Pol. J. Environ. Stud. Vol. XX, No. X (XXX), 1–16
DOI:*10.15244/pjoes/188229* ONLINE PUBLICATION DATE:

Original Research

Transcriptomic Analysis of Venom Glands and Amino Acid Profile of Venom in Different Scorpion Species

Nevin Aslan[1]*, Şahin Toprak[1], İsmail Koyuncu[2], Akın Yiğin[3]

[1]Department of Biology, Faculty of Arts and Sciences, Harran University, Şanlıurfa, Turkey
[2]Department of Medical Biochemistry, Faculty of Medicine, Harran University Şanlıurfa, Turkey
[3]Department of Genetics, Faculty of Veterinary Medicine, Harran University, Şanlıurfa, Turkey

Received: 31 March 2024
Accepted: 2 May 2024

Abstract

Scorpion venom, with its constituents, has the potential to be a drug candidate. Scorpion venom is a mixture of biologically active compounds, including enzymes, toxic peptides, free amino acids, and other metabolites. In this study, the peptide maps and metabolic analysis of the venoms will be elucidated by transcriptomic analysis of mRNA profiles. Tissues were obtained from the venom glands of the scorpion species *Androctonus crassicauda*, *Hottentotta saulcyi,* and *Leiurus abdullahbayrami*. The mRNA was isolated from venom tissues obtained from the venom glands of the different scorpion species. As a result, 8 different proteins were obtained from the species *A. carsicauda* and *H. saulcyi,* and 7 different proteins from the species *L. abdullahbayrami*. According to the LC-MS/MS results, the amino acids taurine, alpha-aminobutyric acid, phosphoethanolamine, lysine, leucine, and arginine were detected in the scorpion venom of *A. crassicauda*. *H. saulcyi*; the amino acids taurine, lysine, phosphoethanolamine, arginine, alpha-aminopimelic acid, and leucine were detected in scorpion venom. *L. abdullahbayrami*; alpha- aminobutyric acid, taurine, lysine, phenylalanine, arginosuccinic acid, leucine, histamine, arginine, proline amino acids were detected in the scorpion venom. It is important to emphasize that this study not only adds to our existing knowledge of scorpion venoms, but may also serve as a basis for future research to define the composition of scorpion venoms and facilitate the identification of new putative toxin families.

Keywords: scorpion, venom gland, transcriptomics, metabolites, omics

Introduction

In order to adapt to their role as predator or prey, animals have developed various strategies. The ability to produce venom in specialized glands or cells and inject it into the target animal through a wound is one of these adaptations. In the animal kingdom, venom and toxic secretions are both complex combinations used for defense, predation, communication, and competition [1, 2]. Bioactive substances such as proteins, peptides, salts,

*e-mail: nevinaslan21@gmail.com

and metabolites such as amines, amino acids, and alkaloids are the various types of molecules that make up venoms. The venom of a single animal can contain hundreds of different components, each serving various biological purposes [3]. Studies of animal venoms have continued to evolve with technological advances, creating opportunities for the discovery of new bioactive molecules. With the introduction of "omics" technologies for venom research, such as proteomics, transcriptomics, genomics, and metabolomics, it has become possible to collect and analyze data on a large scale. One of the recently developed techniques for molecular phylogenetic analysis, biology, medicine, and especially for the identification of novel genes in non-model organisms is transcriptome sequencing [4]. Metabolomics, the most recent development in the field of 'omics' technologies, aims to identify and quantify the entire set of metabolites present in cells, biological fluids, and tissues (the so-called metabolome). The endogenous metabolites are at the end of a series of processes that begin with the genome and continue with the transcriptome and proteome. The metabolome thus closes the gap between a particular genotype and phenotype. The components of the metabolome include classes of compounds such as amines, amino acids, organic acids, steroids, alkaloids, and sugars. The term "metabolomics" was introduced in 1999 by Nicholson, Lindon, and Holmes [5]. Metabolomic and proteomic studies and the identification of new biochemical markers by LC-MS/MS, which are frequently used in diseases, can contribute to the development of preventive and therapeutic strategies [6].

Scorpions belong to the order Arachnid (Arachnida) and have attracted the attention of researchers due to their venomous components and the fact that they are ancient terrestrial animals that are also considered living fossils [7, 8]. The species *A. crassicauda* and *L. abdullahbayrami* are responsible for the majority of scorpion stings in our country, which is one of the endemic regions. In addition, the two most important species responsible for severe clinical findings and mortality are *A. crassicauda* and *L. abdullahbayrami* [9]. The toxicity of the venom of *H. saulcyi* is slightly lower than that of *A. crassicauda*. However, it is important for human health as it causes severe pain and sweating in scorpion stings [10]. *A. crassicauda* is a common scorpion species in Turkey and can cause serious poisoning in humans. *A. crassicauda is* reported to occur mainly in Eastern and Southeastern Anatolia, in the provinces of Elazığ (Palu), Diyarbakır, Şanlıurfa (Harran – Akçakale), Mardin, Adana, Hatay, Malatya and Mersin in Turkey [10, 11]. The venom of *A. crassicauda* contains toxins that are especially effective on sodium channels [12]. The venom of *H. saulcyi* is slightly less toxic than that of *A.crassicauda*. Nevertheless, it is important for human health as it causes severe pain and sweating in scorpion stings. *H. saulcyi* is generally distributed in the provinces of Mardin, Batman, Şırnak and Hakkari in the Southeast Anatolia region of Turkey [13]. *L. abdullahbayrami* is endemic to Turkey and occurs in and around the provinces of Gaziantep, Adıyaman, Kilis, Şanlıurfa, and Mardin provinces in southeastern Anatolia [14]. Scorpion venom consists of a combination of many biologically active substances, including enzymes (such as hyaluronidase, phospholipase, and protease), carbohydrates, free amino acids, lipids, toxic peptides, and other metabolites [15, 16]. Since scorpions use their venom to capture prey, protect themselves from predators, and reproduce, some venom components are essential for the scorpion's survival [15]. Scorpion venom is a complex mixture of biogenic amines, small peptides and large proteins, mucoproteins, and other components produced by the columnar epithelial cells of the venom gland [17].

Mainly due to the heterogeneity, complexity, and high variability of the venom composition and the low solubility of existing methods to separate the venom proteins separation methods, the amount of venom peptides discovered is much lower than expected [18–20]. The first cDNA library for scorpion venom glands was created in 1989 for the scorpion *Androctonus australis*. This library contains collections of cDNA sequences cloned into vectors [21]. Subsequently, transcriptomic analyses of the scorpion venom gland have been greatly improved by the introduction of high-throughput sequencing.

The aim of this study was to perform transcriptomic analysis of mRNA transcripts in tissues collected from venom glands of the scorpion species *A. crassicauda*, *H. saulcyi*, and *L. abdullahbayrami* and to analyze the metabolic venom. As a result of the transcriptome study of these scorpion species, the amino acid sequences in the possible proteins in the venom glands are determined and the venom contents and the amino acids in the venom are compared by analyzing the profile of free amino acids profile analysis. The peptide maps and metabolic analyses of the venoms will be elucidated by transcriptomic analysis of the mRNA profiles of tissues from the venom glands of the scorpions *A. crassicauda*, *H. saulcyi*, and *L. abdullahbayrami*.

Material and Method

Sample Collection and Venom Milking

The species *A. crassicauda* and *L. abdullahbayrami* were collected in the province of Şanlıurfa and the species *H. saulcyi* species were collected in the province of Şırnak. During the field studies, a total of 30 scorpions, 10 of each scorpion species, were collected. The scorpions were captured during the day and at night during field surveys. The captured specimens were placed in plastic containers with lids and brought to the laboratory. The scorpions brought to the laboratory were fed with mealworms once a week to adapt to the environment (Fig. 1). The specimens collected in the field were injected with venom. The venom was milked using the electrical stimulation method by applying current in the range of 12–16 volts. The telson of the scorpions was cut off 3 days after milking. The venom and the telson of the milked scorpions were stored in a freezer at -80°C until the required analyses were performed.

Fig. 1. Adaptation of samples collected from the field in the laboratory environment (A: *A. crassicauda,* B: *H. saulcyi*, C:*L. abdullahbayrami*).

mRNA Isolation and cDNA Library Preparation

Commercial miRNeasy kits (QIAGEN:217004) and a total RNA isolation kit were used for transcriptome analysis of the venom tissues obtained. At the beginning of the isolation, the venom glands of each scorpion species were cut open and placed in separate microcentrifuge tubes. The isolation steps were performed according to the protocol of the kit. RNA concentration and quality control were performed using the RNA 6000 Pico Kit (Agilent Technologies, USA) on an Agilent Bioanalyzer 2100.

RNA Quality Control

For quality control of the isolated total RNA material, measurements were performed with a 4200 TapeStation system (Agilent, USA, #G2991BA).

Library Preparation (mRNA Preparation)

The Illumina Stranded mRNA Preparation Kit (Illumina, USA, 20040534) was used for library preparation of the isolated RNA material. The kit is based on the separation of targeted mRNAs in the whole RNA from the poly-A tails, random fragmentation, and subsequent generation of primary and secondary cDNA chains. In library preparation, capture, fragmentation, mRNA purification, cDNA synthesis, addition of index barcode sequences, and purification were performed according to the manufacturer's instructions.

Sequencing (Illumina NovaSeq 6000)

The NovaSeq 6000 Next Generation Sequencing platform from Illumina was used for sequencing. 150 bp paired-end reads were obtained. Quantification, dilution, and loading of the library was performed according to the manufacturer's instructions.

Examination of the Metabolite Profile of the Venom

The venoms of the species *A. crassicauda*, *H. saulcyi*, and *L. abdullahbayrami* were removed from the -80°C freezer, diluted 1:1 with ultrapure water, and vortexed for 5 minutes. The JASEM amino acid kit was used to analyze the amino acids. For the analysis, it was placed in the tray compartment in the HPLC section of the LC-MS/MS (Shimadzu 8045, Japan) and read. Mobile phase-A and mobile phase-B from the amino acid kit were used as the mobile phase. Restek LC columns were used as columns.

Bioinformatics Analyses

Adapter trimming was necessary because we found nextera transposase sequences above the threshold (>1%) in the read2 fastq files. We performed the trimming using the fastp v0.23.4 tool [22], using the adapter sequences listed in the adapter.fa file in QC_reports and also trimming for quality (min Q30 at 5' and 3' ends), minimum length (100 bp), and ambiguous base concentration. The adaptor sequences used for trimming were the transposase sequence

of nextera (R1 TCGTCGGCAG CGTCAGATGTGTATA AGAGACAG,R2GTCTCGTGGGCTCG-GAGATGTG-TA TAAGAGACAG). The trimmed, filtered reads were mapped to the coding reference sequence of *C. sculpturatus* (https://www.ncbi.nlm.nih.gov/datasets/genome / GCF_000671375.1), the sequenced scorpion species that has gene information. The mapping tool used was BWA (v. 0.7.17-r1188) [23]. We counted the reads mapped to each cds using samtools (v1.15.1) [24] idxstats and calculated the TPM value using the formula with R [25] (version 4.2.1) base functions:

$$TPM = \frac{10^6 \times \text{read count mapped on transcript/transcript length [kb]}}{\sum (\text{read count mapped on transcript/transcript length [kb]})}$$

Results

Investigation of Metabolite Profile of Venoms

Scorpion venom contains macromolecules with a molecular weight of over 5000 Daltons and small micromolecules such as various peptides and enzymes. In this study, the amounts of free amino acids in the venoms of *A.crassicauda*, *H. saulcyi*, and *L. abdullahbayrami* were determined by LC-MS/MS method. The free amino acid profile of the venom was analyzed by LC-MS and the results are presented in Table 1. The presence of the amino acids taurine, alpha-amobutyric acid, phosphoethanolamine, lysine, leucine, and arginine was detected in the scorpion

Table 1. Results of LC-MS/MS analysis of the free amino acid profile of the venoms of *H. saulcyi, A. crassicauda*, and *L. abdullahbayrami* venoms. (Amount detected in the venoms (nmol/L).

Amino acid name	*H. saulcyi* (nmol/L)	*A.crassicauda* (nmol/L)	*L. abdullahbayrami* (nmol/L)
Alanine	17.3	18.3	163.6
Arginine	58.9	49.5	371.9
Asparagine	11.4	15.0	159.1
Aspartic Acid	12.5	4.6	2.9
Citrulline	4.7	3.2	39.2
Glutamine	25.1	43.7	101.5
Glutamic Acid	14.3	7.6	144.2
Glycine	24.6	10.9	14.2
Histidine	4.2	2.8	28.8
Leucine	29.1	56.4	511.6
Isoleucine	20.1	10.6	87.4
Alloisoleucine	0.0	0.0	0.4
Lysine	173.2	70.5	948.2
Methionine	12.3	7.6	156.1
Ornithine	1.5	0.7	54.8
Phenylalanine	28.5	31.5	608.6
Proline	26.1	25.7	175.0
Serine	19.1	11.1	144.4
Threonine	3.4	12.9	29.5
Tryptophan	6.8	5.5	53.9
Tyrosine	6.6	9.1	39.6
Valine	9.4	9.1	88.5
Alphaaminoadipic Acid	1.4	1.2	1.5
Alphaaminopimelic Acid	39.8	39.6	32.3
Anserine	0.0	0.0	0.0
Argininosuccinic Acid	7.2	0.9	601.8

→

Amino acid name	*H. saulcyi* (nmol/L)	*A.crassicauda* (nmol/L)	*L. abdullahbayrami* (nmol/L)
Alphaaminobutyric Acid	3.4	155.4	2050.8
Betaaminoisobutyric Acid	1.6	1.1	4.8
Gammaminobutyric Acid	1.8	3.4	4.6
Beta-Alanine	0.3	0.5	0.0
Sarcosine	0.0	0.0	0.0
Cystathionine	0.2	0.0	4.8
Thiaproline	0.0	0.0	0.0
1-Methylhistidine	0.9	0.0	127.3
3-Methylhistidine	0.0	0.0	0.0
Hydroxylysine	4.7	0.4	119.2
Hydroxyproline	6.4	12.0	90.9
Cystine	0.0	0.0	1.1
Homocystine	0.0	0.0	0.0
Serotonin	0.0	0.0	54.0
Histamine	0.0	0.1	432.2
Etanolamine	0.3	0.2	0.6
Phosphoetanolamine	100.5	84.4	4.1
5-Oh-Trp	0.0	0.0	0.0
Taurine	251.7	251.8	1417.3

venom of *A. crassicauda*. The presence of the amino acids taurine, lysine, phosphoethanolamine, arginine, alpha-aminopimelic acid, and leucine was detected in the scorpion venom *of H. saulcyi*. LC-MS/MS results: The presence of alpha-aminobutyric acid, taurine, lysine, phenylalanine, arginosuccinic acid, arginosuccinic acid, leucine, histidine, arginine, pro-line amino acids were determined in *L. abdullahbayrami* scorpion venom.

The amino acid distribution profiles of the metabolite analyses of the venoms of *H. saulcyi*, *A. crassicauda*, and *L. abdullahbayrami* species are shown in Fig. 2.

Transcriptome Analyses of the Venom Gland

The general statistics of the venom gland transcriptome analyses are given in Table 2. Repeat reads were 81.2% for *A. crassicauda*, 82.5% for *H. saulcyi,* and 75.9% for *L. abdullahbayrami*. GC rates were 37% for *A. crassicauda*, 36% for *H. saulcyi*, and 37% for *L. abdullahbayrami*. Read lengths were 134 bp for *A. crassicauda*, 133 bp for *H. saulcyi,* and 134 bp for *L. abdullahbayrami*, respectively. The total number of sequences was 41.6, 27.0, and 20.7 million for *A. crassicauda*, *H. saulcyi*, and *L. abdullahbayrami*, respectively.

The 8 different protein names, ID codes, similarity ratios, and amino acid sequences obtained as a result of transcriptome analysis in *A. crassicauda* species are given in Table 3. The longest amino acid sequence contains 974 amino acid sequences in colorectal mutant cancer protein-like isoform X1 protein and the shortest amino acid sequence contains 189 amino acid sequences in apoptosis regulator R1-like protein.

The 8 different protein names, ID codes, similarity ratios, and amino acid sequences obtained from the transcriptome analysis of the species *H. saulcyi* are listed in Table 4. The longest amino acid sequence contains 910 amino acid sequences in breast cancer anti-oestrogen resistance protein 3-like protein and the shortest sequence contains 189 amino acid sequences in the apoptosis regulator R1-like protein.

The names, ID codes, similarity ratios, and amino acid sequences of 7 different proteins obtained as a result of transcriptome analysis *L. abdullahbayrami* species are listed in Table 5. The longest sequence contains 970 amino acid sequences in colorectal mutant cancer protein-like protein isoform X2 protein and the shortest sequence contains 189 amino acid sequences in apoptosis regulator R1-like protein.

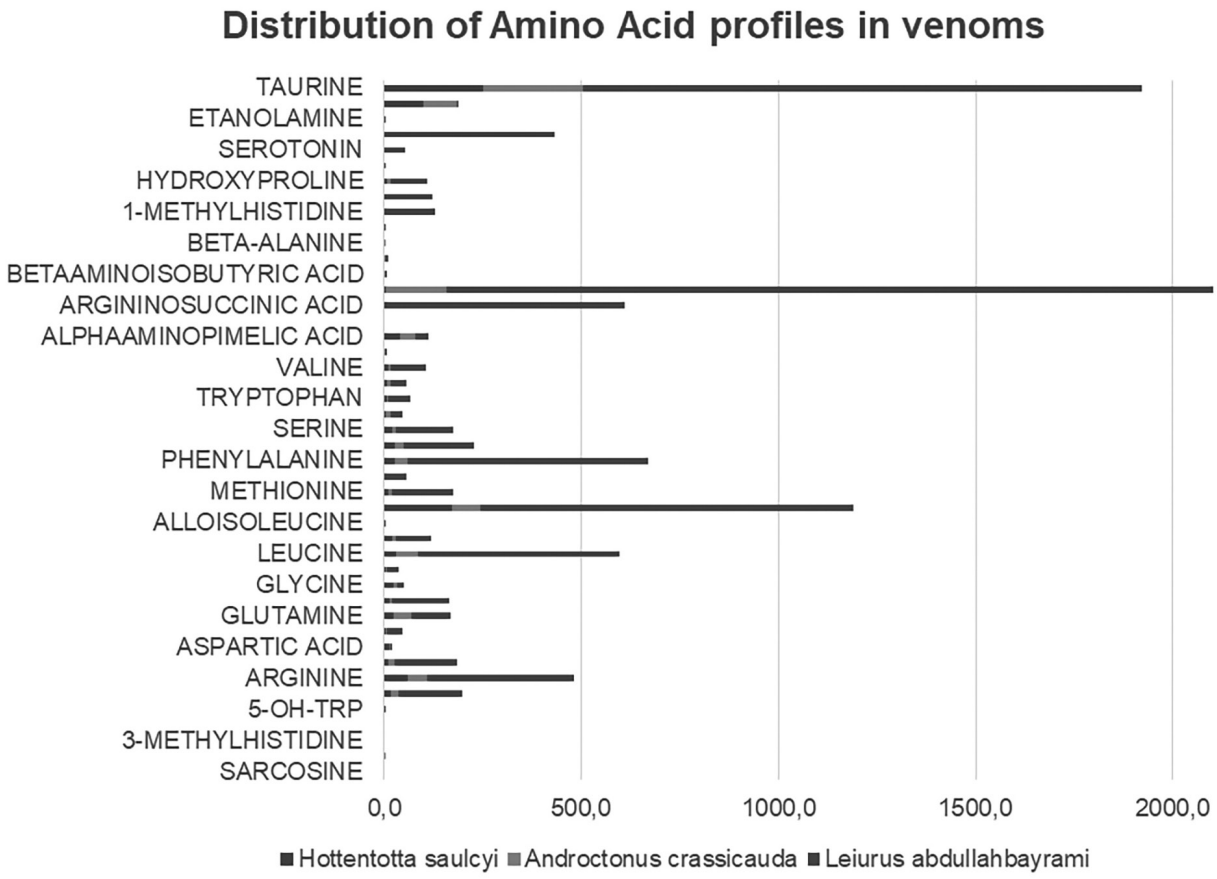

Fig. 2. Distribution of A.A. profiles in metabolite analyses of venoms from the species *H. saulcyi, A. crassicauda,* and *L. abdullahbayrami.*

Table 2. General statistics of venom gland transcriptome analyses in different scorpion species (AC: *A.crassicauda*, HS: *H. saulcyi*, LA: *L. abdullahbayrami*).

Sample Name	% Dups	% GC	Read Length	M Seqs
RP23-025-AC	81.2%	37%	134 bp	41,600,000
RP23-025-HS	82.5%	36%	133 bp	27,000,000
RP23-025-LA	75.9%	37%	134 bp	20,700,000

Table 3. 8 different protein names, ID codes, similarity ratios and amino acid sequences obtained as a result of transcriptome analysis in *A. crassicauda.*

	Protein	ID Code	Similarity	Amino Acid Sequence
1	Breast Cancer Anti-Estrogen Resistance Protein 3-Like	XP_023215813.1-	%97.74	MEPMIGTEASPEELKKALEWELSLSNSDLRSHAWYHGTI PRQRAEELMSNDGDFLIRDCISHPGDYVLTCKWKNSSLH FVINKVVFQPFTVYEKVQYQFEEDSFDTVPDLVTYYVGNK RAVSAASGAIINCPVNRTMPLTYYASRYGLQCQLHYAALAAET DIREADRSPCLARRHSQPCEDGSLSAESHPGTFNKTATSSVAG FATLRPHHQPAGWKGMARIGSDPQLSPTQERDGPPPKPSRLP SQHVYTEPIYKDGNAPSNGHSPNINITDMSTPNLNPPSCYDLNK FATTLLQTENKPLLKEPMIGTEASPEELKKALEWELSLSNSDLR SHAWYHGTIPRQRAEELMSNDGDFLIRDCISHPGDYVLTCKW KNSSLHFVINKVVFQPFTVYEKVQYQFEEDSFDTVPDLVTYY VGNKRAVSAASGAIINCPVNRTMPLTYYASRYGLQCQLHYAAL AAETDIREADRSPCLARRHSQPCEDGSLSAESHPGTFNKTATSS VAGFATLRPHHQPAGWKGMARIGSDPQLSPTQERDGPPPKP SRLPNQHVYTEPIYKDGNAPSNGHSPNINITDMSTPNLNPPSCY DLNKFATTLLQTENKPLESSTLLQIRRLLLENGPRILANHLTRIDLDM

→

	Protein	ID Code	Similarity	Amino Acid Sequence
				LKNSTDFDFGLGVVSGLELITLPQGKQLRMDLMERN HCLRYFVAVTILTCSNEEERANLVHKWIQIAIETKTALGN LYGFTAIMQGLALQQIDRLRSTWLTVRQNFTENAFTYETKL RPTLKSMQECSNPQAPNTCIPYLLSLITILQRHIEVIENNEL NAEPQVEPSGKKTVNILSLGLQWEQSASDYGLQLLLLHLEL GRTFVQQCTTYRRNGEIVLDNVKFDNSILDMFKTEFHLKFLWG SKGANVAAAERYTKFDQVLQVMSERCESS
2	BRCA2-interacting transcriptional repressor EMSY-like isoform X1	XP_023216198.1	%81.19	MMWPMLLDFSRDECKRILRKLELEAYAAIVSAFRAQGELTKDK KKLLQELSTVLSISLERHRAEIRRAVNDERLNTIADRIYGPNTS VDWAIEGRRLIPLLPRLVPQTAFTAVANSVASIQAAKNATMPLP SASGVKEGAIPTTSSAPPTPSRTPTPTQTCRVSLPSSIPLKTTS VNGNNSGNLTGGIITRLPQETNEENQSEDFNGKKRKRSASLD SASLPEKITISSNPLVSDENQAPEKPVTTTNNGMLCSSTPVT TQATITRTFATTPIRITLPSSQKQNAAVSSSAIPKQVVLQQLQTS NASTTANVFQRSVSIPVVKTVSATATTITQKQSKQTVSPS VQGTLLLTGKVSTTTTTGQSHPLNVPIATVGTIARTRPKTLPP PRQSIRPRSSVLSINQNRMQFVGQNFTSQLPTPVAKSSLQIGT PIQVKQVSSDGRTVQIRQEGGVKIVAQGLPVAASKILPKPST SPIVMVSSTQGSNASKFSASSTVIQNQIGTKMVSIASPQN NQSSNKVTAIGNVSLVSRTSQAFSTRTGSAVVQSPGVKPNVIV VHKAQVWPQSQGQGTTIVVSGNPVPKLSTETVQETVTYIQK SEGGRTNSSPAVVSVSSCNPANISALNSTAITNFEQSEKNESST VNQENKTTLNVNSKQADEKGEIDKKNNLLADVIEASGILS DGTESFPPSINAKSALSNQEEKISEELEEPTETVVIEKETEVL NVEETWTDESRLQEIAPLENDEKQNEEVLNVDTSLSLEDIN LPLSQLEDGQILEIQTQDDAGVLSESTRITKEMLELLQQAGL HIQEIEVKSDLSNDSYQGSVPVKSTNEEEAVEMKNKES CELLKHEDNDSCEKLILCQQDEYESSEVDKTINESTQSKNIQSNF
3	Apoptosis regulator R1-like	XP_023217360.1	%96.30	MAEHSTEEMEVSCIVNDYISYQLHMNGYEWRSPTATPTKVNS SLRSLGNEFKTRYRDSFLEMTDSLDIQPSTARATFLGVANELF SEGVTWARIVALFVFASEMALICHRMSHTEIVNDVAEWLSSYIK SNLLTWIRDHDGWDGLVTFHEGYDDLRSESPWPVLKKFVCG VAAGVIGALTVGAILSSKS
4	Venom phosphodiesterase 2-like	XP_023218281.1	%95.56	M M E R Q T L L A L V A F M V V A I V V I F I G G Y F C G V N VAYAQRGQKHLNTGPSSEKLTEQTGPWYQQPCFPERNICPS DFRNPPLLLISLDGFRPDYFDKTPTLKRLAHCGASAPYMK PVFPTKTFPNHYSIVTGLYPESHGIIDNNIYDPKLRRLFSVR RGQDNYNPIWWEAEPLWVTAEKQGKITGTYFWPGSEIKING TRPTYYIKYNNKVSWKTRLDTILSWLELPEESRPSFLTLYINEP DHASHAHGPESPQVNAVLAKLDSFIELMITSLQQRNLLGCI NIIIVSDHGMAKSDCNKVIELDNYINSSEVFLSPGPIGKIRPK GKRTVESVVEILRCNRPELLVYKKNDLPKRYHYSYNNKIEPII IDLPTGWAIENKFDARRCRGGNHGYDYLHPEMNAIFLAFGPS FKKNLLVEPFINIELYELMAELIDVIPNPNNGSKGSLHYILNNPK TLSASSMHKAITACVITNLQSSKFKGCNCPNGSLHQISPPS DFKHKQEVILPWGAPMIKDSEFTSICHLMNPDYVTAFHKKLKQP SWISFPLTRKTSIMSGEARFFHGKCWILDQRIPFENQANCTDYSIL NRDHSAIQQRSLFPPAFAISENSESSVQLMTNAVPMYKTFHSE IWTQFLILLSKWTQKYDSLNIIMGPAFDERGNGLKMSVGEL LERHNNKIPIPTHFFVVITRCEASNVPLNSCRSADLDVLSF LIPHKPFPENCQNINEYFLKHSATVKDIEVLTGLKFFTSLSPY TAIALQTRLTDYEWFNVK
5	Serologically Defined Colon Cancer Antigen 3 Homolog	XP_023224251.1	%73.17	MENGAKSKDNPFSFKNFLKGSQDIDSNPGRRSVNEDADNPFS FRKFLSQGKSSETKRTYNQNVESSRSQDTASSSPKLSSQTTTSSS FEIYSDFVQELPDFIQDHLELKYEDDKIESNEHLYNSNEHYT HTDYSRETYGNSGSSIESGQAEASLESFNGSSVSTSHNES RIENQIKNESYCGLPDFISGTSKSETVATWSSSHLEESNCIGH SQSSLEDEIYKLREENAFLHEQLEKARRTSDIQASKIVELQQK MIELREAEAQETAALEKMVQQVELNLKVTTERAVLAEGQAAK LKQELKSLQALFKNVMKENDKLTGSSSDMECIAEKANSLAS QIKVAADNAEQPLRNLLEGVEHLRLLATLVSSIGKVEEN

→

	Protein	ID Code	Similarity	Amino Acid Sequence
6	C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8	XP_023212124.1	%68.42	MPVGCGEQNMMYMGPTLYTLRFLKVKGDVTPEREGKIYNFIR EGYNRQLTYRKEDGSYGAWIQTPSSTWLTAFTMKVFCQANK FIFVDENVICSGIKWLISHQKDDGAFEDAHPVYHYDMTGGVS GRIPMTAFVLIALQECSSCESEELNVARLKAELFLEISLPTIND PYIMAVVAYALSQSSSDLKHEANEKLKQMSIFDEDLNQRHW KENTRSHSIEVAAYALLTQLALNDLSYSLPIVNWLNTQRLVGGS FPSTQDTVIALQALSQYSMQSQSPDLSLMCNVTSSNDRNLAK TLEFKNDNALVLQQFEVNKVGGTLFFRTAGNGVGSLAVKL RYNVPVPVEKLCKVSRNIKSIIKTLLKIKII
7	Colorectal mutant cancer protein-like isoform X1	XP_023227728.1	%98.30	MASDGNSESSGSDSGSTPEEERIRRLFQTCDGNGDGYIDSQDL LAVCRQLNLEHCVEEIMEQLGADEQGRISYSEFLRRRMQLM NEINALTMQDQLQDPEHRTNTTDIHVAPPQQTSPAPGSGSTGH WPTGSDTSQGAASGKHESWEFDSGARDLSPEPNTLQKLIEAA GGSISGNSSDLLDLANKLHLAALASLKGEILELTNR LQHIGQERDMLERQLNRSQLDRLRLAREHEERLDQQAQRY EERLTELHSIIAELSKKLEIQRTNVIKEEDEFSQQSGEESRQ SEAQSHCTSNDNNATESCSDRDFFLGNDSCVELSRFGETD SHPEVTRTILQDGEEASSGLGDNESSEDRLLQLREVSDLPAL TENRATSCHSLQISQLQEEVLSLRAENTALQEQLGRQEAELNR TRASLASSREERDRLQRRIRELQTKLASLQSQSSQGSQTGSSP TRSKSPHINPTTGERTPTSREEVPIAKMAERVRLKKVES GDRHILGSEISSLGLSTTKVAEHLVHNVQEESHAQEIFQTLF SSGSTVPESKIREFEVEMERLNSRIEHLKSQNDLLNLTLEESK GQCDRLSVLIGKYESNNTALQLALSYSDQTIEAFEVLLAL METEQGLILSNCRAAGLGALVKGSGEDDQEEVTALLKRSQD NRKSAENVARHFLQKLDRSFGIACSIPGCNLSPWEDLSSH SHTTSTTSSASSSSDGDFTKVDEQRLRDYIVQLKGDRAAVRT TVVELESVHVDPTVKDPPTFMDAQKLDLENAVLMQEL MAMKEEKAELKAQVYLLEKEKTALELQLSSREAQEQAY LVQIEHLKSEVKENQQIAERLRTTQGGKVNNQQPWLER SPSGEMAEELAEAMKREMKMKARIQELVNTLEKVTKNSEM RAQQSTEFVNDLKRANSALVTAFEKAKKKQQAKLKKVEQQ MASMVERHSAQIRMLKQRIAMLEGETGHLMQTSNSDLQV
8	Venom metalloproteinase 3-like isoform X1	XP_023234770.1	%70.35	MSFAVIFFGLATLWICGESKFEKFHERLSQEELRKIFHVDLHN NVPEYDVVHVRSLSKRSIPSDDTKRVHLSAFGQNFHLNLKRN HELEDRLKSMKLFAAETKGNELRYKEMEPQEHAISDMGVSYH DENQMAALLVNHAEDGSLRLEGTIGNHLVIKPIPKQLIVVEDNY VDDEMFLDEDGNTSTTHDMQRHLPTLSSVNHVVYKRKQNF PLSHSDYMELEAGYANNSDWTIRSRRTKRKAPTTVWPEVLV VIDYNTFLLHGSDSRALKRYFVSFWNGVDLRYKVLANPKVKIS LAGMVVAKDKDATPYLEKNRLRPPNYDAVDAAGALSDMGKY LYREDRLPTFDLAVVITKLDMCRRQFNGGRCSRGTAGFAYVG GACVVNKRLEKVNSVAIIEDSGGFSGIIVAAHEVGHLLGCVH DGSQPPSYLGGPGASNCPWEDGFIMSDLRHTDRGFQWSAC SIQQFKHFLQGETAVCLYNYPHDNQILSRVLPGTTMSLDDQCRK DRGTTACFKDARVCAQLFCYDTSSGFCVSFRPAAEGSTCGH GQVCRDGRCVGETENIIPDYTHYTPTYAGRHPNFVYRRTDR RGTTPYISRTTREIITRTTPSQSPSSCEDNAQQLAGGLTCTE FLQRYGFRYCKHRYMKQKCCKSQLKLCSHS

Table 4. 8 different protein names, ID codes, similarity ratios and amino acid sequences obtained from the transcriptome analysis of *H. saulcyi*.

	Protein	ID Code	Similarity	Amino Acid Sequence
1	Venom protease-like	XP_023210205.1	%93.33	MLHELFLLLVAVSFLMTSTLSSETVRERQLFVVRRRPPYQLV NNVFGRVNECRMVDGSLGECSHIARCTFMLFDLRRAFCLRNY LIPGVCCPVKNQTKASLVQSTPVPVTPTSVPIVIFTYPPAFQPVT VPALEQCGISKTISRIVGGFESAPSQWPWMAAIFLNTHRG REYWCGGALIDSRYVVTAGHCLSDLREQKYRAKEMSVRLG DHHLFFRGEDANPVDYRVAKVIQHPNFSRHGFFNDIGLLRLE KEVSFDDRIRPICLPERNSSGKNLVGMMATVIGWGTVNYGGAS SGSLHQVSIPIWDNADCDKRYFQPITRGFVCAGFRAGGKDA CQGDSGSPLMLPDNDRRWTIVGLVSFGSRCAQEGYPGVYTRV DTYLDWIRNHTGR
2	Venom metalloproteinase antarease-like TtrivMP_A	XP_023212236.1	%81.25	MFVYLTISLFITITSAVPTGREDVVYPSVETLRSGIKRIKFRAI DQDIDLRLIPAGELISDDFVIYEGDHERRQPTEKIKSLKKKLY RDSEKGAALHIEEDGFTSIHGIISKKLRIEPEETITEGRRAHRV FEVENMERGRIRTNDALMIPPGAIRNVTEAEKRSGICIVVEIVIV SEKTFTHSFGGKYDETAEYYHLRYLYVTITQVQNLMDTMK LTLKVLLIGVITYDNTNEASFIQKSAIKICPNYLDSGDLIENMA EYYKNAKIPLIDEADGVMLVTLTYGIIYVSILMNDDITPVDTAN TIRPDPYLSRNLAECYVYKADKKLSTGYIAIMFVYLTISLFITIT SAVPTGREDVVYPSVETLRSGIKRIKFRAIDQDIDLRLIPAGE LISDDFVIYEGDHERRQPTEKIKSLKKKLYRDSEKGAALHIEE DGFTSIHGIISKKLRIEPEETITEGRRAHRVFEVENMERGRIRTN DALMIPPGAIRNVTEAEKRSGICIVVEIVIVSEKTFTHSFGGKY DETAEYYHLRYLYVTITQVQNLMDTMKLTLKVLLIGVITYD NTNEASFIQKSAIKICPNYLDSGDLIENMAEYYKNAKIPLIDEA DGVMLVTLRPMADVYKRKCDLGTIGIAYVGSVCDNGYKYGI SEDDEGNFYEYCHTFAHELAHMIACPHDEDSPVSYIPGNPGT VECKWKYGYIMSYETNEQNGTKFSPCSRACVQQFVSLSMAS CIVEEC
3	Breast cancer anti-estrogen resistance protein 3-like	XP_023215813.1	%98.08	MEPMIGTEASPEELKKALEWELSLSNSDLRSHAWYHGTIPRQ RAEELMSNDGDFLIRDCISHPGDYVLTCKWKNSSLHFVINKV VFQPFTVYEKVQYQFEEDSFDTVPDLVTYYVGNKRAVSAAS GAIINCPVNRTMPLTYYASRYGLQCQLHYAALAAETDIREAD RSPCLARRHSQPCEDGSLSAESHPGTFNKTATSSVAGFATLRPH HQPAGWKGMARIGSDPQLSPTQERDGPPPKPSRLPSQHVYTE PIYKDGNAPSNGHSPNINITDMSTPNLNPPSCYDLNKFATTLLQ TENKPLLKEPMIGTEASPEELKKALEWELSLSNSDLRSHAWYH GTIPRQRAEELMSNDGDFLIRDCISHPGDYVLTCKWKNSSLHF VINKVVFQPFTVYEKVQYQFEEDSFDTVPDLVTYYVGNKRAV SAASGAIINCPVNRTMPLTYYASRYGLQCQLHYAALAAETDI READRSPCLARRHSQPCEDGSLSAESHPGTFNKTATSSVAGFAT LRPHHQPAGWKGMARIGSDPQLSPTQERDGPPPKPSRLPNQHV YTEPIYKDGNAPSNGHSPNINITDMSTPNLNPPSCYDLNKFATT LLQTENKPLESSTLLQIRRLLLENGPRILANHLTRIDLDMLK NSTDFDFGLGVVSGLELITLPQGKQLRMDLMERNHCLRYF VAVTILTCSNEEERANLVHKWIQIAIETKTALGNLYGFTAIMQG LALQQIDRLRSTWLTVRQNFTENAFTYETKLRPTLKSMQECS NPQAPNTCIPYLLSLITILQRHIEVIENNELNAEPQVEPSGKKT VNILSLGLQWEQSASDYGLQLLLLHLELGRTFVQQCTTYRRN GEIVLDNVKFDNSILDMFKTEFHLKFLWGSKGANVAAAERYT KFDQVLQVMSERCESS
4	BRCA2-interacting transcriptional repressor EMSY-like isoform X2	XP_023216199.1	%100	MWPMLLDFSRDECKRILRKLELEAYAAIVSAFRAQGELTKDK KKLLQELSTVLSISLERHRAEIRRAVNDERLNTIADRIYGPNT SVDWAIEGRRLIPLLPRLVPQTAFTAVANSVASIQAAKNATMP LPSASGVKEGAIPTTSSAPPTPSRTPTPTQTCRVSLPSSIPLKTT SVNGNNSGNLTGGIITRLPQETNEENQSEDFNGKKRKRSASLD SASLPEKITISSNPLVSDENQAPEKPVTTTNNGMLCSSTPVTTQA TITRTFATTPIRITLPSSQKQNAAVSSSAIPKQVVLQQLQTSNAST TANVFQRSVSIPVVKTVSATATTITQKQSKQTVSPSVQGTLLLT GKVSTTTTTGQSHPLNVPIATVGTIARTRPKTLPPPRQSIRPRS

→

	Protein	ID Code	Similarity	Amino Acid Sequence
				SVLSINQNRMQFVGQNFTSQLPTPVAKSSLQIGTPIQVKQVSS DGRTVQIRQEGGVKIVAQGLPVAASKILPKPSTSPIVMVSSTQ GSNASKFSASSTVIQNQIGTKMVSIASPQNNQSSNKVTAIGNVS LVSRTSQAFSTRTGSAVVQSPGVKPNVIVVHKAQVWPQSQGQ GTTIVVSGNPVPKLSTETVQETVTYIQKSEGGRTNSSPAVVSVS SCNPANISALNSTAITNFEQSEKNESSTVNQENKTTLNVNSKQA DEKGEIDKKNNLLADVIEASGILSDGTESFPPSINAKSALSNQE EKISEELEEPTETVVIEKETEVLNVEETWTDESRLQEIAPLEN DEKQNEEVLNVDTSLSLEDINLPLSQLEDGQILEIQTQDDAGV LSESTRITKEMLELLQQAGLHIQEIEVKSDLSNDSYQGSVPVKS TNEEEAVEMKNKESCELLKHEDNDSCEKLILCQQDEYESSEV DKTINESTQSKNIQSNF
5	Apoptosis regulator R1-like	XP_023217360.1	%95.24	MAEHSTEEMEVSCIVNDYISYQLHMNGYEWRSPTATPTKV NSSLRSLGNEFKTRYRDSFLEMTDSLDIQPSTARATFLGVANEL FSEGVTWARIVALFVFASEMALICHRMSHTEIVNDVAEWLSS YIKSNLLTWIRDHDGWDGLVTFHEGYDDLRSESPWPVLKKFV CGVAAGVIGALTVGAILSSKS
6	Venom phosphodies-terase 2-like	XP_023218281.1	%70.91	MMERQTLLALVAFMVVAIVVIFIGGYFCGVNVAYAQRGQKH LNTGPSSEKLTEQTGPWYQQPCFPERNICPSDFRNPPLLLIS LDGFRPDYFDKTPTLKRLAHCGASAPYMKPVFPTKTFPNHY SIVTGLYPESHGIIDNNIYDPKLRRLFSVRRGQDNYNPIWWE AEPLWVTAEKQGKITGTYFWPGSEIKINGTRPTYYIKYN NKVSWKTRLDTILSWLELPEESRPSFLTLYINEPDHASHAHG PESPQVNAVLAKLDSFIELMITSLQQRNLLGCINIIIVSDHGMAK SDCNKVIELDNYINSSEVFLSPGPIGKIRPKGKRTVESVVEILR CNRPELLVYKKNDLPKRYHYSYNNKIEPIIIDLPTGWAIENKF DARRCRGGNHGYDYLHPEMNAIFLAFGPSFKKNLLVEPFINIEL YELMAELIDVIPNPNNGSKGSLHYILNNPKTLSASSMHKAITA CVITNLQSSKFKGCNCPNGSLHQISPPSDFKHKQEVILPWGAP MIKDSEFTSICHLMNPDYVTAFHKKLKQPSWISFPLTRKTSIM SGEARFFHGKCWILDQRIPFENQANCTDYSILNRDHSAIQQRS LFPPAFAISENSESSVQLMTNAVPMYKTFHSEIWTQFLILLS KWTQKYDSLNIIMGPAFDERGNGLKMSVGELLERHNNKIPIP THFFVVITRCEASNVPLNSCRSADLDVLSFLIPHKPFPENCQNI NEYFLKHSATVKDIEVLTGLKFFTSLSPYTAIALQTRLTDYEWF NVK
7	Serologically defined colon cancer antigen 3 homolog	XP_023224251.1	%87.23	MENGAKSKDNPFSFKNFLKGSQDIDSNPGRRSVNEDADNPF SFRKFLSQGKSSETKRTYNQNVESSRSQDTASSSPKLSSQTTT SSSFEIYSDFVQELPDFIQDHLELKYEDDKIESNEHLYNSNEH YTHTDYSRETYGNSGSSIESGQAEASLESFNGSSVSTSHNES RIENQIKNESYCGLPDFISGTSKSETVATWSSSHLEESNCIGH SQSSLEDEIYKLREENAFLHEQLEKARRTSDIQASKIVELQQK MIELREAEAQETAALEKMVQQVELNLKVTTERAVLAEGQA AKLKQELKSLQALFKNVMKENDKLTGSSSDMECIAEKANS LASQIKVAADNAEQPLRNLLEGVEHLRLLATLVSSIGKVEEN
8	Colorectal mutant cancer protein-like isoform X3	XP_023227730.1	%98.39	MEENKCIKGRGQRSRCDLHPQRGELHLAALASLKGEILELT NRLQHIGQERDMLERQLNRSQLDRLRLAREHEERLDQQAQR YEERLTELHSIIAELSKKLEIQRTNVIKEEDEFSQQSGEESRQ SEAQSHCTSNDNNATESCSDRDFFLGNDSCVELSRFGETDSH PEVTRTILQDGEEASSGLGDNESSEDRLLQLREVSDLPALTEN RATSCHSLQISQLQEEVLSLRAENTALQEQLGRQEAELNRT RASLASSREERDRLQRRIRELQTKLASLQSQSSQGSQTGSSPTR SKSPHINPTTGERTPTSREEVPIAKMAERVRLKKVESGDRHIL GSEISSLGLSTTKVAEHLVHNVQEESHAQEIFQTLFSSGSTV PESKIREFEVEMERLNSRIEHLKSQNDLLNLTLEESKGQCDRL SVLIGKYESNNTALQLALSYSDQTIEAFEVLLALMETEQGLILS NCRAAGLGALVKGSGEDDQEEVTALLKRSQDNRKSAENVAR HFLQKLDRSFGIACSIPGCNLSPWEDLSSHSHTTSTTSSASSS SDGDFTKVDEQRLRDYIVQLKGDRAAVRTTVVELESVHVDP TVKDPPTFMDAQKLDLENAVLMQELMAMKEEKAELKAQVY LLEKEKTALELQLSSREAQEQAYLVQIEHLKSEVKENQQIAER

→

	Protein	ID Code	Similarity	Amino Acid Sequence
				LRTTQGGKVNNQQPWLERSPSGEMAEELAEAMKREMKMKA RIQELVNTLEKVTKNSEMRAQQSTEFVNDLKRANSALVTAFE KAKKKQQAKLKKVEQQMASMVERHSAQIRMLKQRIAMLE GETGHLMQTSNSDLQV

Table 5. Names, ID codes, similarity ratios and amino acid sequences of 7 different proteins obtained as a result of transcriptome analysis in *L. abdullahbayrami.*

	Protein	ID Code	Similarity	Amino Acid Sequence
1	Breast cancer anti-estrogen resistance protein 3-like	XP_023215813.1	%80.65	MEPMIGTEASPEELKKALEWELSLSNSDLRSHAWYHGTI PRQRAEELMSNDGDFLIRDCISHPGDYVLTCKWKNSSLH FVINKVVFQPFTVYEKVQYQFEEDSFDTVPDLVTYYVGNK RAVSAASGAIINCPVNRTMPLTYYASRYGLQCQLHYAAL AAETDIREADRSPCLARRHSQPCEDGSLSAESHPGTFNK TATSSVAGFATLRPHHQPAGWKGMARIGSDPQLSPTQERDGPP PKPSRLPSQHVYTEPIYKDGNAPSNGHSPNINITDMSTPN LNPPSCYDLNKFATTLLQTENKPLLKEPMIGTEASPEELK KALEWELSLSNSDLRSHAWYHGTIPRQRAEELMSNDG DFLIRDCISHPGDYVLTCKWKNSSLHFVINKVVFQPFTVY EKVQYQFEEDSFDTVPDLVTYYVGNKRAVSAASGAIINCPVN RTMPLTYYASRYGLQCQLHYAALAAETDIREADRSPCLARRH SQPCEDGSLSAESHPGTFNKTATSSVAGFATLRPHHQPAGWK GMARIGSDPQLSPTQERDGPPPKPSRLPNQHVYTEPIYKDG NAPSNGHSPNINITDMSTPNLNPPSCYDLNKFATTLLQTEN KPLESSTLLQIRRLLLENGPRILANHLTRIDLDMLKNSTDFD FGLGVVSGLELITLPQGKQLRMDLMERNHCLRYFVAVTILTC SNEEERANLVHKWIQIAIETKTALGNLYGFTAIMQGLALQ QIDRLRSTWLTVRQNFTENAFTYETKLRPTLKSMQECSN PQAPNTCIPYLLSLITILQRHIEVIENNELNAEPQVEPSGK KTVNILSLGLQWEQSASDYGLQLLLLHLELGRTFVQQCT TYRRNGEIVLDNVKFDNSILDMFKTEFHLKFLWGSKGAN VAAAERYTKFDQVLQVMSERCESS
2	BRCA2-interacting transcriptional repressor EMSY-like isoform X2	XP_023216199.1	%76.09	MWPMLLDFSRDECKRILRKLELEAYAAIVSAFRAQGELTKDK KKLLQELSTVLSISLERHRAEIRRAVNDERLNTIADRIYGPNTS VDWAIEGRRLIPLLPRLVPQTAFTAVANSVASIQAAKNATM PLPSASGVKEGAIPTTSSAPPTPSRTPTPTQTCRVSLPSSIPLK TTSVNGNNSGNLTGGIITRLPQETNEENQSEDFNGKKRKRSA SLDSASLPEKITISSNPLVSDENQAPEKPVTTTNNGMLCSST PVTTQATITRTFATTPIRITLPSSQKQNAAVSSSAIPKQVVLQQL QTSNASTTANVFQRSVSIPVVKTVSATATTITQKQSKQTVS PSVQGTLLLTGKVSTTTTTGQSHPLNVPIATVGTIARTRPK TLPPPRQSIRPRSSVLSINQNRMQFVGQNFTSQLPTPVAKSS LQIGTPIQVKQVSSDGRTVQIRQEGGVKIVAQGLPVAASKILP KPSTSPIVMVSSTQGSNASKFSASSTVIQNQIGTKMVSIASPQN NQSSNKVTAIGNVSLVSRTSQAFSTRTGSAVVQSPGVKPNVIV VHKAQVWPQSQGQGTTIVVSGNPVPKLSTETVQETVTYIQK SEGGRTNSSPAVVSVSSCNPANISALNSTAITNFEQSEKNESST VNQENKTTLNVNSKQADEKGEIDKKNNLLADVIEASGILS DGTESFPPSINAKSALSNQEEKISEELEEPTETVVIEKETEVL NVEETWTDESRLQEIAPLENDEKQNEEVLNVDTSLSLEDIN LPLSQLEDGQILEIQTQDDAGVLSESTRITKEMLELLQQAGL HIQEIEVKSDLSNDSYQGSVPVKSTNEEEAVEMKNKES CELLKHEDNDSCEKLILCQQDEYESSEVDKTINESTQSKNIQS NF
3	Apoptosis regulator R1-like	XP_023217360.1	%96.30	MAEHSTEEMEVSCIVNDYISYQLHMNGYEWRSPTATPT KVNSSLRSLGNEFKTRYRDSFLEMTDSLDIQPSTARATFLG VANELFSEGVTWARIVALFVFASEMALICHRMSHTEIVND VAEWLSSYIKSNLLTWIRDHDGWDGLVTFHEGYDDLRSESP WPVLKKFVCGVAAGVIGALTVGAILSSKS

→

	Protein	ID Code	Similarity	Amino Acid Sequence
4	Venom phosphodiesterase 2	XP_023218281.1	%95.65	MMERQTLLALVAFMVVAIVVIFIGGYFCGVN VAYAQRGQKHLNTGPSSEKLTEQTGPWYQQPCFPERNICPS DFRNPPLLLISLDGFRPDYFDKTPTLKRLAHCGASAPYMK PVFPTKTFPNHYSIVTGLYPESHGIIDNNIYDPKLRRLFSVR RGQDNYNPIWWEAEPLWVTAEKQGKITGTYFWPGSEIKING TRPTYYIKYNNKVSWKTRLDTILSWLELPEESRPSFLTLYINEP DHASHAHGPESPQVNAVLAKLDSFIELMITSLQQRNLLGCI NIIIVSDHGMAKSDCNKVIELDNYINSSEVFLSPGPIGKIRP KGKRTVESVVEILRCNRPELLVYKKNDLPKRYHYSYNNKIE PIIIDLPTGWAIENKFDARRCRGGNHGYDYLHPEMNAIFLAF GPSFKKNLLVEPFINIELYELMAELIDVIPNPNNGSKGSLHY ILNNPKTLSASSMHKAITACVITNLQSSKFKGCNCPNGSLH QISPPSDFKHKQEVILPWGAPMIKDSEFTSICHLMNPDYVTAF HKKLKQPSWISFPLTRKTSIMSGEARFFHGKCWILDQRIPFEN QANCTDYSILNRDHSAIQQRSLFPPAFAISENSESSVQLMT NAVPMYKTFHSEIWTQFLILLSKWTQKYDSLNIIMGPAF DERGNGLKMSVGELLERHNNKIPIPTHFFVVITRCEASNV PLNSCRSADLDVLSFLIPHKPFPENCQNINEYFLKHSATVKDI EVLTGLKFFTSLSPYTAIALQTRLTDYEWFNVK
5	Serologically defined colon cancer antigen 3 homolog	XP_023224251.1	%86.67	MENGAKSKDNPFSFKNFLKGSQDIDSNPGRRSVNEDAD NPFSFRKFLSQGKSSETKRTYNQNVESSRSQDTASSSP KLSSQTTTSSSFEIYSDFVQELPDFIQDHLELKYEDDKIESNE HLYNSNEHYTHTDYSRETYGNSGSSIESGQAEASLESFNGSS VSTSHNESRIENQIKNESYCGLPDFISGTSKSETVATWSSSH LEESNCIGHSQSSLEDEIYKLREENAFLHEQLEKARRTS DIQASKIVELQQKMIELREAEAQETAALEKMVQQVELNLKVT TERAVLAEGQAAKLKQELKSLQALFKNVMKENDKLTGSSS DMECIAEKANSLASQIKVAADNAEQPLRNLLEGVEHLRL LATLVSSIGKVEEN
6	C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8	XP_023212124.1	%66.67	MPVGCGEQNMMYMGPTLYTLRFLKVKGDVTPEREGKI YNFIREGYNRQLTYRKEDGSYGAWIQTPSSTWLTAFTMKVF CQANKFIFVDENVICSGIKWLISHQKDDGAFEDAHPVYHY DMTGGVSGRIPMTAFVLIALQECSSCESEELNVARLKAEL FLEISLPTINDPYIMAVVAYALSQSSSDLKHEANEKLKQM SIFDEDLNQRHWKENTRSHSIEVAAYALLTQLALNDL SYSLPIVNWLNTQRLVGGSFPSTQDTVIALQALSQYSMQSQSP DLSLMCNVTSSNDRNLAKTLEFKNDNALVLQQFEVNKVG GTLFFRTAGNGVGSLAVKLRYNVPVPVEKLCKVSRNIKSIIK TLLKIKII
7	Colorectal mutant cancer protein-like isoform X2	XP_023227729.1	%97.49	MASDGNSESSGSDSGSTPEEERIRRLFQTCDGNGDGYIDSQDL LAVCRQLNLEHCVEEIMEQLGADEQGRISYSEFLRRRMQLM NEINALTMQDQLQDPEHRTNTTDIHVAPPQQTSPAPGSGSTGH WPTGSDTSQGAASGKHESWEFDSGARDLSPEPNTLQKLIEAA GGSISGNSSDLLDLANKLHLAALASLKGEILELTNR LQHIGQERDMLERQLNRSQLDRLRLAREHEERLDQQAQRY EERLTELHSIIAELSKKLEIQRTNVIKEEDEFSQQSGEESRQ SEAQSHCTSNDNNATESCSDRGNDSCVELSRFGETDSHPE VTRTILQDGEEASSGLGDNESSEDRLLQLREVSDLPALTEN RATSCHSLQISQLQEEVLSLRAENTALQEQLGRQEAELNR TRASLASSREERDRLQRRIRELQTKLASLQSQSSQGSQTGSS PTRSKSPHINPTTGERTPTSREEVPIAKMAERVRLKKVES GDRHILGSEISSLGLSTTKVAEHLVHNVQEESHAQEIFQTLF SSGSTVPESKIREFEVEMERLNSRIEHLKSQNDLLNLTLEESK GQCDRLSVLIGKYESNNTALQLALSYSDQTIEAFEVLLAL METEQGLILSNCRAAGLGALVKGSGEDDQEEVTALLKRSQD NRKSAENVARHFLQKLDRSFGIACSIPGCNLSPWEDLSSH SHTTSTTSSASSSSDGDFTKVDEQRLRDYIVQLKGDRAAVRT TVVELESVHVDPTVKDPPTFMDAQKLDLENAVLMQEL MAMKEEKAELKAQVYLLEKEKTALELQLSSREAQEQAY LVQIEHLKSEVKENQQIAERLRTTQGGKVNNQQPWLER SPSGEMAEELAEAMKREMKMKARIQELVNTLEKVTKNSEM RAQQSTEFVNDLKRANSALVTAFEKAKKKQQAKLKKVEQQ MASMVERHSAQIRMLKQRIAMLEGETGHLMQTSNSDLQV

Discussion

Transcriptomic analyses are among the most important molecular studies of recent years. In particular, the sequencing of transcripts encoding venom components from venom glands has made it possible to identify structures with low density or low ionizable potential in proteomic analyses. In this way, it will be possible to understand these small potential new structures and their functions that cannot be detected due to the physicochemical properties of venom components. Sanger sequencing of expressed sequence tags (ESTs) was used to perform the first transcriptome analysis of scorpion venom glands [26].

Currently, NGS technologies, including pyrosequencing, form the basis for transcriptome analysis and expression-based sequencing (Illumi-na) [26]. Previous studies reported 37 transcriptomes of scorpion venom glands belonging to 7 of the 20 existing scorpion families.

Our transcriptome analyses of the venom glands of various scorpion species led to the identification of 35,540 peptides. Of these, 9,254 were unidentified proteins, 124 were venom proteins (venom allergens, venom factor-like proteins, venom metallo-proteinases) and the rest were enzymes, AMP*s*, and cancer-related proteins. The overall statistics of our venom gland transcriptome analyses are as follows: 81.2% for *A. crassicauda*, 82.5% for *H. saulcyi,* and 75.9% for *L. abdullahbayrami*. GC rates were 37% for *A. crassicauda*, 36% for *H. saulcyi,* and 37% for *L. abdullahbayrami*, respectively. The read lengths were 134 bp in *A. crassicauda*, 133 bp in *H. saulcyi,* and 134 bp in *L. abdullahbayrami*, respectively. Finally, the total number of sequences amounted to 41.6 million in *A. crassicauda*, 27.0 in *H. saulcyi,* and 20.7 million in *L. abdullahbayrami*. Quintero-Hernández et al.[8] 83,812,864 raw reads were obtained from the transcriptome of the venom gland of the scorpion species *U. yaschenkoi*. After removing concordances, ambiguous reads, and low-quality reads, the sequence data yielded 83,808,178 assignable reads. The average read length was 101 nt and the transcriptome size corresponded to 8,464,625,978 nucleotides (8.4 Gb) [8]. *U. yaschenkoi* in the scorpion species, only 51% of 62,505 individual genes in the databases searched were found to be significant in the databases searched. Only 3,900 of these sequences had similar sequences to toxins, venom-related components (such as hyaluronidase, phospholipase, and other enzymes), and housekeeping genes (e.g. heat shock protein, β-actin, RNA-binding protein). For the toxin genes, they obtained a total of 3900 sequences (unigenes) that were compatible with the sequences published in GenBank. These sequences were then manually analyzed to select only those coding for toxins, antimicrobial peptides, and venom-specific components. They identified 210 sequences encoding 111 unique amino acid sequences, including venom toxins and proteins involved in venom production [26].

SO et al. [27] using a combination of transcriptomic and proteomic approaches, putative conserved toxins from the venom glands of the scorpions *Liocheles australasiae*, *Mesobuthus martensii,* and *Scorpio maurus palmatus* were identified and compared. The transcriptome sequencing on the three scorpion venom glands produced 40,644,404−45,175,588 clean reads. Subsequent assembly processes produced 43,481, 24,054, and 40,432 transcript sequences for *M. martensii*, *S. maurus palmatus,* and *L. australasiae*, respectively. Among these assembled transcripts, 22,215, 10,045 and 14,145 sequences could be functionally annotated. For the venom gland proteomes, a total number of 1,644 (*M. martensii*), 830 (*S. maurus palmatus*), and 1,269 (*L. australasiae*) [27].

Zhong et al. [28] fully investigated the composition of venom gland peptides in *Hadogenes troglodytes* using a transcriptomic approach. They discovered 121 new peptides from scorpions, including 20 new types of peptides cross-linked with one, two, three, four, or seven disulfide bridges, 11 new K+ channel toxin-like peptides, and 2 new ryanodine receptor-specific toxins, respectively [28]. In the species *H. saulcyi* studied by us, 8 different proteins were obtained as a result of the transcriptome analysis; Venom protease-like (amino acid sequence length 390), venom metalloproteinase antarease-like TtrivMP_A (amino acid sequence length 739), breast cancer antiestrogen resistance protein 3-like (amino acid sequence length 910), BRCA2 interacting transcriptional repressor EMSY-like isoform X2 (amino acid sequence length 888), Apoptosis regulator R1-like (amino acid sequence length 189), venom phosphodiesterase 2-like (amino acid sequence length 770), serologically defined colorectal cancer antigen 3 homolog (amino acid sequence length 377) and colorectal cancer mutant cancer protein-like isoform X3 (amino acid sequence length 812). In the scorpion species *U. yaschenkoi*, [8] the venom protein f-spondin-1-like, the venom protein (spondin-like) venom protein-9, the venom insulin-like venom protein 302-like and the toxin-like venom protein-7 were detected. The proteins detected in our study in the scorpion species *H. saulcyi* were venom metalloproteinase antarease-like TtrivMP_A, venom protease-like, and venom phosphodiesterase-2-like. In this study, 8 different proteins were identified and selected in the species *A. crassicauda* and *H. saulcyi,* and 7 different proteins in the species *L. abdullahbayrami*. The selection was based on the fact that the selected proteins are venom- and cancer-relevant proteins. Based on the data obtained in this study, proteins were identified that could have a potential for cancer treatment. Kayhan et al. [29] statistically analyzed the results of the cytotoxic effects of raw scorpion venom on MCF-7 and A549 cancer cells and L929 fibroblast cells separately using the Mann-Whitely U test. It was found that while the percentage viability value of L929 cells was very close to the A-549 value, the percentage viability value of MCF-7 cells was higher than the L929 value, but the difference in the study was not statistically significant. When the proliferation curves of these cells were examined, it was found that the highest dose of venom slightly decreased the proliferation of MCF-7 and L929 cells, but did not affect the proliferation of A549 cells. They observed the apoptotic and antiproliferative effects of the crude scorpion venom of *M. gibbosus* on MCF-7 and A549 cancer cells to some extent [29]. Valdez-Velázquez et al. [30]

reported comprehensive venom characterization using high-throughput proteomic and Illumina transcriptomic sequencing performed with purified RNA from *Centruroides hirsutipalpus* scorpion venom glands. 2,553,529 reads were assembled into 44,579 transcripts. From these transcripts, 23,880 were successfully annotated using Trinotate. Using specialized databases and by performing bioinformatic searches, it was possible to identify 147 putative venom protein transcripts. The venom of *C. hirsutipalpus* contains the highest reported number (77) of transcripts encoding NaScTxs, which are the components responsible for human fatalities [30].

Setayesh-Mehr et al. [31] focused on the fraction of *Hemiscorpius lepturus* scorpion venom. Two bioactive fractions (F2 and F4) were separated and collected by RP-HPLC. The cytotoxicity of F2 and F4 against human red blood cells and the MCF-7 cell line was then investigated. In the analyses, they determined that Bax and p53 expression increased in MCF-7 cells treated with fractions compared to untreated control cells, while Bcl-2 expression decreased. In vivo analysis results determined that the level of IFN-γ increased in mice treated with F2 and F4 at concentrations of 5 and 10 mg/kg compared to control animals, whereas the level of IL-4 decreased significantly (10). $p < 0.0001$). In conclusion, two extract fractions of *H. lepturus* scorpion venom may exhibit anti-cancer properties and can modulate the immune system and induce apoptosis in cancer cells [31].

In the *A. crassicauda* species examined by us, 8 different proteins were identified as a result of the transcriptome analysis. These proteins are Breast Cancer Anti- Oestrogen Resistance Protein 3-like (amino acid sequence length 910), BRCA2 interacting transcriptional repressor EMSY-like isoform X1 (amino acid sequence length 889), apoptosis regulator R1-like (amino acid sequence length 189), venom phosphodiesterase 2-like (amino acid sequence length 770), serologically defined colorectal cancer antigen 3 homolog (amino acid sequence length 377), C3 and PZP-like alpha-2-macroglobulin domain containing protein8 (amino acid sequence length 367), colorectal mutant cancer protein-like isoform X1 (amino acid sequence length 974) and venom metalloproteinase 3-like isoform X1 (amino acid sequence length 658). The venom neuropeptide was identified as a venom protein in the scorpion species *Parabuthus stridulus*. In our study, we identified a venom protein similar to phosphodiesterase-2 in *A. crassicauda*. Naderisoorki et al. [32] 11 isolated ESTs associated with the sodium channel in the scorpion species *Odonthubuthus doriae* sodium channel. The coding sequences and peptide sequences of these 11 isolated ESTs were determined and then the signal peptides of each EST were predicted using signal peptide prediction tools. Finally, the mature peptides were predicted by deleting the signal peptides. Analysis of these mature ODNaTx peptides showed that the total number of amino acids in 8 of these peptides was similar to the ODNaTx peptides in the Buthidae family [32].

Romero-Gutierrez et al. [33] reported a detailed examination of venom gland transcripts and venom composition of the Mexican scorpion *Thorellius atrox*. By RNA-seq performed with the Illumina protocol, more than 20,000 assembled transcripts were obtained. Following a database search and annotation strategy, 160 transcripts were identified that potentially encode venom components. Mass fingerprinting by LC-MS identified 135 individual venom components; five of these matched the theoretical masses of putative peptides translated from the transcriptome. LC-MS/MS de novo sequencing allowed the reconstruction and identification of 42 proteins encoded by the assembled transcripts [33]. In the species *L. abdullahbayrami* examined by us, 7 different proteins were found as a result of transcriptome analysis; Breast cancer anti-oestrogen resistance protein 3-like (amino acid sequence length 910), BRCA2 interacting transcriptional repressor EMSY-like isoform X2 (amino acid sequence length 888), apoptosis regulator R1-like (amino acid sequence length 189), venom phosphodiesterase 2 (amino acid sequence length 770), serologically defined homolog of colorectal cancer antigen 3 (amino acid sequence length 377), C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8 (amino acid sequence length 367) and colorectal cancer mutant protein-like isoform X2 (amino acid sequence length 970).

Uzair et al. [34] conducted research on the biological activities of scorpion venom in antimicrobial, antiviral, anti-cancer, and immune diseases. This review also describes the evolutionary perspective of peptides derived from different scorpion venoms. The most important poison peptides are Ctriporin, Chlorotoxins (cltx), Neopladine I and II, Meucin 24, Meucin 25, and Hp 1090. The most well-known scorpion species with pharmaceutical activity; *Pandinus imperator*, Chaerilustricostatus, *Buthus martensii*, *Mesobuthus eupeus*, *Leiurus quinnquestriatus*, *Tityus disrepans,* and *Heterometrus bengalensis*. The role of peptides in cardiovascular events and osteoporosis treatment demonstrates their importance. The role of peptides against pathogens, skin infections, pain-relieving effects, and antimalarial and anti-viral effects are discussed in detail [34]. Suhas [35] compiled the latest developments regarding the therapeutic applications of scorpion peptides. In most cases, the mechanism of action, as well as the primary sequence of the peptides, are given, and it therefore seems appropriate to investigate lead molecules for further development. Therefore, the relevant findings presented here can serve as a valuable resource for both improving the pharmacological profile and developing new drugs [35].

The end products of cellular regulatory mechanisms are called metabolites. The content of metabolites is the final response of biological systems to genetic and environmental changes [36]. The goal of metabolomics is to analyze all small molecules in a biological system. Although metabolomics is not yet as mature as genomics and proteomics, it is a more promising technology than others for finding biomarkers for diseases such as cancer [36]. There are few publications investigating the low molecular weight fractions of venoms (frogs, spiders, snakes, scorpions, and wasps) [5]. Hu et al. detected metabolites in the secretions of live scorpions and frogs using ionization mass spectrometry. Lysine, serotonin, tryptophan, and some lipids were identified in the MS spectra [37]. Amino acids

found in the secretions of spiders, scorpions, and frogs serve as precursors of monoamines. Both classes of metabolites have variable biological functions in venoms or intoxications. Lysine and tryptophan in the venom of the scorpion *Tityus serrulatus* immobilize the prey and prevent it from escaping [38].

Conclusion

In this study, the amounts of free amino acids in the venoms of *A.crassicauda*, *H. saulcyi* and *L. abdullahbayrami* were determined by LC-MS/MS methods. According to the LC-MS/MS results, the presence of the amino acids taurine, alpha-aminobutyric acid, phosphoethanolamine, lysine, leucine and arginine was detected in the scorpion venom of *A. crassicauda*. *H. saulcyi* in scorpion venom; the amino acids taurine, lysine, phosphoethanolamine, arginine, alpha-aminopimelic acid and leucine were detected. The presence of the amino acids alpha-aminobutyric acid, taurine, lysine, phenylalanine, arginosuccinic acid, leucine, histamine, arginine and proline was detected in the venom of the scorpion *L.abdullahbayrami*. Our study was the first worldwide to perform transcriptomic analyses of mRNA transcripts in the venom glands of the scorpion species of *A. carsicauda*, *H.saulcyi*, and *L.abdullahbayrami* and metabolic analyses of their venom in their tissues. In the transcriptome study of these scorpion species, amino acid sequences in possible proteins in the venom glands were determined, and the venom contents and amino acids in the venom were compared by examining the profile of free amino acids. In addition, by revealing the amino acid profile of the scorpion venom using molecular biology methods, data were obtained that will contribute to the development of new treatment methods in cancer treatment, and proteins that may have the potential for cancer treatment were identified with the data obtained from this study.

Acknowledgments

This study was supported by HUBAK with project number 22013. We thank HUBAK for its support.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. VON REUMONT B.M., CAMPBELL L.I., JENNER R.A. Quo vadis venomics? A roadmap to neglected venomous invertebrates. Toxins, **6** (12), 3488, **2014.**
2. SUNAGAR K., MORGENSTERN D., REITZEL A.M., MORAN Y. Ecological venomics: How genomics, transcriptomics and proteomics can shed new light on the ecology and evolution of venom. Journal of Proteomics, **135,** 62, **2016.**
3. UTKIN Y.N. Modern trends in animal venom research-omics and nanomaterials. World Journal of Biological Chemistry, **8** (1), 4, **2017.**
4. BUDAK Ş.Ö., DÖNMEZ S. Gıda Biliminde Yeni Omik Teknolojileri. Gıda, **37** (3), 173, **2012.**
5. NICHOLSON J.K., LINDON J.C., HOLMES E. Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica, **29** (11), 1181, **1999.**
6. GONEL A., KOYUNCU I., AKSOY M., CELIK H. Metabolomics in Lipoid Proteinosis. Current Metabolomics and Systems Biology Formerly: Current Metabolomics, **7** (1), 32, **2020.**
7. DUNLOP J.A. Geological history and phylogeny of Chelicerata. Arthropod Structure & Development, **39** (2-3), 124, **2010.**
8. QUINTERO-HERNÁNDEZ V., RAMIREZ-CARRETO S., ROMERO-GUTIÉRREZ M., VALDEZ-VELÁZQUEZ L., BECERRIL B., D.POSSANI L., ORTIZ E. Transcriptome analysis of scorpion species belonging to the Vaejovis genus. PloS One, **10** (2), e0117188, **2015.**
9. ÇAGLAR A. Çocukluk Çağı Zehirlenmelerine Acil Yaklaşım. 1. Baskı; Türkiye Klinikleri: Ankara, pp. 95, **2021** [In Turkish].
10. FILAZI A., ÖZKAN Ö. Türkiye'de akrep serumunun tarihi. Türk Hijyen ve Deneysel Biyoloji /Turkish Bulletin of Hygiene & Experimental Biology, **78** (1), 107, **2021.**
11. TRUGUT T. Turkey scorpions and the paraspecific effect of anti-*Androctonus crassicauda* scorpion serum prepared in Turkey. Turkish Journal of Hygiene and Experimental Biology, **24** (2), 146, **1964.**
12. CALISKAN F., GARCÍA B.I., CORONAS F.I., BATISTA C.V., ZAMUDIO F.Z., POSSANI L.D. Characterization of venom components from the scorpion *Androctonus crassicauda* of Turkey: peptides and genes. Toxicon, **48** (1), 12, **2006**.
13. VACHON M. A propos de quelques scorpions de Turquie collectes par M. Le Professeur Dr.Curt Kosswing. İstanbul Üniversitesi Fen Fakültesi Dergisi Seri No, **16** (4), 341, **1951.**
14. YAĞMUR E.A., KOÇ H., KUNT K.B. Description of a new species of *Leiurus ehrenberg*, 1828 (Scorpiones: Buthidae) from Southeastern Turkey. Euscorpius, **2009** (85), 1, **2009.**
15. BRUNTZ R.C., LINDSLEY C.W., BROWN H.A. Phospholipase D signaling pathways and phosphatidic acid as therapeutic targets in cancer. Pharmacological Reviews, **66** (4), 1033, **2014.**
16. BARADARAN M., JALALI A., JOLODAR A., GHASEMIAN S. New caerin-like antibacterial peptide from the venom gland of the Iranian scorpion *Mesobuthus eupeus*: cDNA amplification and sequence analysis. African Journal of Biotechnology, **11** (44), 10176, **2012.**
17. AL-ASMARI A.K., KUNNATHODI F., AL SAADON K., IDRIS M.M. Elemental analysis of scorpion venoms. Journal of Venom Research, **7**, 16, **2016.**
18. VETTER I., DAVIS J., RASH L., ANANGI R., MOBLI M., ALEWOOD P., LEWIS J.R., KING F.G. Venomics: a new paradigm for natural products-based drug discovery. Amino Acids, **40**, 15, **2011.**
19. AHMADI S., KNERR M.J., ARGEMİ L., BORDON C.F.K., PUCCA B.M., CERNİ A.F., ARANTES C.E., ÇALIŞKAN F., LAUSTSEN H.A. Scorpion venom: Detriments and benefits. Biomedicines, **8** (5), 118, **2020.**
20. MA Y., HE Y., ZHAO R., WU Y., LI W., CAO Z. Extreme diversity of scorpion venom peptides and proteins revealed

by transcriptomic analysis: implication for proteome evolution of scorpion venom arsenal. Journal of Proteomics, **75** (5), 1563, **2012.**
21. BOUGIS P.E., ROCHAT H., SMITH L.A. Precursors of *Androctonus australis* scorpion neurotoxins: structures of precursors, processing outcomes, and expression of a functional recombinant toxin II. Journal of Biological Chemistry, **264** (32), 19259, **1989.**
22. CHEN S., ZHOU Y., CHEN Y., GU J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, **34** (17), i884, **2018.**
23. arXiv preprint arXiv:1303.3997, **2013.** Available online: https://arxiv.org/abs/1303.3997 (accessed on 22 September 2023).
24. DANECEK P., BONFIELD K.J., LIDDLE J., MARSHALL J., OHAN V., POLLARD O.M., WHITWHAM A., KEANE T., MCCARTHY A.S., DAVIES M.R., LI H. Twelve years of SAMtools and BCFtools. Gigascience, **10** (2), giab008, **2021.**
25. R. R Core Team. R: A language and environment for statistical computing, **2013.** Available online: https://www.R-project.org (accessed on 22 September 2023).
26. CID-URIBE J.I., VEYTIA-BUCHELI J.I., ROMERO-GUTIERREZ T., ORTIZ E., POSSANI L.D. Scorpion venomics: a 2019 overview. Expert Review of Proteomics, **17** (1), 67, **2020.**
27. SO W.L., LEUNG T.C., NONG W., BENDENA W.G., NGAI S.M., HUI J.H. Transcriptomic and proteomic analyzes of venom glands of the scorpions *Liocheles australasiae, Mesobuthus martensii* and *Scorpio maurus palmatus.* Peptides, **146,** 170643, **2021.**
28. ZHONG J., ZENG X.C., ZENG X., NIE Y., ZHANG L., WU S., BAO A. Transcriptomic analysis of the venom glands from the scorpion *Hadogenes troglodytes* revealed unique and extremely high diversity of the venom peptides. Journal of Proteomics, **150**, 40, **2017.**
29. KAYHAN N.Y., OCAL İ.Ç., BUYUKKARTAL O. Investigation of the Mesobuthus Gibbosus (Scorpiones: Buthidae) Akrep Ham Zehrinin Kanser Hücre Hatları Üzerine Sitotoksik ve Antiproliferatif Etkilerinin Araştırılması. Abant Tıp Dergisi, **12** (1), 19, **2023.**
30. VALDEZ-VELÁZQUEZ L.L., CID-URIBE J., ROMERO-GUTIERREZ M.T., OLAMENDI-PORTUGAL T., JIMENEZ-VARGAS J.M., POSSANI L.D. Transcriptomic and proteomic analyses of the venom and venom glands of *Centruroides hirsutipalpus*, a dangerous scorpion from Mexico. Toxicon, **179**, 21, **2020.**
31. SETAYESH-MEHR Z., ASOODEH A., GHASEMI L.V. The Anti-cancer Effect of Two Extract Fractions from the *Hemiscorpius lepturus* Scorpion Venom. Applied Biochemistry and Microbiology, **59** (6), 850, **2023.**
32. NADERISOORKI M., GALEHDARI H., BARADARAN M., JALALI A. First venom gland transcriptomic analysis of Iranian yellow scorpion "*Odonthubuthus doriae*" with some new findings. Toxicon, **120**, 69, **2016.**
33. ROMERO-GUTIERREZ T., PEGUERO-SANCHEZ E., CEVALLOS M.A., BATISTA C.V., ORTIZ E., POSSANI L.D. A deeper examination of Thorellius atrox scorpion venom components with omic techonologies. Toxins, **9** (12), 399, **2017.**
34. UZAIR B., BINT-E-IRSHAD S., KHAN B.A., AZAD B., MAHMOOD T., REHMAN M.U., BRAGA V.A. Scorpion venom peptides as a potential source for human drug candidates. Protein and Peptide Letters, **25** (7), 702, **2018.**
35. SUHAS R. Structure, function and mechanistic aspects of scorpion venom peptides-A boon for the development of novel therapeutics. European Journal of Medicinal Chemistry Reports, **6**, 10068, **2022.**
36. KARAHALIL B. Overview of systems biology and omics technologies. Current Medicinal Chemistry, **23** (37), 4221, **2016.**
37. HU B., WANG L., YE WC., YAO ZP. In vivo and real-time monitoring of secondary metabolites of living organisms by mass spectrometry. Scientific Reports, **3** (1), 1, **2013.**
38. HASSANI O., MANSUELLE P., CESTÈLE S., BOURDEAUX M., ROCHAT H., SAMPIERI F. Role of lysine and tryptophan residues in the biological activity of toxin VII 102 (Ts γ) from the scorpion *Tityus serrulatus*. European Journal of Biochemistry, **260** (1), 76, **1999.**