping function, as expected, but we found novel proteins in the skin. As frog defense peptides are better characterized, we enriched the functional annotation using DADP database. Using this approach, there were several classes of AMP that are summarized in Table 2. The contigs that presented the highest values of TPM belong to tachykynins, temporin, peroniin and buforin classes. *Bufo bufo*, *B. verrucosissimus* and *B. variabilis* skin secretions act as antimicrobial agent against *Escherichia fecalis* and *E. faecium* with lower minimum inhibitory concentration than ampicillin [18].

The molecular classes of AMPs found in this transcriptome are discussed below.

Kassinin and Rugosaperolein

Kassinin and rugosaperolein AMPs belong to the tachykinin family of neuropeptides. Tachykinins are well present in amphibian secretions and higher organisms, acting as neurotransmitters and neuroprotective agents in the latter. They present a C-terminal conserved region Phe-X-Gly-Leu-Met, that is known for this family's activity [32]. Regarding amphibians secretions, kassinin, a dodecapeptide, was the first tachykinin described from *Kassina senegalensis* methanol skin extract, which shows similar activity as substance P, physalaemin and eldosein, being able to stimulate salivary secretion, act as hypotensive agent and stimulate smooth muscle [33, 34]. Contig TRINITY_DN127618_c2_g1_i1 is very similar to the first kassinin ever identified.

Rugosaperolein was discovered in *Uperoleia rugosa* skin methanol extracts and was named after the tachykinin uperolein [35].

Tachykinins are widely studied mainly in higher organisms due their interesting neuroprotective activity. Despite it is known that they are present in amphibian skin, their role as AMP is poorly studied. As their role as neurotransmitters is well established, here we hypothesize they act not as a AMP, but as a host defense molecule that prevents the toad from predation along with micro molecules that are already known to play this role [36]. We use the same hypotheses to the presence of peroniin. Although we cannot discard the possibility of peroniin present any antimicrobial activity, we believe it is produced to prevent *R. schneideri* to be swallowed, once tachykinins are known to show neuroactivity being able to cause gut tissue contraction [37].

Temporin

Temporins were discovered in a cDNA library from *Rana temporaria* skin and later obtained in *R. temporalis*' skin secretions. They are effective against gram positive *Bacillus megaterium* and *Escherichia coli* with different sensitivities. This AMP family comprises

Table 3 List of other peptides and protein of interest from the transcriptome

Protein class	Accession number	Identification	ENA transcript identification	TPM
Cobatoxin	gb DR51150.1	Cobatoxin	TRINITY_DN69643_c0_g1_i1	94.407
	gb AKJ54497.1	Cobatoxin A	TRINITY_DN121110_c0_g1_i1	0.40738
Galectin	gb KPJ04718.1	Galectin-4	TRINITY_DN74940_c0_g1_i1	0.372625
	gij 847,127,031	Galectin-8	TRINITY_DN78589_c0_g2_i1	149.955
	gij 768,932,680	Galectin-4-like	TRINITY_DN78656_c0_g1_i1	224.999
	gij 692,190,428	Galectin 9	TRINITY_DN80231_c0_g1_i1	0.180761
			TRINITY_DN80231_c0_g1_i2	0.188734
	gij 591,365,832	Galectin-9C-like	TRINITY_DN83688_c1_g1_i1	101.232
			TRINITY_DN83688_c1_g1_i4	170.619
	gij 928,062,140	Galectin-4-like	TRINITY_DN83688_c1_g1_i3	0.702454
	gij 512,835,424	Galectin-12	TRINITY_DN86454_c0_g1_i1	197.028
	gij 512,835,420	Galectin-12	TRINITY_DN86454_c0_g1_i2	189.779
		TRINITY_DN86454_c0_g1_i3	0.440812	
		TRINITY_DN86454_c0_g1_i4	0.561889	
Ficolin	gij 2,554,855	Galectin A Chain A,	TRINITY_DN92920_c1_g1_i1	672.677
	gij 512,864,759	Ficolin-2-like	TRINITY_DN75601_c0_g1_i1	0.832349
	gij 530,624,987	Veficolin-1-like	TRINITY_DN81241_c1_g5_i1	369.808
	gij 611,978,444	Ficolin-2-like	TRINITY_DN85932_c0_g1_i3	104.045
	gij 512,864,763	Ficolin-2-like	TRINITY_DN86611_c9_g1_i1	932.407
	gij 148,230,483	Ficolin	TRINITY_DN86611_c9_g2_i1	120.763
		TRINITY_DN86611_c9_g2_i2	205.043	
Phospholipase A ₂	gij 512,862,492	Cytosolic Phospholipase A2 zeta	TRINITY_DN47903_c0_g1_i1	0.447251
	gij 512,929,364	Group XIIA secretory phospholipase A2	TRINITY_DN86978_c1_g1_i1	0.340838
	gij 530,606,756	Phospholipase A2, minor isoenzyme-like	TRINITY_DN69641_c0_g1_i1	0.724376
	gij 657,561,888	Group 10 secretory phospholipase A2-like	TRINITY_DN71078_c0_g1_i1	111.032
	gij 700,363,984	Cytosolic phospholipase A2 gamma	TRINITY_DN84494_c4_g4_i1	0.294445
			TRINITY_DN84494_c4_g5_i1	0.141515
	gij 847,119,410	85/88 kDa calcium-independent phospholipase A2	TRINITY_DN79625_c2_g1_i1	0.48039
			TRINITY_DN79625_c2_g1_i2	0.508773
	gij 884,758,924	Cytosolic phospholipase A2 gamma	TRINITY_DN60871_c0_g1_i1	0.0554201
			TRINITY_DN60871_c0_g2_i1	0.273762
			TRINITY_DN60871_c0_g3_i1	0.081637
			TRINITY_DN60871_c0_g4_i1	0.278435
	gij 512,862,492	Cytosolic phospholipase A2 zeta	TRINITY_DN54445_c0_g1_i1	162.823
	gij 148,223,595	Cytosolic phospholipase A2	TRINITY_DN71015_c0_g1_i1	107.048
		TRINITY_DN71015_c0_g2_i1	0.0297482	
gij 48,429,036	Phospholipase A2 crotoxin basic subunit CBc	TRINITY_DN72825_c0_g1_i1	320.644	
gij 129,456	Phospholipase A2 homolog crotoxin acid subunit CA	TRINITY_DN72825_c0_g2_i1	388.741	
gij 512,834,221	Group XV phospholipase A2 isoform X2	TRINITY_DN74674_c0_g1_i1	202.101	

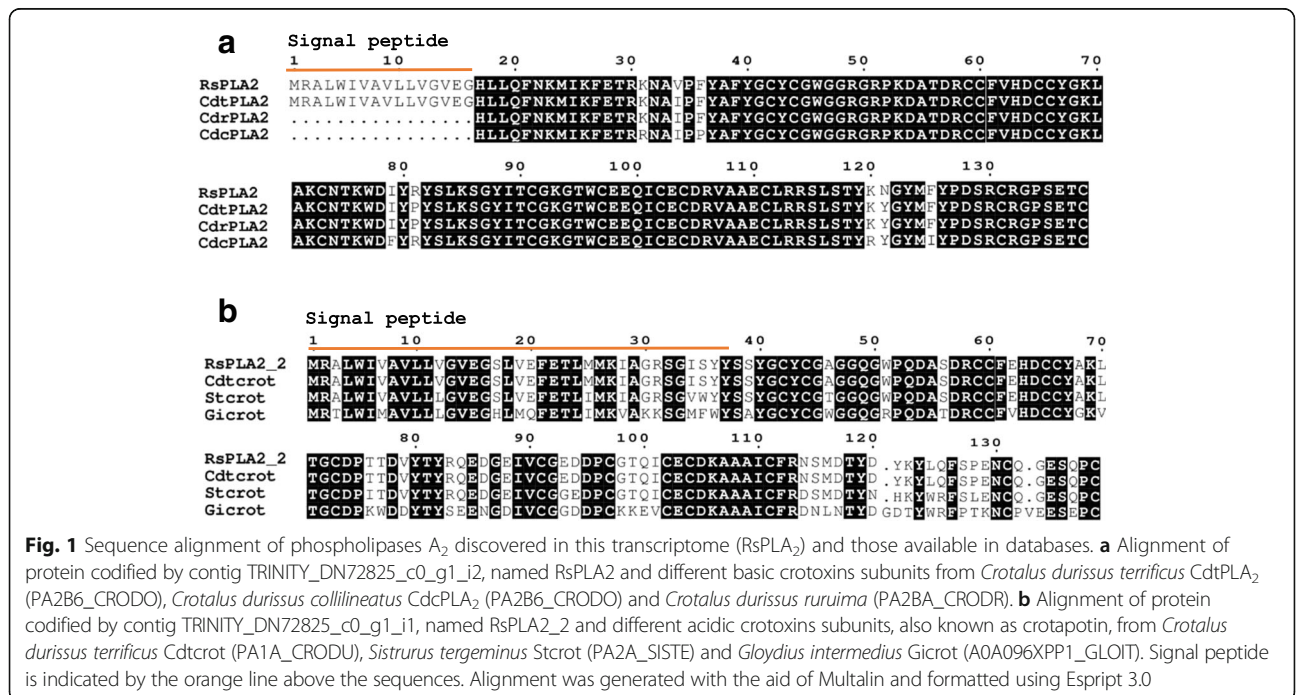
Table 3 List of other peptides and protein of interest from the transcriptome (Continued)

Protein class	Accession number	Identification	ENA transcript identification	TPM
Metallo proteases	gij163,916,015	Phospholipase A2, group X1A	TRINITY_DN74706_c0_g1_i1	250.869
	gij148,222,675	Disintegrin and metalloproteinase domain-containing protein 22 precursor	TRINITY_DN11900_c0_g1_i1	0.414536
Serine proteases	gb ACI32835.1	Serine proteinase-like protein 1	TRINITY_DN79566_c0_g1_i1	0.486137
	gij913,306,165	Serine proteinase stubble	TRINITY_DN86427_c0_g1_i1	0.326702
	gb ACI32835.1	Serine proteinase-like protein 1	TRINITY_DN87507_c0_g1_i1	17.516
			TRINITY_DN87507_c0_g1_i3	126.092
			TRINITY_DN87507_c0_g1_i4	156.004
			TRINITY_DN87507_c0_g1_i5	0.772103
	gij380,875,411	Thrombin-like Enzyme gyroxin B1.4	TRINITY_DN87578_c1_g1_i2	0.172693
			TRINITY_DN87578_c1_g1_i5	0.125578
			TRINITY_DN87578_c1_g1_i6	0.658672
	gb AEJ32000.1	Serine proteinase 6	TRINITY_DN87578_c1_g1_i3	0.266027
TRINITY_DN87578_c1_g1_i4			0.246148	
TRINITY_DN87578_c1_g1_i8			0.29959	
TRINITY_DN87890_c1_g6_i3			0.193263	
gij827,563,139	Serine proteinase-like protein isoform X1	TRINITY_DN87890_c1_g6_i3	0.193263	
gb AAM69353.1	Serine proteinase-like protein 2	TRINITY_DN87890_c1_g6_i5	0.508857	

anionic, hydrophobic, small peptides (8–17 amino acids) that are folded in alpha-helices and which potency is related to the final net charge [38, 39]. Following its first discovery, temporins were also found in secretions of different species of the genus *Euphyctis*, *Limnonectes*, *Hypsiboas*, *Amolops*, *Hylarana* and *Lithobates* [39]. Regarding their biotechnological applications, there are

more than 20 deposited patents related to them, which varies from their pharmacological use as anti-HIV to obtaining of transgenic plant resistant to pathogens [39].

Deposited contig TRINITY_DN86407_c15_g2_i1 is similar to temporin-E_{Ca}, from *Euphyctis cyanophlyctis*, the skittering frog. As other temporins, it shows activity against gram-positive bacteria *E. coli*, *K. pneumonia*,



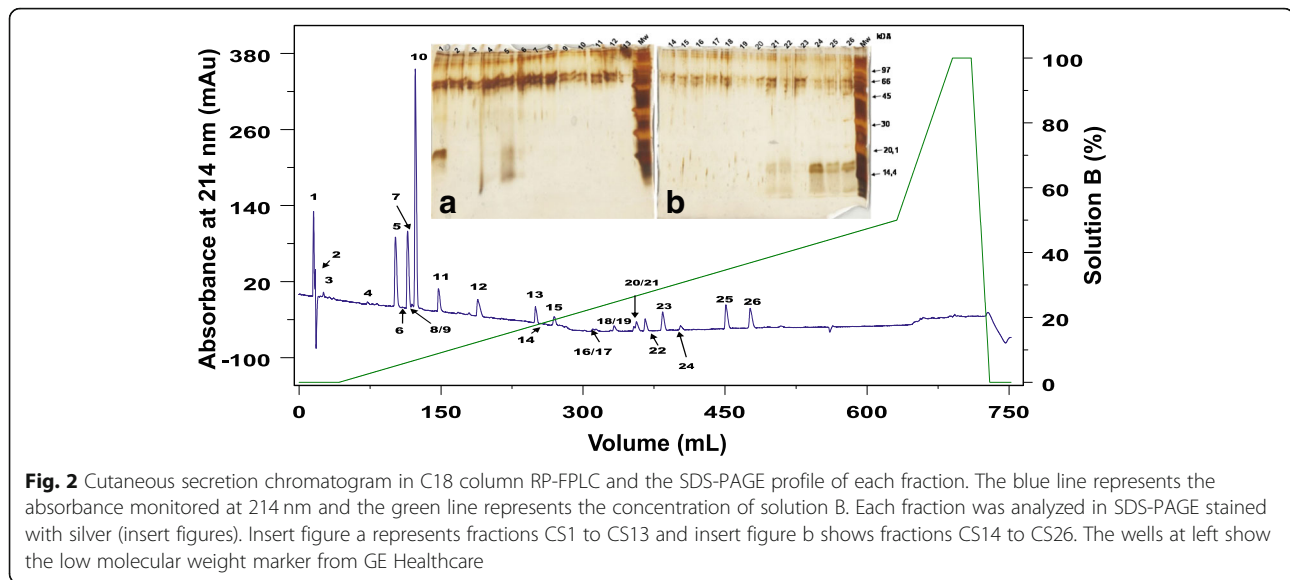


Fig. 2 Cutaneous secretion chromatogram in C18 column RP-FPLC and the SDS-PAGE profile of each fraction. The blue line represents the absorbance monitored at 214 nm and the green line represents the concentration of solution B. Each fraction was analyzed in SDS-PAGE stained with silver (insert figures). Insert figure a represents fractions CS1 to CS13 and insert figure b shows fractions CS14 to CS26. The wells at left show the low molecular weight marker from GE Healthcare

Micrococcus luteus and *Staphylococcus aureus*, and low hemolytic activity [40].

Peroniin

Peroniins discovery was made through mass spectrometry analysis of *Litoria peronii* skin secretions in winter and summer. Albeit they are considered AMP, there is no report of their test against neither gram positive nor gram negative bacteria. In fact, peroniins are the major component of *L. peronii* secretions both summer and winter and they possess activity over smooth muscle causing its contraction. There is only one report of peroniin up to date [41].

Buforin

This AMP family was first discovered in *Bufo Bufo garzarizans* stomach. This family comprises small peptides (approximately 6.5 kDa) and are effective against several gram positive and negative bacteria [42]. They are derived from histone H2A and belong to the toads' innate immune system. While buforin I is secreted in the stomach protecting the toad against pathogen ingested microorganisms and further binding to the mucosa biofilm enhancing its protection, buforin II does not have its mechanism of action fully elucidated, but it is hypothesized they bind to microorganism's nucleic acid destroying it [43]. In this study, contig TRINITY_DN54614_c0_g1_i1 is related to buforin-EC, isolated from skin secretions of frog *Euphlyctis cyanophlyctis*, which has shown activity against *Staphylococcus aureus* and *Escherichia coli* [40]. We believe temporins and buforins, along with other AMPs, act as protection against microorganisms, due to the well-established activity of these molecules against gram

positive and negative bacteria. The diversity of AMP found in the skin proves that glandular secretions work as a biochemical enriched barrier for the toads' protection.

In addition to the AMPs many other protein components, which probably have relevant roles for frog defense, have been identified in the transcriptome, among them cobatoxins, PLA₂, proteases, ficolin and galectins. These molecular classes are discussed below. Interestingly, ficolin was one of the major contigs found in the transcriptome survey.

Cobatoxin

Cobatoxins were firstly reported in the *Centruroides noxius* scorpion poison as potassium channel-blocking toxins, belonging to the α -K-toxins, subfamily 9. They are moderate affinity blockers of K⁺ voltage-dependent channels Shaker and K_v1.1 [44]. Cobatoxin from *Helicoverpa armigera* is mostly connected to the insect defense. Its level of expression has significantly raised after infection with both gram positive and negative bacteria [45]. Regarding *Spodopetra exigua* cobatoxin, gene expression analysis from the insect midgut after the exposition of the insect to *B. thuringiensis* toxins revealed an increase of the mRNA coding for this protein, indicating that it plays a fundamental role in the insect defense [46]. Furthermore, an analysis of *Galleria mellonella* challenged against *Micrococcus luteus* has shown that cobatoxin is likely to maximize the potential of other innate AMPs from the insect [47].

Here, we hypothesize that cobatoxin is also part of the immune system of the toad.

We also found other proteins of interest that are not catalogued as AMP with the aid of non-redundant NCBI database annotation:

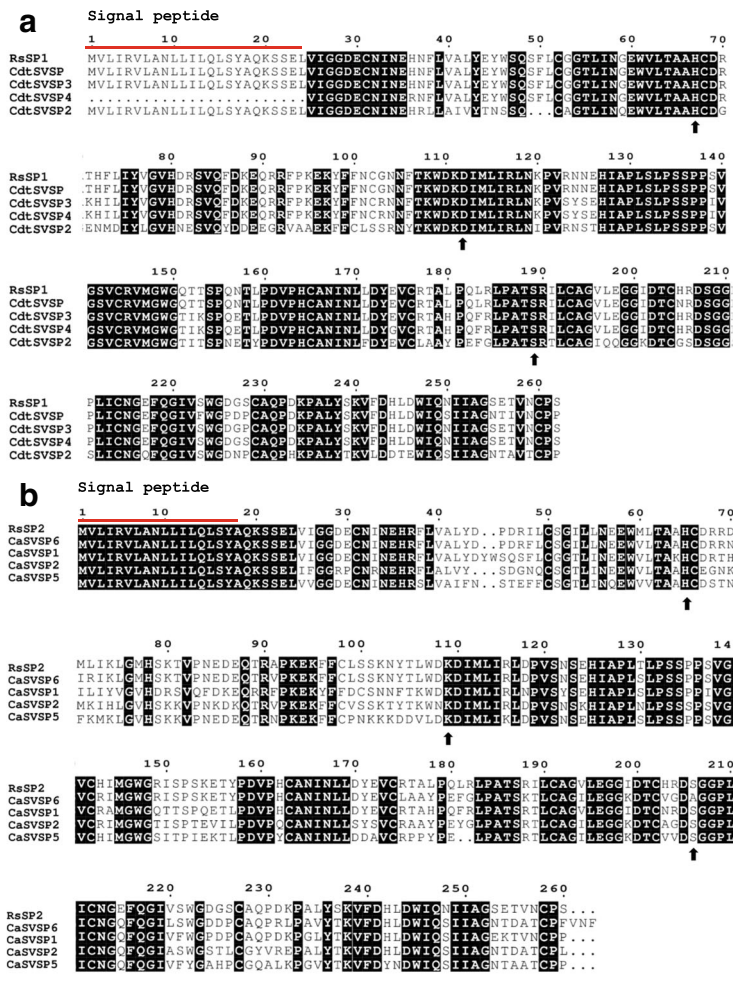


Fig. 3 Sequence alignment of serine protease discovered in this transcriptome (RsSP) and those from different snake venoms (SVSP) available in databases. **a** Alignment of protein codified by contigs TRINITY_DN87578_c1_g1_j3, TRINITY_DN87578_c1_g1_j4 and TRINITY_DN87578_c1_g1_i8, named RsSP1 and gyroxin serine proteases from snake venoms (SVSPs) from *Crotalus durissus terrificus* (CdtSVSP3, VSP13_CRODU), CdtSVSP 2 (VSP21_CRODU), CdtSVSP4 (VSP14_CRODU). **b** Alignment of serine protease codified by contigs TRINITY_DN87578_c1_g1_i2, TRINITY_DN87578_c1_g1_i5 and TRINITY_DN87578_c1_g1_i6, named RsSP2 and gyroxin SVSP from *Crotalus adamanteus* 1, CaSVSP1 (VSP1_CROAD), CaSVSP2 (VSP2_CROAD), CaSVSP5 (VSP5_CROAD) and CaSVSP6 (A0A1W7R8B4_CROAD). Signal peptide is indicated above the sequences and the arrows beside the alignment indicates the amino acids that are important to the catalytic activity. Alignment was generated with the aid of Multalin and formatted using Esprint 3.0

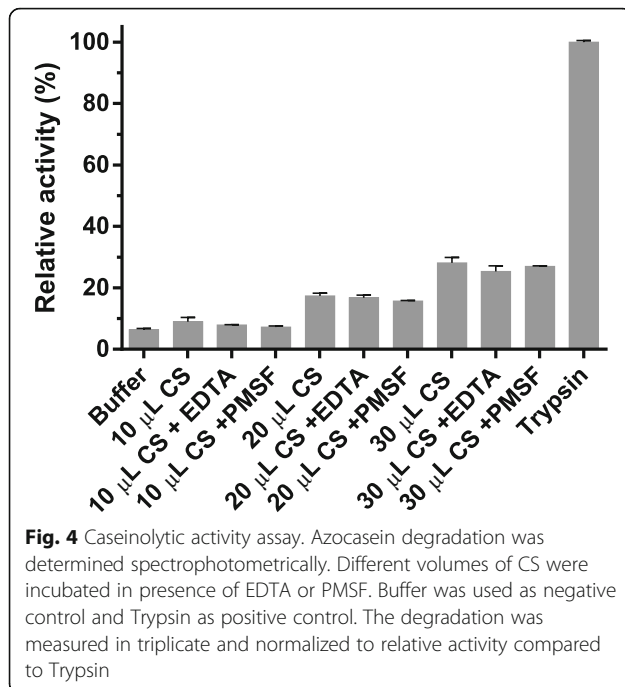
Phospholipase A₂

Phospholipase A₂ (PLA₂) catalyze the hydrolysis of phospholipids in the sn2 position releasing arachidonic acid and lysophosphatidic acid, which are precursors of signaling molecules in immune response, inflammation, pain and cell regulatory processes [48–50]. They

can be found in different tissues and organelles and are often small proteins (14–18 kDa), and their stability varies according to the number of disulfide bonds. Secreted PLA₂s are one of the major components of Elapidae and Viperidae snake venoms [51]. In fact, the crotoxin was the first toxin isolated almost 100 years

Table 4 N-terminal sequences from the poison fractions by Edman's degradation technique

Fraction	Protein sequence	Protein family
CS1	RWEECEDCDDDQDDQQQQCAKDDDDDE--Q-QQD	Lectin
CS24a	GLLEFNKMIKFETRKNAIIFYAFYGCYCWGGRRRPK	PLA ₂
CS24b	-TPFYKGAAGQQVPQDIVNYYAFGGCQK—EP-PMRY-EY-VNAGGDQ-----D	Galectin
CS25	-L-EFNKMIKFETRKNAIIFYAFY	PLA ₂
CS26	-L-EFNKMIKFETRKNAIIFYAFYGCYCWGGRRRPKDA	PLA ₂



ago and its sequence has been determined for more than 30 years now, and the cloning was successfully obtained in the 80's [52, 53].

In *Crotalus* snake venoms, crotoxin is composed of two non-covalently bound subunits (one acidic and one basic). The basic component (CB) is a catalytic active PLA₂ whereas the acidic component (CA) is a PLA₂ catalytically inactive responsible to direct CB towards specific sites that lead to neurotoxic actions [54–58].

In humans, secreted PLA₂ group IIA can be found in tears and it is most likely to play a defensive role in eyes defense against gram-positive bacteria, but no response to gram-negative bacteria [59]. These molecules are also found in dromedary tears and showed activity against both gram positive and negative bacteria [60]. Thus, a PLA₂ isolated from *Daboia Russelli* venom was able to strongly inhibit gram negative bacteria and also showed activity against gram positive bacteria isolated from human [61]. Therefore, we also assume the protective action in toad skin.

Serine and metalloproteases

The contig TRINITY_DN11900_c0_g1_i1 is highly homologous to a metallo-disintegrin protease that has been related to *Xenopus laevis* reproduction [62] and neural crest development [63]. Although it was possible to obtain a full-length transcript in neural cells, the metallo-disintegrin did not have the catalytic domain, which indicates this protein acts as a transmembrane receptor [63].

Several full-length sequences were related to serine proteases. Three of them containing the same coding sequence (herein named RsSVSP) highly related to gyroxin, a non-lethal serine protease with neurotoxic effects that causes a neurological syndrome in mice known for the animal's movements as rotation to a barrel roll [64]. The contigs TRINITY_DN87578_c1_g1_i3, TRINITY_DN87578_c1_g1_i4 and TRINITY_DN87578_c1_g1_i8 encode a protein, named RsSVSP2, very similar to serine protease 6 from *C. adamanteus* (gi|338,855,342), also found in a transcriptoma survey [65].

Ficolin and galectin

Ficolins are a group of oligomeric lectin that present fibrinogen-like and collagen-like domains and possess a carbohydrate binding domain (CRD), being N-acetylglucosamine (GlcNAc) the carbohydrate that presents the major number of galectins specificity [66, 67]. They are capable of activating complement system via lectin pathway, and aggregate some bacteria enhancing phagocytosis showing their relevant role in organisms' defense [68].

Galectins are lectins which major ligand is β -galactose-containing glycoconjugates and show their CRD conserved. They may bind to cell-surface and matrix glycans, being able to control intracellular signaling, and protein interactions dependent pathways [69]. Due to its action over neurological system, we believe it also plays a possible role as host defense peptides (HDP) and may provide protections against animal predation adding or even enhancing the symptoms of micro molecules present in parotoid glands that show neurological effects [10].

The contig TRINITY_DN92920_c1_g1_i1 is similar to galectin, an S-type lectin, from *Rhinella arenarum* ovary, probably playing a developmental regulatory role [70]. In contrast, *Xenopus* galectins from the animal's skin are believed to act as HDP [71]. Recently, a galectin was found in parotoid secretion from *R. schneideri* using mass spectrometry de novo sequencing [11].

Crude secretion has shown a different profile from parotoid poison presented before [11, 12]. In the cutaneous secretion (Additional file 2) it is possible state there are a richness of protein content which molecular weight varies from high molecular mass (around 100 kDa) to low molecular mass (less than 14 kDa), but there is a strong band with approximately 60 kDa that is common to the profile presented in male and female parotoid secretions from Piauí, Brazil [11]. After analyzing the variation of mass, crude secretion was submitted to RP-FPLC to further investigation.

The RP-FPLC of the secretion resulted in 26 fractions, which were further analyzed by SDS-PAGE. The fractions CS1, CS24-CS26 were submitted to Edman's

degradation sequencing, because they showed protein bands on SDS-PAGE. Other fractions might contain micro molecules from poison and this method of chromatography was already used for isolation of those molecules from *R. schneideri* parotoid poison before [15].

Fractions CS24, CS25 and CS26 showed sequences regarding a PLA₂ from snake venom [72]. Fractions CS24a, CS25 and CS26 contain PLA₂ isoforms similar to a basic PLA₂ isolated from *Crotalus durissus terrificus* (PA2BF_CRODU). As expected, we were able to predict the protein sequence from the contig TRINITY_DN72825_c0_g1_i1. We also retrieved contigs related to the acidic PLA₂. In crotoxin, one of the major components of *Crotalus spp.* venom, both units (basic and acidic PLA₂) are bonded by a non-covalent bound. The acid subunit (also named crotapotin) is devoid of activity, acting as a stabilizer, and the basic subunit is catalytically active and toxic [52]. Despite its role in the snake venom, we hypothesize this PLA₂ is related to toad's immune system, similarly to secreted PLA₂ in human [59] and dromedary tears [60], especially facing the probability of the toad infection by amoeba.

CS24b sequencing also presents a protein that showed similarity to a galectin-1 from *Rhinella arenarum* (gi|2558855; sp.|P5621). It is a beta-galactosyl-binding lectin discovered in toad ovary and similar to the mammalian one that binds to different carbohydrates intra and extra-cellular, probably regulating developmental process in toads oocytes [70]. There are three contigs related to this protein, but none of them is complete. *Xenopus laevis* is known to secrete a lectin, along with other defense peptides, which plays a role not only in defense against pathogenic microorganisms, but also against predators and another unknown structural role [71].

CS1 sequenced revealed this fraction contains more than one protein; the first putative protein is similar to a C-type lectin isolated from *Helicoverpa armigera* (gi|385,202,653) that is up regulated in the presence of *Escherichia coli* injection, indicating its role as a defense protein [45], but it was not found in the transcriptome. CS1 is also similar to the lectin from *Naegleria glaberrima* (gi|290,983,012), an amoeba that can be found in moist habitats and freshwater, that was unrevealed in its genome [73]. We could not retrieve any contigs related to this protein, which indicates a possible contamination of the toad skin by this microorganism. This contamination may explain the high levels of defense peptides and proteins found in the transcriptome and other fractions. It is known that toads might produce toxins that are specific against pathogens [74].

Beyond the AMP assessment and the discovery of other peptides and proteins of defense, we also retrieved some proteases. As the transcriptome showed

some evidence of serine and metalloproteases, we decided to carry out an exploratory experiment to survey the presence of these proteins in the poison. Azocasein degradation assay is a classic test to determine the presence of either of these enzymes since proteolytic activity is evaluated. Raw secretion can result in dose dependent azocasein degradation. We used EDTA, which is a chelator agent that inhibits metalloproteases, and PMSF that inhibits serine proteases to investigate which class is most likely to be present in the raw secretion and cause the casein degradation. Despite our best efforts, it was not possible to determine which class of protease was more active, once the controls with PMSF and EDTA did not show any statistical relevance. However, we believe serine protease are the main responsible for this result, since there are more full-length contigs with higher TPM in the transcriptome and only one low expressed (TPM <1) full-length contig coding for a metalloprotease. In addition, a serine protease was found in *Bufo bufo* parotoid secretion through a proteomic analysis [17], indicating it may be produced in this toad secretion too, but further characterization is needed.

Here we presented the first transcriptome survey from *R. schneideri* skin. As the results showed, this study paves the way for discovering new molecules besides characterizing an important secretion and the glands where they are produced.

Conclusions

Although *Rhinella schneideri* toads are known to possess bioactive molecules in its secretion, it is still poorly studied when compared to other venomous and poisonous animals. Most of the studies comprises the parotoid micro molecules secretions in detriment of cutaneous secretion. This is the first study to use a high throughput RNA-seq technology to investigate *R. schneideri* cutaneous secretions and the first one to focus on defense peptides and proteins. Furthermore, using the milked secretion and skin from the same toad allowed an accurate analysis of protein expression, once individual variation was dismissed. The results obtained herein showed evidence of novel AMP and enzymes that need to be further explored.

Additional files

Additional file 1: Gene ontology of *Rhinella schneideri* skin transcriptome. The Gene Ontology is divided in biological process, molecular function and cellular component. (DOCX 13 kb)

Additional file 2: Electrophoretic profile of *Rhinella schneideri*'s cutaneous secretion. Different volumes (5, 10 and 20 µL) of CS were analyzed by 12.5 % SDS-PAGE and stained with Coomassie Blue PhastGel™ R-350. MW-molecular weight marker; 5 µL- 5 µL of CS; 10 µL- 10 µL of CS; 20 µL - 20 µL of CS. All the samples were reduced in the presence of

β -mercaptoethanol and boiled for 10 minutes before application in the SDS-PAGE. (PPTX 189 kb)

Abbreviations

ACN: Acetonitrile; AMP: Antimicrobial peptides; CS: Cutaneous secretion; DADP: Anuran antimicrobial peptides database; DEPC: Diethylpyrocarbonate; EDTA: Ethylenediamine tetraacetic acid; ENA: European Nucleotide Archive; GO: Gene ontology; HDP: Host defense peptides; NCBI: National Center for Biotechnology Information; pident: Protein identity; PLA₂: Phospholipase A₂; PMSF: Phenylmethylsulfonyl fluoride; PVDF: Polyvinylidene fluoride; RNA-seq: RNA sequencing; RP-FPLC: Reversed phase fast protein liquid chromatography; SDS-PAGE: Denaturing polyacrylamide electrophoresis gel; SVSP: Snake venom serine protease; TCA: Trichloroacetic acid; TFA: Acid trifluoroacetic; TPM: Transcripts Per Kilobase Million

Acknowledgements

The authors are thankful for Karla de Castro Figueiredo Bordon, PhD, for the N-terminal sequencing, Iara Aimê Cardoso for technical support and biologist Luiz Henrique Anzaloni Pedrosa.

Funding

The São Paulo Research Foundation (FAPESP; grant n. 2015/18432-0 and scholarships to PYTS n.2014/15644-3, CTC n. 2013/26200-6 and GAW n. 201/06170-8) supported this work. This work was supported in part by the Coordination for the Improvement of Higher Education Personnel (CAPES) through Programa Editoração CAPES (edital n. 13/2016, auxílio n. 0722/2017, processo n. 88881.142062/2017-01) and by the National Council for Scientific and Technological Development (CNPq) through Programa Editorial CNPq/CAPES (process n. 440954/2017-7).

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the European Nucleotide Archive- ENA repository, <http://www.ebi.ac.uk/ena>. The accession number for the transcriptome is: PRJEB27393 and for the coding sequences mentioned in this paper is: ERZ693103.

Authors' contributions

PYTS was responsible for carrying out the transcriptome data analysis, toad's poison biochemical analysis and was a major contributor in writing the manuscript. CTC also was responsible for data analysis, design experiment and writing the manuscript. RMF performed the de novo assembling and FPM calculations. GAW performed part of the functional annotation and deposited the data. CUV supervised the in silico analysis and assisted in the project design. ECA coordinated the whole team, assisted in the experiments design and advised PYTS. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare there are no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Laboratory of Animal Toxins, School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, Avenida do Café s/n, Ribeirão Preto, SP, Brazil.

²Laboratory of genetics – LABGEN, Institute of Genetics and Biochemistry, Campus Umuarama, Federal University of Uberlândia, Avenida Pará, Uberlândia, MG 1720, Brazil. ³Department of Physics and Chemistry, University of São Paulo, School of Pharmaceutical Sciences of Ribeirão Preto, Av. do Café s/n°, Monte Alegre, Ribeirão Preto, SP 14040-903, Brazil.

Received: 4 July 2018 Accepted: 7 November 2018

Published online: 29 November 2018

References

1. Koehn FE, Carter GT. The evolving role of natural products in drug discovery. *Nat Rev Drug Discov*. 2005;4(3):206–20.
2. Newman DJ, Cragg GM, Snader KM. The influence of natural products upon drug discovery. *Nat Prod Rep*. 2000;17(3):215–34.
3. Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG. Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol Evol*. 2012;28(4):219–29.
4. Fry BG, Roelants K, Champagne DE, Scheib H, Tyndall JD, King GF, et al. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. *Annu Rev Genomics Hum Genet*. 2009;10:483–511.
5. Chan WY, Ng TB, Yeung HW. Examination for toxicity of a Chinese drug, the toad glandular secretory product chan su, in pregnant mice and embryos. *Biol Neonate*. 1995;67(5):376–80.
6. Frost DR. *Amphibian Species of the World: an Online Reference*. vol. 6.0; 2018. <http://research.amnh.org/vz/herpetology/amphibia/>
7. Regis-Alves E, Jared SGS, Maurício B, Mailho-Fontana PL, Antoniazzi MM, Fleury-Curado MC, et al. Structural cutaneous adaptations for defense in toad (*Rhinella icterica*) parotoid macroglands. *Toxicon*. 2017;137:128–34.
8. Jared C, Antoniazzi MM, Jordao AE, Silva JR, Greven H, Rodrigues MT. Parotoid macroglands in toad (*Rhinella jimi*): their structure and functioning in passive defence. *Toxicon*. 2009;54(3):197–207.
9. Clarke BT. The natural history of amphibian skin secretions, their normal functioning and potential medical applications. *Biol Rev Camb Philos Soc*. 1997;72(3):365–79.
10. Daly JW, Spande TF, Garraffo HM. Alkaloids from amphibian skin: a tabulation of over eight-hundred compounds. *J Nat Prod*. 2005;68(10):1556–75.
11. Sousa-Filho LM, Freitas CD, Lobo MD, Monteiro-Moreira AC, Silva RO, Santana LA, et al. Biochemical profile, biological activities, and toxic effects of proteins in the *Rhinella schneideri* parotoid gland secretion. *J Exp Zool A Ecol Genet Physiol*. 2016;325(8):511–23.
12. Sciani JM, Angeli CB, Antoniazzi MM, Jared C, Pimenta DC. Differences and similarities among parotoid macrogland secretions in South American toads: a preliminary biochemical delineation. *Sci World J*. 2013;2013:ID 937407.
13. Schmeda-Hirschmann G, Quispe C, Theoduloz C, de Sousa PT Jr, Parizotto C. Antiproliferative activity and new argininylyl bufadienolide esters from the "cururu" toad *Rhinella (Bufo) schneideri*. *J Ethnopharmacol*. 2014;155(2):1076–85.
14. Anjolette FA, Leite FP, Bordon KC, Azzolini AE, Pereira JC, Pereira-Crott LS, et al. Biological characterization of compounds from *Rhinella schneideri* poison that act on the complement system. *J Venom Anim Toxins incl Trop Dis*. 2015;21:25. <https://doi.org/10.1186/s40409-015-0024-9>.
15. Shibao PY, Anjolette FA, Lopes NP, Arantes EC. First serine protease inhibitor isolated from *Rhinella schneideri* poison. *J Venom Anim Toxins incl Trop Dis*. 2015;21:30. <https://doi.org/10.1186/s40409-015-0029-4>.
16. Freitas CDT, Sousa-Filho LM, Lima MSCS, Pereira ACTC, Ferreira GP, Lopes FES, et al. Proteins from the *Rhinella schneideri* parotoid gland secretion exhibit anti-nociceptive effect against nociception induced by inflammation. *Biomed Pharmacother*. 2017;93:705–8.
17. Kowalski K, Marciniak P, Rosinski G, Rychlik L. Toxic activity and protein identification from the parotoid gland secretion of the common toad *Bufo bufo*. *Comp Biochem Physiol C Toxicol Pharmacol*. 2018;205:43–52.
18. Nalbantsoy A, Karis M, Yalcin HT, Göcmen B. Biological activities of skin and parotoid gland secretions of bufonid toads (*Bufo bufo*, *Bufo verrucosissimus* and *Bufo variabilis*) from Turkey. *Biomed Pharmacother*. 2016;80:298–303.
19. König E, Bininda-Emonds OR, Shaw C. The diversity and evolution of anuran skin peptides. *Peptides*. 2014;63:96–117.
20. Boldrini-Franca J, Cologna CT, Pucca MB, Bordon KC, Amorim FG, Anjolette FA, et al. Minor snake venom proteins: structure, function and potential applications. *Biochim Biophys Acta Gen Subj*. 2016;1861(4):824–38.
21. Huang L, Li J, Anboukaria H, Luo Z, Zhao M, Wu H. Comparative transcriptome analyses of seven anurans reveal functions and adaptations of amphibian skin. *Sci Rep*. 2016;6:24069.
22. Yang W, Qi Y, Fu J. Genetic signals of high-altitude adaptation in amphibians: a comparative transcriptome analysis. *BMC Genet*. 2016;17:134.
23. Zhao F, Yan C, Wang X, Yang Y, Wang G, Lee W, et al. Comprehensive transcriptome profiling and functional analysis of the frog (*Bombina maxima*) immune system. *DNA Res*. 2014;21(1):1–13.

24. Andrews S. FastQC: a quality control tool for high throughput sequence data; 2010.
25. Chen TW, Gan RC, Fang YK, Chien KY, Liao WC, Chen CC, et al. FunctionAnnotator, a versatile and efficient web tool for non-model organism annotation. *Sci Rep*. 2017;7:10430.
26. Novković M, Simunić J, Bojović V, Tossi A, Juretić D. DADP: the database of anuran defense peptides. *Bioinformatics*. 2012;28(10):1406–7.
27. Schägger H, von Jagow G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem*. 1987;166(2):368–79.
28. Edman P, Begg G. A protein sequenator. *Eur J Biochem*. 1967;1(1):80–91.
29. Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res*. 1988;16(22):10881–90.
30. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;42:W320–4.
31. Wang WJ, Shih CH, Huang TF. A novel P-I class metalloproteinase with broad substrate-cleaving activity, agkistrolysin, from *Agkistrodon acutus* venom. *Biochem Biophys Res Commun*. 2004;324(1):224–30.
32. Severini C, Improta G, Falconieri-Erspamer G, Salvadori S, Erspamer V. The tachykinin peptide family. *Pharmacol Rev*. 2002;54(2):285–322.
33. Anastasi A, Montecucchi P, Erspamer V, Visser J. Amino acid composition and sequence of kassinin, a tachykinin dodecapeptide from the skin of the African frog *Kassina senegalensis*. *Experientia*. 1977;33(7):857–8.
34. Erspamer GF, Erspamer V, Piccinelli D. Parallel bioassay of physalaemin and kassinin, a tachykinin dodecapeptide from the skin of the African frog *Kassina senegalensis*. *Naunyn Schmiedeberg's Arch Pharmacol*. 1980;311(1):61–5.
35. Nakajima T, Yasuhara T, Erspamer V, Erspamer GF, Negri L, Endean R. Physalaemin- and bombesin-like peptides in the skin of the Australian leptodactylid frog *Uperoleia rugosa*. *Chem Pharm Bull (Tokyo)*. 1980;28(3):689–95.
36. Sakate M, Oliveira PCL. Toad envenoming in dogs: effects and treatment. *J Venom Anim Toxins*. 2000;6(1):52–62 www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-7930200000100003.
37. Carter MS, Krause JE. Structure, expression, and some regulatory mechanisms of the rat preprotachykinin gene encoding substance P, neurokinin A, neuropeptide K, and neuropeptide gamma. *J Neurosci*. 1990;10(7):2203–14.
38. Simmaco M, Mignogna G, Canofeni S, Miele R, Mangoni ML, Barra D. Temporins, antimicrobial peptides from the European red frog *Rana temporaria*. *Eur J Biochem*. 1996;242(3):788–92.
39. Ladram A, Nicolas P. Antimicrobial peptides from frog skin: biodiversity and therapeutic promises. *Front Biosci (Landmark Ed)*. 2016;21:1341–71.
40. Asoodeh A, Azam AG, Chamani J. Identification and characterization of novel antibacterial peptides from skin secretions of *Euphylyctis cyanophlyctis*. *Int J Pept Res Ther*. 2012;18(2):107–15.
41. Bilusich D, Jackway RJ, Musgrave IF, Tyler MJ, Bowie JH. The host-defence skin peptide profiles of Peron's tree frog *Litoria peronii* in winter and summer. Sequence determination by electrospray mass spectrometry and activities of the peptides. *Rapid Commun Mass Spectrom*. 2009;23(17):2628–36.
42. Park CB, Kim MS, Kim SC. A novel antimicrobial peptide from *Bufo bufo* gargarizans. *Biochem Biophys Res Commun*. 1996;218(1):408–13.
43. Cho JH, Sung BH, Kim SC. Buforins: histone H2A-derived antimicrobial peptides from toad stomach. *Biochim Biophys Acta*. 2008;1788(8):1564–9.
44. Selisko B, Garcia C, Becerril B, Gomez-Lagunas F, Garay C, Possani LD. Cobatoxins 1 and 2 from *Centruroides noxius* Hoffmann constitute a subfamily of potassium-channel-blocking scorpion toxins. *Eur J Biochem*. 1998;254(3):468–79.
45. Wang Q, Liu Y, He HJ, Zhao XF, Wang JX. Immune responses of *Helicoverpa armigera* to different kinds of pathogens. *BMC Immunol*. 2010;11:9.
46. Crava CM, Jakubowska AK, Escrache B, Herrero S, Bel Y. Dissimilar regulation of antimicrobial proteins in the midgut of *Spodoptera exigua* larvae challenged with *Bacillus thuringiensis* toxins or baculovirus. *PLoS One*. 2015;10(5):e0125991.
47. Bolouri Moghaddam MR, Tonk M, Schreiber C, Salzig D, Czernek P, Vilcinskas A, et al. The potential of the *Galleria mellonella* innate immune system is maximized by the co-presentation of diverse antimicrobial peptides. *Biol Chem*. 2016;397(9):939–45.
48. Narumiya S, Sugimoto Y, Ushikubi F. Prostanoid receptors: structures, properties, and functions. *Physiol Rev*. 1999;79(4):1193–226.
49. Lin ME, Herr DR, Chun J. Lysophosphatidic acid (LPA) receptors: signaling properties and disease relevance. *Prostaglandins Other Lipid Mediat*. 2010;91(3–4):130–8.
50. Murakami M, Kudo I. Phospholipase A2. *J Biochem*. 2002;131(3):285–92.
51. Zambelli VO, Picolo G, Fernandes CAH, Fontes MRM, Cury Y. Secreted phospholipases A2 from animal venoms in pain and analgesia. *Toxins (Basel)*. 2017;9(12):406.
52. Aird SD, Kaiser II, Lewis RV, Kruggel WG. A complete amino acid sequence for the basic subunit of crotoxin. *Arch Biochem Biophys*. 1986;249(2):296–300.
53. Bouchier C, Ducancel F, Guignery-Frelat G, Bon C, Boulain JC, Menez A. Cloning and sequencing of cDNAs encoding the two subunits of Crotoxin. *Nucleic Acids Res*. 1988;16(18):9050.
54. Bon C. Synergism of the two subunits of crotoxin. *Toxicon*. 1982;20(1):105–9.
55. Bon C, Changeux JP, Jeng TW, Fraenkel-Conrat H. Postsynaptic effects of crotoxin and of its isolated subunits. *Eur J Biochem*. 1979;99(3):471–81.
56. Canziani G, Seki C, Vidal JC. The mechanism of inhibition of phospholipase activity of crotoxin B by crotoxin A. *Toxicon*. 1983;21(5):663–74.
57. Habermann E, Walsch P, Breithaupt H. Biochemistry and pharmacology of the crotoxin complex. II. Possible interrelationships between toxicity and organ distribution of phospholipase a, crotopotin and their combination. *Naunyn Schmiedeberg's Arch Pharmacol*. 1972;273(4):313–30.
58. Rübtsamen K, Breithaupt H, Habermann E. Biochemistry and pharmacology of the crotoxin complex. I. Subfractionation and recombination of the crotoxin complex. *Naunyn Schmiedeberg's Arch Pharmacol*. 1971;270(3):274–88.
59. Qu XD, Lehrer RI. Secretory phospholipase A2 is the principal bactericide for staphylococci and other gram-positive bacteria in human tears. *Infect Immun*. 1998;66(6):2791–7.
60. Ben Bacha A, Abid I. Secretory phospholipase A2 in dromedary tears: a host defense against staphylococci and other gram-positive bacteria. *Appl Biochem Biotechnol*. 2013;169(6):1858–69.
61. Sudharshan S, Dhananjaya BL. Antibacterial potential of a basic phospholipase A2 (VRV-PL-VIIa) from *Daboia russelii pulchella* (Russell's viper) venom. *J Venom Anim Toxins incl Trop Dis*. 2015;21:17. <https://doi.org/10.1186/s40409-015-0014-y>.
62. Shilling FM, Krätzschar J, Cai H, Weskamp G, Gayko U, Leibow J, et al. Identification of metalloprotease/disintegrins in *Xenopus laevis* testis with a potential role in fertilization. *Dev Biol*. 1997;186(2):155–64.
63. Cai H, Krätzschar J, Alfandari D, Hunnicutt G, Blobel CP. Neural crest-specific and general expression of distinct metalloprotease-disintegrins in early *Xenopus laevis* development. *Dev Biol*. 1999;204(2):508–24.
64. Alexander G, Grothusen J, Zepeda H, Schwartzman RJ. Gyroxin, a toxin from the venom of *Crotalus durissus terrificus*, is a thrombin-like enzyme. *Toxicon*. 1988;26(10):953–60.
65. Rokyta DR, Wray KP, Lemmon AR, Lemmon EM, Caudle SB. A high-throughput venom-gland transcriptome for the eastern diamondback rattlesnake (*Crotalus adamanteus*) and evidence for pervasive positive selection across toxin classes. *Toxicon*. 2011;57(5):657–71.
66. Matsushita M. Ficolins in complement activation. *Mol Immunol*. 2013;55(1):22–6.
67. Stryer JMB, John LT, Stryer L. Lectins are specific carbohydrate-binding proteins; 2002.
68. Matsushita M. Ficolins: complement-activating lectins involved in innate immunity. *J Innate Immun*. 2010;2(1):24–32.
69. Yang RY, Rabinovich GA, Liu FT. Galectins: structure, function and therapeutic potential. *Expert Rev Mol Med*. 2008;10:e17.
70. Bianchet MA, Ahmed H, Vasta GR, Amzel LM. Soluble beta-galactosyl-binding lectin (galectin) from toad ovary: crystallographic studies of two protein-sugar complexes. *Proteins*. 2000;40(3):378–88.
71. Marschal P, Herrmann J, Leffler H, Baronides SH, Cooper DN. Sequence and specificity of a soluble lactose-binding lectin from *Xenopus laevis* skin. *J Biol Chem*. 1992;267(18):12942–9.
72. Hernandez-Oliveira S, Toyama MH, Toyama DO, Marangoni S, Hyslop S, Rodrigues-Simioni L. Biochemical, pharmacological and structural characterization of a new PLA2 from *Crotalus durissus terrificus* (south American rattlesnake) venom. *Protein J*. 2005;24(4):233–42.
73. Fritz-Laylin LK, Prochnik SE, Ginger ML, Dacks JB, Carpenter ML, Field MC, et al. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. *Cell*. 2010;140(5):631–42.
74. Sun T, Zhan B, Gao Y. A novel cathelicidin from *Bufo bufo* gargarizans Cantor showed specific activity to its habitat bacteria. *Gene*. 2015;571(2):172–7.