

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE BIOLOGIA GERAL
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA



DISSERTAÇÃO DE MESTRADO

**Estratégias de Anotação do Transcriptoma do
Télson de *Tityus serrulatus***

INCIPIT VITA NOVA

ORIENTADA: Camila Marani Abreu Lopes

ORIENTADOR: Evanguedes Kalapothakis

CO-ORIENTADORA: Izinara Rosse da Cruz

Belo Horizonte

Agosto – 2018

CAMILA MARANI ABREU LOPES DE OLIVEIRA

Estratégias de Anotação do Transcriptoma do Télson de *Tityus serrulatus*

Dissertação submetida ao Programa de Pós-graduação em Genética do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de Mestre em Genética.

ORIENTADOR: Evanguedes Kalapothakis

CO-ORIENTADORA: Izinara Rosse da Cruz

UFMG – ICB – LBMM

Belo Horizonte-Minas Gerais

Agosto – 2018

043 Oliveira, Camila Marani Abreu Lopes de.
Estratégias de Anotação do Transcriptoma do Têlson de *Tityus serrulatus*
[manuscrito] / Camila Marani Abreu Lopes de Oliveira. – 2018.

167 f. : il. ; 29,5 cm.

Orientador: Evanguedes Kalapothakis. Co-orientador: Izinara Rosse da Cruz.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de
Ciências Biológicas.

1. Genética. 2. *Tityus serrulatus* - Teses. 3. Escorpião Veneno - Teses. 4.
Transcriptômica - Teses. 5. Anticorpos - Teses. I. Kalapothakis, Evanguedes.
II. Cruz, Izinara Rosse da. III. Universidade Federal de Minas Gerais.
Instituto de Ciências Biológicas. IV. Título.

CDU: 575

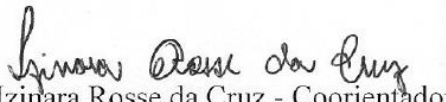


"Estratégias de Anotação do Transcriptoma do Télson de *Tityus serrulatus*"

Camila Marani Abreu Lopes de Oliveira

Dissertação aprovada pela banca examinadora constituída pelos Professores:


Evanguedes Kalapothakis - Orientador
UFMG


Izinara Rosse da Cruz - Coorientador
UFMG


Fabiano Sviatopolk Mirsky Pais
FIOCRUZ


Bárbara Bruna Ribeiro de Oliveira Mendes
UFMG

Belo Horizonte, 07 de agosto de 2018.

“Like the legend of the phoenix. All ends with beginnings. What keeps the planet spinning? The force of love beginning. We’ve come too far to give up who we are. So let’s raise our lives and our souls to the stars.”

Daft Punk – *Get Lucky* (adaptado)

Agradecimentos

Agradeço muito a Deus pela vida, saúde, oportunidade, sustento e ânimo. Durante este período (2016-2018) passei por uma gestação, pelo CTI, por um parto, por uma depressão profunda pós-parto, pela mudança de projeto do mestrado entre outras coisas. Se eu consegui chegar até aqui, é porque não vim sozinha. Por isso, agradeço também ao meu esposo, Yuri, por todo suporte, apoio, incentivo e amor. À minha sogra, Eunice, por ter cuidado com maestria da minha casa e do meu filho, Bento, o que me deixou muito mais tranquila para fazer este trabalho acontecer. [Alguns diriam que não é apropriado contar este tipo de vivência pessoal nos Agradecimentos, mas a Academia é feita de pessoas; pessoas passam por problemas; se estas pessoas tiverem uma rede de apoio, elas vão conseguir realizar sua pesquisa, assim como eu consegui. É importante que histórias como a minha se espalhem, assim como as histórias, que chegaram aos meus ouvidos, de que as alunas grávidas e puérperas não conseguem ser aprovadas nas suas defesas ou qualificações.]

Ao sr. Antônio Filizzola por me ceder um espaço/escritório para que eu pudesse estudar mais próxima de casa. À minha mãe, Cleusa, que, na minha juventude, me levou até a porta da Universidade e me apresentou um mundo mais amplo do que eu podia imaginar; também sou grata a ela por me receber no escritório inúmeras vezes com marmitinhas e cafés-da-tarde. Ao meu pai, Duarte, por todo amor e por se importar.

Ao professor Evanguedes pela oportunidade de cursar o mestrado sob sua supervisão, pela liberdade com que me deixou cumpri-lo e pelo apoio e auxílio na obtenção de bolsa e na confecção deste trabalho. À minha coorientadora Izinara, pelo conhecimento que me transmitiu, pela paciência, apoio e suporte. Sem ela este trabalho não teria saído do campo das ideias e eu teria desistido.

À minha colega e parceira de pesquisa, Thaís Silva, que foi uma amiga incentivadora e auxiliadora; sem o companheirismo dela eu teria perdido oportunidades e aprendido menos. Da mesma forma, agradeço a minha outra colega e, também, parceira, Alessandra Bedore, por ter me ensinado assuntos relativos a este trabalho e por ter me paparicado tanto durante a gravidez.

Aos doutores e colegas Ana Paula Martins e Anderson do Carmo, cujas explanações me pouparam horas, quiçá dias, de angústia atrás de conhecimento. Ao Dr. Tetsu Sakamoto e ao prof. Miguel Ortega que com expertise me ensinaram muito a respeito de bioinformática e de como lidar com a sagarana; eles me proporcionaram todo apoio e suporte com uma solicitude e prontidão desconcertantes. Ao professor Francisco Lobo pelas conversas, dicas e textos lá no princípio, quando o projeto ainda era outro, mas contribuiu muito para meu aprendizado.

Às minhas amigas, Mariana Oliveira e Nazaré Lúcio. A todos os colegas do LBMM (Laboratório de Biotecnologia e Marcadores Moleculares) pelo ambiente agradável, amigável, acolhedor e informativo.

Ao programa de Pós-Graduação e à toda comunidade da UFMG, essa universidade é um organismo vivo, ativo, produtor de conhecimento e debates que enobrecem o ser humano, o cientista, o Brasil! Às agências de fomento à pesquisa CAPES, CNPq e FAPEMIG pelo suporte, e pela bolsa, ainda que intermitente.

Sumário

| | |
|--|------|
| Sumário | viii |
| LISTA DE FIGURAS..... | x |
| LISTA DE TABELAS..... | xii |
| LISTA DE ABREVIATURAS..... | xiii |
| RESUMO | xiv |
| ABSTRACT..... | xv |
| 1. INTRODUÇÃO | 16 |
| 1.1 Escorpiões | 16 |
| 1.1.1 A morfologia dos escorpiões..... | 16 |
| 1.1.2 Características gerais dos escorpiões..... | 18 |
| 1.1.3 Família Buthidae e o gênero <i>Tityus</i> | 20 |
| 1.1.4 A espécie <i>T. serrulatus</i> | 20 |
| 1.2 O escorpionismo no Brasil..... | 23 |
| 1.3 História e produção do soro antiescorpiônico | 26 |
| 1.4 Composição da peçonha | 26 |
| 1.4.1 Toxinas atuantes em canais iônicos | 27 |
| 1.4.2 Proteínas com pontes dissulfeto..... | 28 |
| 1.4.3 Proteínas sem pontes dissulfeto (NDBPs)..... | 29 |
| 1.4.4 Antimicrobianos | 29 |
| 1.4.5 Hipotensinas..... | 30 |
| 1.4.6 Hialuronidases..... | 30 |
| 1.4.7 Proteases..... | 31 |
| 1.4.8 Evolução dos componentes da peçonha..... | 32 |
| 1.5 Transcriptoma como estratégia de identificação de genes | 33 |
| 1.5.1 Estudo transcriptômico por RNA-Seq | 34 |
| 1.5.2 Estudo transcriptômico em glândulas de peçonha escorpiônica..... | 35 |
| 1.5.3 A bioinformática nas análises transcriptômicas..... | 35 |
| 1.6 Anotação do transcriptoma | 36 |
| 1.6.1 Bancos de Dados utilizados na anotação | 37 |
| 2. JUSTIFICATIVA | 38 |
| 3. OBJETIVOS..... | 40 |
| 3.1 Objetivo geral..... | 40 |
| 3.2 Objetivos específicos..... | 40 |
| 4. MATERIAL..... | 40 |
| 4.1 A obtenção do transcriptoma | 40 |

| | | |
|-------|--|-----|
| 5. | MÉTODO..... | 43 |
| 5.1 | Busca por genes ortólogos | 43 |
| 5.2 | Anotação por buscas de similaridade | 43 |
| 5.2.1 | Utilização de blastx contra NR e SwissProt | 43 |
| 5.2.2 | Utilização de tblastn contra ArachnoServer e bancos de dados customizados.. | 44 |
| 5.2.3 | Escolha dos parâmetros de filtragem dos dados após realização de blastx ou tblastn | 48 |
| 5.3 | Utilização do BLAST2GO..... | 48 |
| 5.4 | Anotação por termos ontológicos (GO) | 49 |
| 5.5 | Anotação por vias metabólicas (KAAS-KEGG) | 50 |
| 5.6 | Análises dos <i>contigs no match</i> | 50 |
| 5.7 | Junção dos resultados | 51 |
| 6. | RESULTADOS E DISCUSSÃO | 51 |
| 6.1 | Busca por sequências ortólogas..... | 51 |
| 6.2 | Buscas por similaridade..... | 52 |
| 6.2.1 | Similaridade por blastx e tblastn..... | 52 |
| 6.2.2 | Escolha dos parâmetros para filtrar os alinhamentos locais | 63 |
| 6.3 | Análises do BLAST2GO | 64 |
| 6.4 | Análises do HMMER2GO..... | 68 |
| 6.5 | Análise comparativa dos programas BLAST2GO vs. HMMER2GO | 69 |
| 6.6 | Estratégia de anotação da montagem por vias metabólicas KAAS-KEGG | 75 |
| 6.7 | Resultados gerais | 78 |
| 6.8 | Junção dos resultados das diferentes estratégias de anotação..... | 79 |
| 6.9 | Breve comparação entre os estudos transcriptômicos de <i>Tityus serrulatus</i> | 82 |
| 7. | CONCLUSÃO | 84 |
| 8. | PERSPECTIVAS | 85 |
| | REFERÊNCIAS..... | 87 |
| | ANEXO A – Lista dos organismos utilizados para montagem do banco de dados KAAS-KEGG .. | 98 |
| | APÊNDICE A – Lista da anotação realizada por KEGG | 99 |
| | APÊNDICE B – As 25 vias metabólicas (KEGG) mais abundantes na montagem K24_B145 | 111 |
| | APÊNDICE C – Componentes de peçonha anotados pelo BLAST2GO | 112 |
| | APÊNDICE D – Lista das proteínas 100% similares às ORFs dos <i>contigs</i> da montagem K24_B145 | 117 |

LISTA DE FIGURAS

- Figura 1** – Morfologia externa dorsal e ventral de escorpião (*Tityus aba*). os = prossoma, ms = mesossoma, mt = metassoma, qla = quela, pd = pedipalpos, ql = quelíceras, pe = perna, tl = télson, et = esterno, pn = pente, es = espiráculos. Fonte: Cândido e colaboradores (2005) (adaptado)..... 16
- Figura 2** - Vista frontal e ventral do prossoma de *Tityus serrulatus*. om = olhos medianos, ol = olhos laterais, ql = quelíceras, et = esterno, op = opérculo genital, pn = pente, es = espiráculos. Foto: Tiago Jordão Porto (BRAZIL&PORTO, 2010) 17
- Figura 3** – À esquerda: morfologia externa do télson escorpiônico, vista ventral. tl = télson, fe = ferrão ou agulhão, ac = acúleo, an = ânus, segV = quinto segmento do metassoma, pl = pêlos ou cerdas. À direita: esquema histológico do corte transversal do télson escorpiônico. cut = cutícula, cm = músculo compressor da glândula, gp = glândula de peçonha. Fonte: SISSOM, 1990 (adaptado)..... 17
- Figura 4** – Exemplares de escorpiões fósseis, prova da antiguidade da Ordem Scorpiones. A-) *Proscorpius oshorni*, New York, EUA – espécie marinha do Siluriano Fonte: SISSOM, 1990. B-) *Protoischnurus axelrodorum*, Ceará, Brasil – espécie terrestre do Cretáceo. Foto W. R. Lourenço (BRAZIL&PORTO, 2010)..... 18
- Figura 5** - Gradiente de riqueza de espécies da família Buthidae no Brasil. Quanto mais escura a cor, maior a diversidade de espécies. Todos os estados apresentam registro de escorpiões. Fonte: BRAZIL&PORTO (2010)..... 20
- Figura 6** - *Tityus serrulatus*. Note o dorso marrom e as patas, palpos e metassoma amarelados, o espinho subaculear e as serrilhas do 4º segmento. Fonte do desenho ampliado da cauda BORTOLUZZI et al., 2007. Foto: LOURENÇO, 2015 (figura adaptada) 21
- Figura 7** – Notar a semelhança entre duas espécies distintas de *Tityus*. A - *Tityus serrulatus* e B - *Tityus stigmurus*. Foto: T. J. Porto (BRAZIL&PORTO, 2010) 22
- Figura 8** – Números dos casos de acidentes envolvendo animais peçonhentos no Brasil, registrados para o ano de 2016. Escorpiões são responsáveis por 52,65% dos acidentes. Fonte: (<http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinannet/cnv/animaisbr.def>) 24
- Figura 9** – Gráfico da qualidade geral das reads obtidas por RNA-Seq do télson de *Tityus serrulatus*. Os valores de qualidade Q de Phred são distribuídos de acordo com o tamanho, em número de bases, das reads. A-) Qualidade das reads brutas, antes do tratamento (trimagem) dos dados. Percebe-se como a qualidade cai à medida que a read se estende. B-) Qualidade das reads após trimagem. No gráfico de caixas a linha vermelha indica o valor da mediana, as barras amarelas indicam os quartis 25% e 75%, os traços que limitam inferior e superiormente indicam os 10% e 90% dos dados. A linha azul indica a qualidade média das bases sequenciadas. A faixa verde, laranja e rosa indicam os valores considerados bons, razoáveis e ruins respectivamente.(FastQC 0.10.1 - <http://www.bioinformatics.s.babraham.ac.uk/projects/fastqc/>). 42
- Figura 10** – Metodologias empregadas pelo Blastx e tBlastn de acordo com os parâmetros escolhidos segundo o exemplo da linha de comando anteriormente justaposta à figura. 47
- Figura 11** - Número de genes ortólogos catalogados para artrópodes presentes na montagem K24_B145 52
- Figura 12** - Número de contigs que apresentaram hit com os bancos de dados em maior quantidade. As barras cinzas representam os 26.287 contigs totais da montagem K24_B145 e as barras azuis quantos contigs que cada banco de dados conseguiu anotar..... 55
- Figura 13** - Proporção de sequências dos bancos de dados que apresentaram hit com a montagem K24_B145. As barras cinza representam 100% de cada banco de dados e as barras

| | |
|--|----|
| azuis representam qual a porcentagem destes bancos estão presentes na montagem. Este gráfico também serve para validação da montagem..... | 56 |
| Figura 14 – Número total de proteínas com ao menos 70% de similaridade com os contigs da montagem K24_B145. Quantidade de proteínas, não redundantes, que cada banco acrescenta à anotação final, por alinhamento local..... | 57 |
| Figura 15 - Alinhamento entre os aminoácidos do contig_369 montado neste trabalho, a toxina descrita por Alvarenga et al. (2012) denominada Ts17, a toxina Tbt5 de <i>Tityus bahiensis</i> e a toxina atuante em canal para sódio, Ts5, de <i>Tityus serrulatus</i> . Notar como são similares. Ts17 é 100% idêntica ao contig_369. Tbt5 é 95,35% idêntica e 97,67% similar ao contig_369 e Ts5 é 84,6% idêntica e 86,1% similar ao contig_369..... | 62 |
| Figura 16 - Resultado geral da análise bruta da montagem K24_B145 (26.287 contigs totais) pelo programa BLAST2GO. No eixo x - o número de contigs. No eixo y - as etiquetas que o programa confere a cada contig dependendo do processo de anotação ao qual se enquadram. | 64 |
| Figura 17 - As 29 espécies que mais aparecem entre os resultados de BLAST da montagem K24_B145. No eixo x - quantidade de contigs. No eixo y - nome das espécies. *Espécies também presentes entre as Top-Hit do estudo transcriptômico para <i>Tityus serrulatus</i> de Alvarenga e colaboradores (2012). | 65 |
| Figura 18 - Proporção de termos do segundo nível do GO para cada uma das três ontologias (Componente Celular, Processo Biológico e Função Molecular). Fonte: BLAST2GO..... | 67 |
| Figura 19 - Resumo da anotação funcional realizada pelo BLAST2GO para a montagem K24_B145. No eixo x – o número de contigs. No eixo y – as categorias a que os contigs pertencem..... | 68 |
| Figura 20 - Comparação da anotação funcional relativa a ontologias do GO realizadas pelos programas BLAST2GO e HMMER2GO. No eixo y – a primeira comparação, da esquerda para a direita, é referente ao número de contigs. Nas demais comparações os números são relativos à quantidade de termos do GO distribuídos..... | 70 |
| Figura 21 - Comparação entre os resultados de anotação funcional, dos programas BLAST2GO e HMMER2GO, para os termos GO relacionados a Comp. Celulares. Eixo y – Nº de termos do GO..... | 71 |
| Figura 22 - Comparação entre os resultados de anotação funcional, dos programas BLAST2GO e HMMER2GO, para os termos do GO relacionados a F. Molecular. Eixo y – Nº de termos do GO..... | 72 |
| Figura 23 - Comparação entre os resultados de anotação funcional, dos programas BLAST2GO e HMMER2GO, para os termos GO relacionados a P. Celular. Eixo y – Nº de termos GO (parte 1) | 73 |
| Figura 24 - Comparação entre os resultados de anotação funcional, dos programas BLAST2GO e HMMER2GO, para os termos GO relacionados a P. Celular. Eixo y – Nº de termos GO (parte 2) | 74 |
| Figura 25 – Número de termos do GO anotados em maior quantidade pelo HMMER2GO comparado ao BLAST2GO. | 75 |
| Figura 26 - Número de contigs dentro de cada uma das principais categorias do KEGG..... | 76 |
| Figura 27 - Fluxo da análise dos contigs da montagem K24_B145 para anotação. Os números de contigs anotados por cada estratégia estão em marrom, dentro de seus respectivos quadros. Os 10.410 contigs anotados por similaridade são a soma de 2.292 contigs encontrados por similaridade pelo BLAST2GO mais os 8.118 contigs dos outros alinhamentos locais realizados por blastx e tblastn. Os 12.648 contigs anotados por busca por ontologias são | |

a soma dos 1.840 contigs anotados via HMMER2GO mais os 10.808 contigs anotados via BLAST2GO..... 79

Figura 28 - Visualização da anotação por contigs. Número de contigs que foram anotados por estratégia empregada. 80

LISTA DE TABELAS

Tabela 1 - Número de ocorrências de escorpionismo no Brasil, por ano e por região. O total de acidentes, por ano e o número de óbitos registrados no período. Fonte: tabnet.datasus.gov.br/cgi/tabcgi.exe?sinanet/cnv/animaisbr.de f (2018) 23

Tabela 2 - Valores apresentados por cada um dos bancos de dados personalizados, ou não, após filtragem. | A. Gland – Arachnida Gland; A. Muscle – Arachnida Muscle; Epithelium G. – Epithelium glandular; Outros – Outros Animais. 53

Tabela 3 - Número de proteínas pertencentes a determinadas famílias de proteínas na montagem K24_B145..... 57

Tabela 4 - Componentes da peçonha de *Tityus serrulatus*, anteriormente descritos, agora encontrados neste estudo transcriptômico. *(c200)= Esta sequência já havia sido descrita para *T. serrulatus* na tese de Martins, 2016; mas foi depositada no NCBI em 2017 por outro grupo que também a encontrou em estudos transcriptômicos. Martins, nomeia a sequência de c200. | Alvarenga *et al.*, 2012= no seu estudo transcriptômico de télson de *T. serrulatus* descreve estas sequências. | ^A= Alvarenga *et al.*, 2012 encontra componentes de peçonha anteriormente descritas para outras espécies de escorpiões. | ^B= Componentes da peçonha de outros escorpiões que o SwissProt retorna quando realizamos a busca apenas para a espécie *Tityus serrulatus*. | ^C= Verano-Braga e colab. (2013) em seu estudo de proteoma do *Tityus serrulatus* encontraram proteínas de outros animais entre os componentes da peçonha. | MESEU= *Mesobuthus eupeus* (escorpião); HOTJU= *Hottentotta judaicus* (escorpião), pttv.= putative; TITCL= *Tityus clathratus*; TITBA= *Tityus bahiensis*; TITCO= *Tityus costatus* e DIPPU= *Diploptera punctata* (abelha) 59

Tabela 5 - Quantidade de contigs pertencentes às categorias KEGG de primeiro (cinza) e segundo nível (branco)..... 77

Tabela 6 - Resultados das estratégias de anotação. 81

Tabela 7 - Comparação entre os estudos transcriptômicos já realizados para *Tityus serrulatus*. 84

LISTA DE ABREVIATURAS

| | |
|-----------|---|
| aa | Aminoácidos |
| ACE | Angiotensin Converting Enzyme (Enzima conversora de angiotensina) |
| Ala | Alanina |
| AMP | Antimicrobial Peptide (peptídeo antimicrobiano) |
| BD | Banco de dados |
| BLAST | Basic Local Alignment Search Tool |
| BPP | Bradykinin-Potentiating Peptide (peptídeo potencializador de bradicinina) |
| BUSCO | Benchmarking Universal Single-Copy Orthologs |
| C- | Carboxi |
| CC | Componente Celular |
| cDNA | DNA complementar |
| cm | Centímetros |
| CPU | Central Processing Unit (Unidade central de processamento) |
| Da | Dalton |
| DNA | Ácido dexossirribonucléico |
| EMBL | European Molecular Biology Laboratory |
| EST | Expressed Sequence Tag (etiqueta de sequencia expressa) |
| FM | Função Molecular |
| Gb | Gigabyte |
| GO | Gene Ontology |
| HMM | Hidden Markov Model |
| Ile | Isoleucina |
| KAAS | Kegg Automatic Annotation Server |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| KO | KEGG Orthology |
| LBMM | Laboratório de Biotecnologia e Marcadores Moleculares |
| mRNA | RNA mensageiro |
| N- | Amino |
| NCBI | National Center for Biotechnology Information |
| NDBP | Non Disulfide Bridge Peptide (Peptídeo sem ponte dissulfeto) |
| NGS | Next Generation Sequencing (sequenciamento de segunda geração) |
| NR | Banco de dados Não Redundante do GenBank |
| ORF | Open Reading Frame (Janelas abertas de leitura) |
| pb | Pares de bases |
| PB | Processo Biológico |
| PCR | Reação em cadeia de polimerase |
| PHRED | Phil's Read Editor |
| Pro | Prolina |
| RNA | Ácido ribonucléico |
| RNA-Seq | Técnica de sequenciamento massivo de cDNA |
| SwissProt | Protein knowledgebase of Swiss Institute |
| UniProt | The Universal Protein Resource |
| VAMP | Vesicle-associated Membrane Proteins |

RESUMO

O escorpião amarelo, *Tityus serrulatus*, é uma espécie muito adaptada ao ambiente urbano, de comportamento oportunista e reprodução partenogenética, qualidades que tornam difícil seu controle populacional. Por isso o escorpionismo vem ultrapassando em números absolutos todos os registros de acidentes com animais peçonhentos, inclusive os de cobras e aranhas, no Brasil. Neste país, cerca de uma pessoa foi a óbito a cada três dias, entre 2014 e 2016, metade delas crianças. Sua peçonha é a mais agressiva entre os escorpiões do território brasileiro. A descoberta detalhada dos seus componentes auxiliaria medidas de saúde pública, contribuindo para a produção de um soro antiescorpiônico mais específico e eficaz, que provoca menos reações adversas. Também auxiliaria no desenvolvimento de produtos com potenciais biotecnológicos e farmacêuticos, além de poder proporcionar uma visão evolutiva deste aracnídeo. A transcriptômica é uma potente linha de estudo que auxilia no desvelar dos transcritos de determinados tecidos num determinado momento, sob certas condições fisiológicas. Uma biblioteca de cDNA, proveniente do télson de *Tityus serrulatus*, foi sequenciada por RNA-Seq e uma montagem foi realizada considerando os parâmetros *padrão* do programa. Em seguida esta montagem foi avaliada quanto a sua possível completude gênica e foi verificado que ela apresenta 77,3% dos genes ortólogos descritos para artrópodes. Posteriormente a montagem foi submetida a diferentes estratégias de anotação. Buscas por similaridades, utilizando BLAST2GO e BLAST®; pesquisas por domínios e motivos proteicos com auxílio do INTERPRO; associações aos termos do GO pelo BLAST2GO e HMMER2GO e a verificação de participação em vias metabólicas pelo KAAS-KEGG foram realizadas. Essas técnicas se mostraram complementares, aumentando o número de *contigs* anotados e trazendo mais informação sobre uma parcela deles. No total, 73,35% da montagem escolhida foi anotada, encontrando-se 72% das proteínas já descritas para *T. serrulatus*. Já existem outros dois estudos transcriptômicos do télson de *Tityus serrulatus* na literatura, mas cada um deles extraíram mRNA e sequenciaram cDNA utilizando estratégias diferentes das empregadas no presente trabalho. Essa divergência é suficiente para revelar novas moléculas presentes na peçonha do escorpião, além de apontar novas caracterizações ontológicas e de vias metabólicas do transcriptoma. Várias proteínas putativas, provenientes de outros estudos transcriptômicos, que se supõe serem produzidas pelas glândulas de peçonha de *T. serrulatus*, também foram encontradas. Isso reforça a importância de serem investigadas por métodos experimentais, principalmente as de atividade tóxica.

Palavras-chave: Transcriptoma, Anotação, *Tityus serrulatus*, escorpião, RNA-Seq.

ABSTRACT

The yellow scorpion, *Tityus serrulatus* is well adapted to the urban environment, with opportunistic behavior and parthenogenetic reproduction, qualities that make them hard to perform a population control. That is why scorpionism has surpassed in absolute numbers all the records of accidents with venomous animals in Brazil, including those with snakes and spiders. Between 2014 and 2016, about one person died every three days in Brazil, half of them children. *Tityus serrulatus* venom is the most aggressive among Brazilian scorpions. The detailed discovery of its components would aid public health measures, contributing to the production of a more specific and effective scorpion antivenom, with less adverse reactions. It would also assist in the development of products, with biotechnological and pharmaceutical potential and to provide an evolutionary view of this arachnid. Transcriptomy is a powerful field of study that helps in unveiling transcripts of certain tissues at a given time, under certain physiological conditions. A cDNA library from *Tityus serrulatus* telson was sequenced by RNA-Seq and an assembly was performed considering the program's default parameters. The assembly was evaluated for possible gene completeness, by calculating the presence of the orthologous genes described for arthropods. The result was 77.3%. Subsequently, the assembly was submitted to several annotation strategies. Searches for similarities, using BLAST2GO and BLAST®; searches for domains and protein motifs with support from INTERPRO; associations with GO terms by BLAST2GO and HMMER2GO and verification of involvement in metabolic pathways by the KAAS-KEGG were performed. These techniques were complementary, increasing the number of contigs annotated and bringing more information about a portion of them. We annotated the totality of 73,35% of the chosen assembly and 72% of the proteins already described for *T. serrulatus* were found. There are already two other transcriptomic profiles of *Tityus serrulatus* telson described in the literature; however, different methodologies for mRNA extraction and cDNA sequencing were applied. This divergence is sufficient to reveal new molecules present in the venom of the scorpion, besides pointing out new ontological characterizations and metabolic pathways of the transcriptome. Several putative proteins, from other transcriptomic studies, which are believed to be produced by the venomous gland of *T. serrulatus*, were also found. This reinforces the relevance of experimental methods in the evaluation of those proteins, especially the ones with toxic activity.

Key words: Transcriptome, Annotation, *Tityus serrulatus*, scorpion, RNA-Seq.

1. INTRODUÇÃO

1.1 Escorpiões

1.1.1 A morfologia dos escorpiões

Escorpiões, lacraus, ou surajus como alguns índios brasileiros os denominam, são artrópodes terrestres pertencentes ao Subfilo Chelicerata, Classe Arachnida, Ordem Scorpiones. Seu corpo é dividido em duas partes: prossoma e opistossoma. O prossoma constitui o cefalotórax e o opistossoma se divide em duas partes: mesossoma (tronco) e metassoma (cauda) (Figura 1).

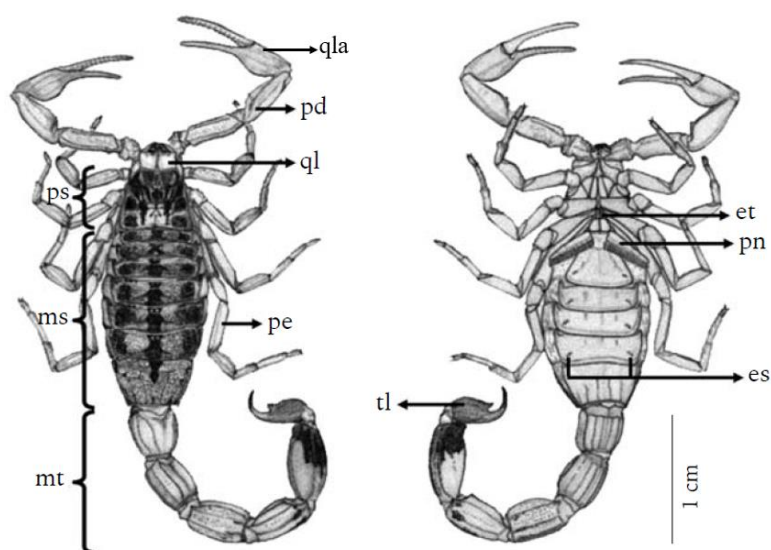


Figura 1 – Morfologia externa dorsal e ventral de escorpião (*Tityus aba*). os = prossoma, ms = mesossoma, mt = metassoma, qla = quela, pd = pedipalpos, ql = quelíceras, pe = perna, tl = télson, et = esterno, pn = pente, es = espiráculos. Fonte: Cândido e colaboradores (2005) (adaptado)

No dorso do prossoma encontramos um par de olhos medianos e dois a cinco pares de olhos laterais, que são muito sensíveis à luz, uma vez que estes artrópodes têm hábitos noturnos (MINEO, *et al.*, 2003). Também encontramos quelíceras trisegmentadas, possuindo dedo móvel e fixo, que trituram as presas (Figura 2A). Pela vista ventral, partem do prossoma quatro pares de pernas locomotoras e um par de pedipalpos, cuja as extremidades apresentam quelas com funções de percepção sensorial, imobilização de presas, defesa e condução do parceiro sexual na corte. Já no opistossoma, na parte mesossomal, em vista ventral, é possível identificar o opérculo genital, pequenos espiráculos que levam a pulmões foliáceos e um par de pentes (ou pectenes) que apresentam funções sensoriais (mecano e quimiorreceptoras). Nos machos, os pentes, auxiliam o processo de corte e nas fêmeas, podem auxiliar no parto (POLIS & SISSOM, 1990; BECCALONI, 2009) (Figura 2B). Além dos pentes, outras estruturas sensoriais se fazem presentes como cerdas cuticulares mecanorreceptoras que recobrem todo o corpo dos

escorpiões. Algumas cerdas das queléceras, pedipalpos e tarsos são muito provavelmente quimiorreceptoras também. Células fotossensíveis estão localizadas na cauda (ROOT, 1990).

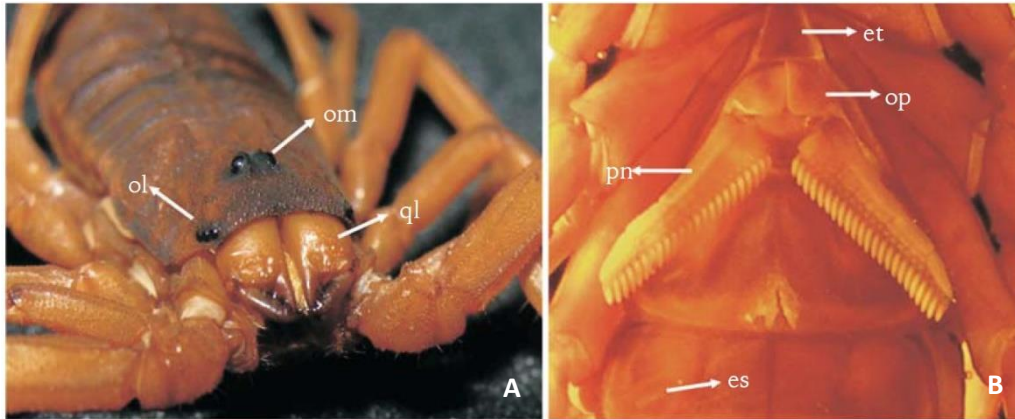


Figura 2 - Vista frontal e ventral do prossoma de *Tityus serrulatus*. om = olhos medianos, ol = olhos laterais, ql = queléceras, et = esterno, op = opérculo genital, pn = pente, es = espiráculos. Foto: Tiago Jordão Porto (BRAZIL&PORTO, 2010)

Enquanto prossoma e mesossoma apresentam achatamento dorsoventral, o mesmo não ocorre com o metassoma, onde há um estreitamento denominado cauda. Na extremidade distal do metassoma há uma estrutura chamada télson que abriga duas glândulas de peçonha e o aguilhão contendo dois orifícios, utilizado na inoculação da peçonha (Figura 3). Ambas glândulas estão circundadas por músculos que ejetam a peçonha por dois ductos que têm seu lúmen na ponta do aguilhão.

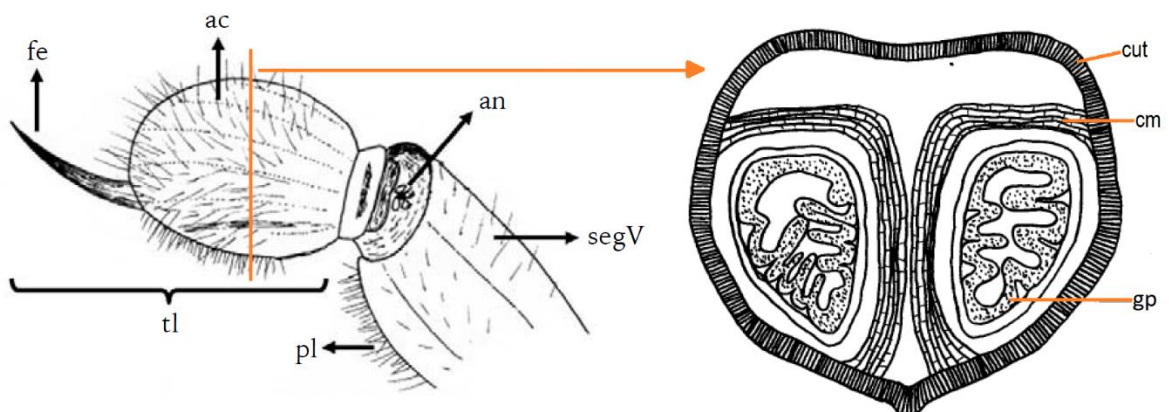


Figura 3 – À esquerda: morfologia externa do télson escorpiônico, vista ventral. tl = télson, fe = ferrão ou aguilhão, ac = acúleo, an = ânus, segV = quinto segmento do metassoma, pl = pêlos ou cerdas. À direita: esquema histológico do corte transversal do télson escorpiônico. cut = cutícula, cm = músculo compressor da glândula, gp = glândula de peçonha. Fonte: SISSOM, 1990 (adaptado)

1.1.2 Características gerais dos escorpiões

Há mais de 4.000 anos os escorpiões despertam o interesse da sociedade. (CLOUDSLEY-THOMPSON, 1990; BECCALONI, 2009). A atenção que eles têm recebido se deve principalmente à sua importância médica, mas também à sua abundância, distribuição geográfica, importância para análise da filogenia de Chelicerata e participação no balanço dos ecossistemas (SISSOM, 1990).

A Ordem Scorpiones surgiu há aproximadamente 450 milhões de anos (Período Siluriano) no ambiente marinho (SISSOM, 1990; JEYAPRAKASH & HOY, 2009). Os registros terrestres datam do final do Devoniano, início do Carbonífero (325 a 350 milhões de anos). São, portanto, os aracnídeos mais antigos que se tem conhecimento (Figura 4) (BECCALONI, 2009).

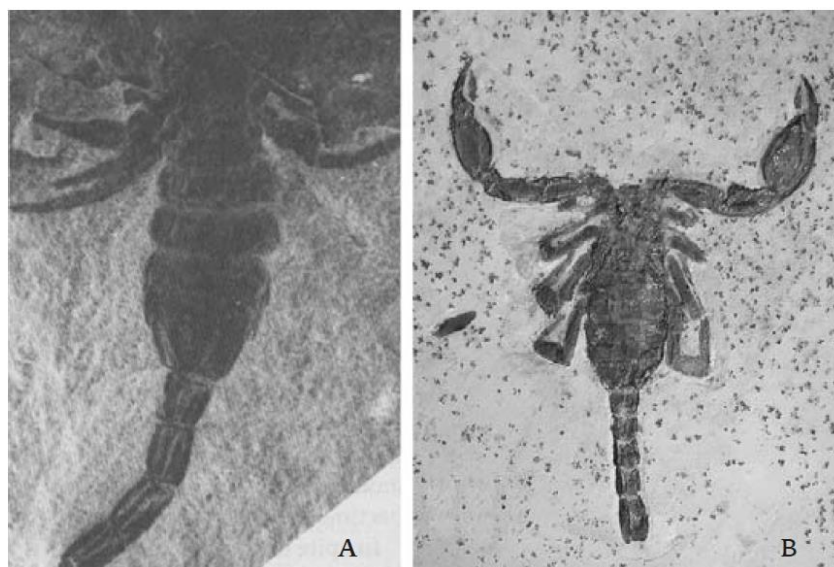


Figura 4 – Exemplos de escorpiões fósseis, prova da antiguidade da Ordem Scorpiones. A-) *Proscorpius oshorni*, New York, EUA – espécie marinha do Siluriano Fonte: SISSOM, 1990. B-) *Protoischnurus axelrodorum*, Ceará, Brasil – espécie terrestre do Cretáceo. Foto W. R. Lourenço (BRAZIL&PORTO, 2010).

Em regiões tropicais eles são mais ativos nos meses quentes e chuvosos do ano. Seus principais predadores são macacos, quatis, camundongos, anuros, lagartos, aves e outros artrópodes. Em ambiente natural, estes artrópodes ficam escondidos sob pedras ou folhagem, em frestas de rochas, cascas de árvores, em troncos caídos no solo, fundo de tocas e em cavernas. Os escorpiões adaptados ao ambiente urbano também são muito encontrados em cemitérios, instalações de rede elétrica, cantos e frestas e sistemas de ar condicionado, enfim, locais onde encontrem abrigo e pouca luz (BRAZIL&PORTO, 2010; BRASIL, 2009). Escorpiões são carnívoros, predadores de baratas, grilos, larvas de insetos e aranhas, sua estratégia é do tipo senta-espera.

Seguram as presas com as quelas e se elas oferecem resistência, os escorpiões inoculam sua peçonha com o objetivo de paralisá-las (BRAZIL&PORTO, 2010).

Sua reprodução sexual é externa. O macho deposita o espermatóforo no solo e com o uso das quelas conduz a fêmea numa “dança” da corte até que esta estrutura se encaixe no opérculo genital feminino, iniciando a transferência indireta dos gametas (MATTHIESEN, 1968). A reprodução assexuada por partenogênese é rara na Ordem, mas já foi reportada para onze espécies no mundo, sendo cinco delas presentes no Brasil (*Tityus serrulatus*, *T. stigmurus*, *T. metuendus*, *T. trivittatus* e *T. uruguayensis*) (LOURENÇO, 2008). São animais vivíparos e o período gestacional varia de dois meses (característica de alguns escorpiões da família Buthidae) a 22 meses. Para o gênero *Tityus*, em geral, é de três meses (BRASIL, 2009).

As ninhadas podem apresentar de um a 105 filhotes, que demoram cerca de 5 a 30 dias para sair do dorso materno, para *T. serrulatus* este período é de 14 dias. O tempo que este cuidado parental perdura é relativo à primeira ou duas primeiras trocas de exoesqueleto dos filhotes (LOURENÇO, 2002b; BRAZIL&PORTO, 2010; BECCALONI, 2009). O desenvolvimento até a chegada da maturidade sexual leva de seis meses a sete anos, dependendo da espécie (BRAZIL&PORTO, 2010). O tempo de vida da maioria dos escorpiões está entre 2 e 10 anos, mas há espécies que chegam a sobreviver por até 25 anos (BRAZIL&PORTO, 2010). *Tityus serrulatus* sobrevive em média por 4 anos (BRASIL, 2009).

Estão presentes em todos os continentes, exceto na Antártida (POLIS, 1990; RUPPERT & BARNES, 2005). Nas américas eles são encontrados desde o sul do Canadá até a Patagônia, na Argentina (RUPPERT & BARNES, 2005). Habitam quase todos os ecossistemas terrestres. Já foram registrados desde 5.500 metros de altitude, nos Andes Peruanos, dentro de cavernas, sob pedras cobertas de gelo; a desertos dos mais áridos e em florestas das mais úmidas. Porém, em sua maioria, os escorpiões não são animais de hábitos generalistas, ao contrário, apresentam padrões biogeográficos e ecológicos bastante previsíveis e localizados (LOURENÇO, EICKSTEDT, 2009 *apud* BRAZIL&PORTO, 2010).

A Ordem Scorpiones possui 13-20 famílias, 209 gêneros e cerca de 2.259 espécies (AMNH, 2018). Mais da metade deles ocorre na região Neotropical (América latina): 8 famílias, 48 gêneros e cerca de 800 espécies (LOURENÇO, 2002a *apud* BRAZIL&PORTO, 2010). O Brasil está entre os países com a maior diversidade de escorpiões do mundo: 131 espécies, de 23 gêneros e quatro famílias habitam todas as regiões e biomas brasileiros (LOURENÇO, 2002c; BRAZIL&PORTO, 2010). As regiões Norte e Nordeste são as responsáveis pela alta riqueza de espécies escorpiônicas (BRAZIL&PORTO, 2010). Ainda assim, a escorpiofauna brasileira é

considerada subestimada por causa das lacunas na amostragem, falta de incentivo à pesquisa e carência de especialistas nesta área (BRAZIL&PORTO, 2010).

1.1.3 Família Buthidae e o gênero *Tityus*

A família Buthidae é a maior e mais amplamente distribuída da Ordem Scorpiones, possuindo representantes em todos os continentes, inclusive em todo o território brasileiro (Figura 5). Os membros dela são reconhecidos por apresentar o esterno subtriangular (Figura 2b) e a patela do pedipalpo sem tricobótrias ventrais (SISSOM, 1990; LOURENÇO, 2002c *apud* BRAZIL&PORTO, 2010). A família destaca-se, também, por conter todos os escorpiões de importância médica do mundo, exceto uma (*Hemiscorpius lepturus*, Hemiscorpiidae). Ela é representada por cerca de 100 gêneros, alguns já extintos, e mais de 1000 espécies; entre elas 220, aproximadamente, são do gênero *Tityus* (C. L. KOCH, 1836) (LOURENÇO, 2015). A característica mais marcante deste gênero é a presença de um espinho sob o ferrão (Figura 6) (BRASIL, 2009).

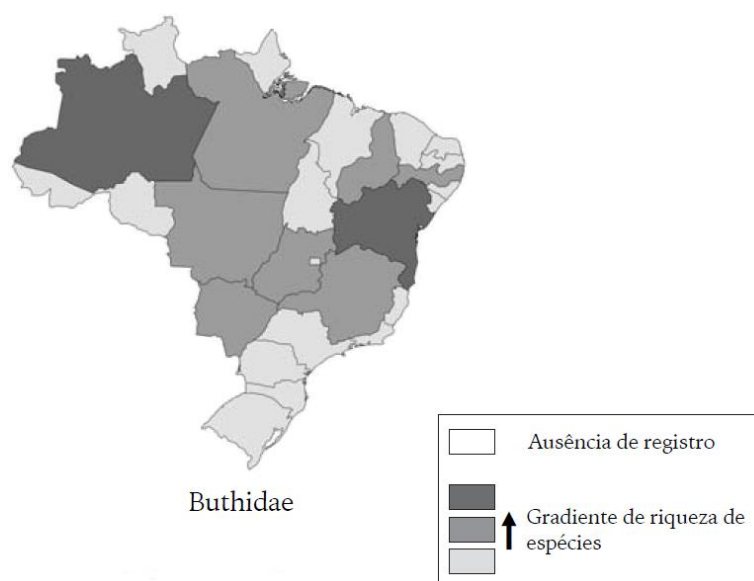


Figura 5 - Gradiente de riqueza de espécies da família Buthidae no Brasil. Quanto mais escura a cor, maior a diversidade de espécies. Todos os estados apresentam registro de escorpiões. Fonte: BRAZIL&PORTO (2010)

1.1.4 A espécie *T. serrulatus*

Tityus serrulatus (LUTZ & MELLO, 1922), ou escorpião amarelo como é popularmente conhecido, é de tamanho médio atingindo entre 5,5 cm até 7,2 cm de comprimento total, possui grânulos, em formato de espinhos, no dorso do 3º e 4º segmentos do metassoma com 4 a 6 dentes, formando uma serrilha que dá origem ao seu nome em latim (LOURENÇO, 2015) (Figura

6). É a principal espécie causadora de acidentes graves no Brasil, com registro de óbitos. Um pouco mais da metade das vítimas são crianças e pré-adolescentes (BRASIL, 2009; GUIA DE VIGILÂNCIA EPIDEMIOLÓGICA, 2009).

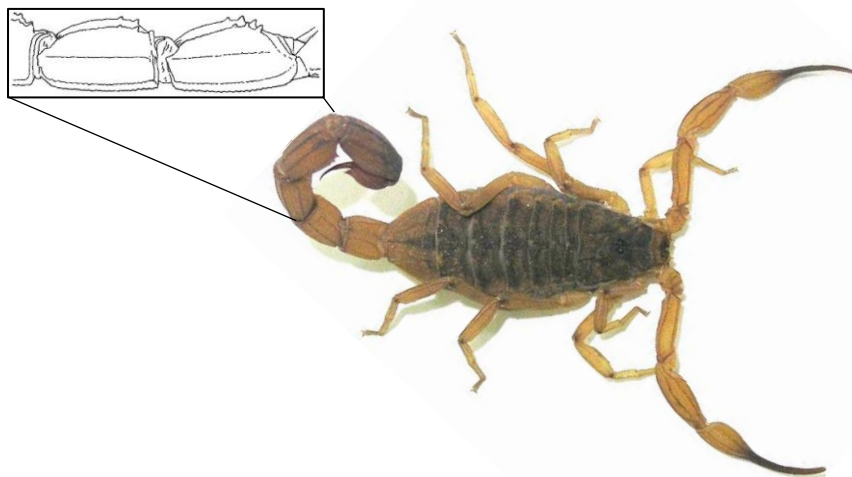


Figura 6 - *Tityus serrulatus*. Note o dorso marrom e as patas, palpos e metassoma amarelados, o espinho subaculear e as serrilhas do 4º segmento. Fonte do desenho ampliado da cauda BORTOLUZZI *et al.*, 2007. Foto: LOURENÇO, 2015 (figura adaptada)

Seus hábitos são de característica oportunista (LOURENÇO, 2015; BRAZIL&PORTO, 2010; BORTOLUZZI *et al.*, 2007) e, além disso, ele é capaz de reproduzir-se por partenogênese (MATTHIESEN, 1962). Uma das características de espécies oportunistas e partenogênicas é sua regulação de população independente da densidade, por isso se a espécie encontra condições adequadas pode proliferar muito e causar surtos (BRAZIL&PORTO, 2010; BRASIL 2009; LOURENÇO, 2015). Esse tipo de estratégia de reprodução é vantajosa porque permite uma rápida colonização e ampla dispersão sem a necessidade do acasalamento. Uma vez transportado passivamente, instala-se e prolifera principalmente em ambientes urbanos (LOURENÇO E CUELLAR, 1995; LOURENÇO, 2002b; SANTOS *et al.*, 2014; BRASIL 2009).

Populações sexuadas e assexuadas (partenogênicas) da mesma espécie ocorrem em diferentes habitats, esse padrão é conhecido por partenogênese geográfica (VANDEL, 1928 *apud* LOUREÇO 2002b; LYNCH, 1984; SANTOS *et al.*, 2014). As populações sexuadas de *T. serrulatus* possuem uma restrita distribuição geográfica, enquanto populações assexuadas são amplamente distribuídas, ocupando preferencialmente áreas urbanas (LOURENÇO, 2008). Ser partenogênético permite ocupar áreas em que populações sexuais têm dificuldades em estabelecer colônias (LYNCH, 1984) além disso pode aumentar a capacidade reprodutiva em até duas vezes porque *T. serrulatus* é partenogênético do tipo *thelytoky* (só produz fêmeas)

(MATTHIESEN, 1962). Cada fêmea tem em média dois partos por ano com cerca de 20 filhotes por vez, podendo gerar 160 filhotes ao longo da vida (BRASIL, 2009). A partenogênese é a principal estratégia reprodutiva de *T. serrulatus* devido a alguns fatores, como o baixíssimo número de machos tombados em três das maiores coleções de escorpiões do Brasil (Instituto Butantan – SP, Museu Nacional – RJ e MZUFBA) e pela experiência de campo dos especialistas da área (SANTOS *et al.*, 2014). O dimorfismo sexual existe, mas é discreto, o metassoma de machos é mais robusto e o pedipalpo deles é maior do que o das fêmeas (SOUZA *et al.*, 2009).

Sua distribuição geográfica, inicialmente restrita apenas ao estado de Minas Gerais, se expandiu e já foi registrado em todo o país, exceto no Rio Grande do Sul, Tocantins, Pará, Amazonas, Acre, Amapá, Paraíba, Roraima, Alagoas e Maranhão. Também já foi descrito na Argentina e Bolívia. Devido a sua extraordinária capacidade adaptativa e seu comportamento oportunista ele se expandiu rapidamente ao longo das décadas pelo Brasil, desalojando espécies antes predominantes, por simples competição (BRASIL, 2009).

Há uma hipótese de que *Tityus serrulatus* (LUTZ & MELLO, 1922) era uma das espécies de menor abundância do país (LOURENÇO, 2015). As explicações vão desde a descrição da espécie ter ocorrido décadas depois de várias outras espécies que habitavam a mesma localidade; até o fato de que o soro antiescorpiônico para *T. bahiensis* (PERTY, 1833) era produzido no Brasil há cinco anos quando houve a descrição de *T. serrulatus* (MAURANO, 1915 *apud* BRAZIL&PORTO, 2010; LOURENÇO, 2015). Registros dos anos 1920 sugerem que *Tityus stigmurus* (THORELL, 1876) era a espécie mais comum de Minas Gerais, até o século 19 (LOURENÇO, 2015). No entanto, é possível que *T. stigmurus* e *T. serrulatus* tenham sido mal identificadas ao longo dos anos devido a sua semelhança (Figura 7). Desde sua descrição, os registros da sua presença ficaram mais frequentes e, em consonância com a urbanização e o desmatamento, a ocorrência de surtos populacionais começaram a ocorrer (LOURENÇO, 2015).



Figura 7 – Notar a semelhança entre duas espécies distintas de *Tityus*. A - *Tityus serrulatus* e B - *Tityus stigmurus*. Foto: T. J. Porto (BRAZIL&PORTO, 2010)

1.2 O escorpionismo no Brasil

Escorpionismo é o tipo de acidente que envolve a picada de escorpião, considerado um problema de saúde pública em locais de alta incidência. No entanto, das cerca de 1500 espécies de escorpiões existentes, apenas 30, aproximadamente, têm importância médica (CHIPPAUX & GOYFFON, 2008). As regiões de acidentes escorpiônicos são bem definidas geograficamente, assim como os principais causadores. Os escorpiões que causam acidentes graves, no Brasil, pertencem ao gênero *Tityus*: *T. bahiensis*, *T. stigmurus*, *T. obscurus* e *T. serrulatus*. Sendo este último o responsável pelos casos fatais ou de maior gravidade (BRAZIL&PORTO, 2010; BRASIL 2009; LOURENÇO, 2015).

Um fato preocupante é o aumento de casos envolvendo *T. serrulatus* em áreas que antes predominavam outras espécies menos danosas (CUPO; AZEVEDO-MARQUES; HERING, 2009 *apud* BRAZIL&PORTO, 2010; BRASIL, 2009). Em 2016 foram registrados 91.422 acidentes com 121 óbitos. Isso significa que, por dia, cerca de 250 pessoas procuraram atendimento médico no Brasil e que a cada três dias uma pessoa morreu em decorrência das complicações causadas pela peçonha (DATASUS, 2018) (Tabela 1). Neste mesmo ano os escorpiões foram responsáveis por 52,65% dos casos de acidentes provocados por animais peçonhentos, ultrapassando em número absolutos as ocorrências de ofidismo e araneísmo juntas (<http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinannet/cnv/animaisbr.def>) (Figura 8).

Tabela 1 - Número de ocorrências de escorpionismo no Brasil, por ano e por região. O total de acidentes, por ano e o número de óbitos registrados no período. Fonte: tabnet.datasus.gov.br/cgi/tabcgi.exe?sinannet/cnv/animaisbr.def (2018)

| Regiões | 2014 | 2015 | 2016 |
|--------------------------------|---------------|---------------|---------------|
| Norte | 3.645 | 3.635 | 3.223 |
| Nordeste | 43.117 | 38.337 | 39.438 |
| Sudeste | 34.947 | 37.927 | 41.954 |
| Sul | 1.857 | 2.508 | 2.238 |
| Centro-Oeste | 4.235 | 3.997 | 4.569 |
| Nº Total de ocorrências | 87.801 | 86.404 | 91.422 |
| Nº Total de óbitos | 100 | 105 | 121 |

Apesar de ele não atacar deliberadamente vertebrados e só picar para se defender, quando se sente ameaçado, dados mostram o aumento na ocorrência do escorpionismo no Brasil. Ano após ano o número de casos só aumenta, inclusive o número de óbitos (DATASUS, 2018). As regiões Sudeste e Nordeste lideram os *rankings* de acidentes (Tabela 1). Regiões bem urbanizadas em detrimento das rurais também apresentam mais ocorrências, sendo os indivíduos da faixa etária entre 20-39 anos os mais afetados em números absolutos (DATASUS, 2018). Explicações para este aumento da incidência de escorpionismo estão diretamente relacionados aos hábitos do agente causal: sua forma de reprodução, hábitos alimentares e comportamento, aliados às circunstâncias geradas pelo homem.

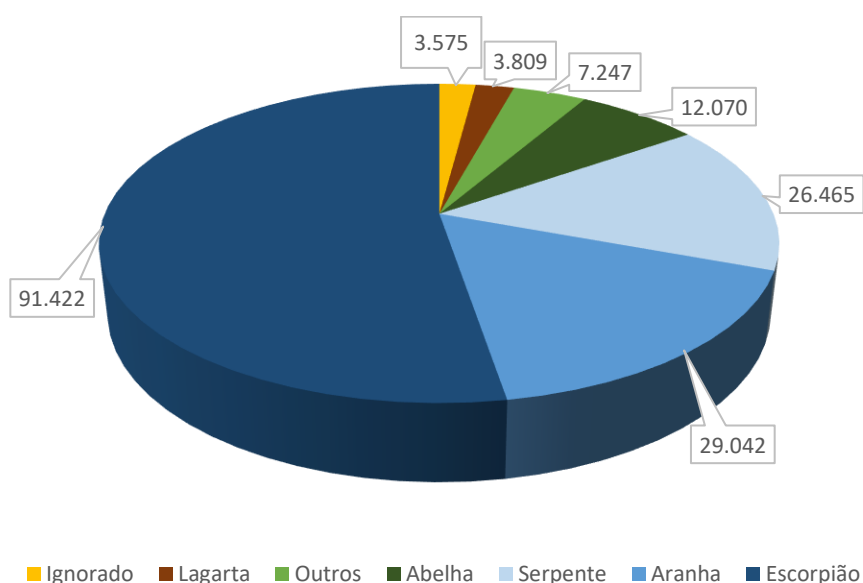


Figura 8 – Números dos casos de acidentes envolvendo animais peçonhentos no Brasil, registrados para o ano de 2016. Escorpiões são responsáveis por 52,65% dos acidentes. Fonte: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinannet/cnv/animaisbr.def>

O controle das populações se faz necessário para a diminuição do número de acidentes escorpiônicos, uma vez que a erradicação dos escorpiões não é possível nem viável, afinal, escorpiões têm importante papel no equilíbrio ecológico. Ações de controle populacional se restringem à busca ativa, captura e eliminação dos possíveis abrigos e alimentos dos escorpiões, porque o controle químico não é eficaz. O hábito dos escorpiões de se esconder e permanecer meses sem se movimentar torna o tratamento químico ineficaz, uma vez que eles não entram em contato com as substâncias e soma-se a isto a capacidade de fechar seus estigmas pulmonares por longos períodos. Outro fator que dificulta o processo de eliminação dos escorpiões é que eles não formam ninhos, mas podem deslocar-se o tempo todo, sem necessariamente retornar ao lugar de origem (BRASIL, 2009).

A gravidade do acidente depende de alguns elementos como idade e sensibilidade da vítima às toxinas; local da picada; tempo decorrido entre a picada e o tratamento; quantidade e toxicidade da peçonha injetada, tamanho e espécie do escorpião (PUCCA *et al.*, 2015; NUNAN *et al.*, 2003). As crianças, por possuírem menor massa corporal, apresentam maior concentração de toxinas (toxinas/peso) além de possuírem maior permeabilidade a pequenos peptídeos pela barreira sanguínea cerebral, facilitando a ação das neurotoxinas. Idosos também são mais vulneráveis à ação da peçonha por apresentarem problemas de saúde prévios como cardiopatias e baixa imunidade, por exemplo (PUCCA *et al.*, 2015; NUNAN *et al.*, 2003). A peçonha de *T. serrulatus* afeta todos os sistemas anatômicos do corpo humano, ela é potente e fatal. Sua letalidade é baixa, porque além da gravidade do envenenamento depender dos vários motivos, acima citados, o volume injetado durante uma picada varia entre 0,1 e 0,9 microlitros, mil vezes menor que de uma serpente (PUCCA *et al.*, 2015).

Os sintomas podem ser locais (dor, edema, hiperemia, piloereção) ou sistêmicos (manifestações gastrointestinais, respiratórias, cardiocirculatórias e neurológicas). Em 97% dos acidentes registrados o prognóstico é leve, pacientes relatam apenas a sintomatologia local, os sinais inflamatórios da picada são pouco evidentes e é incomum a visualização da marca do ferrão (ESPINO-SOLIS *et al.*, 2009; MINISTÉRIO DA SAÚDE, 2001). O tratamento se limita a prescrição de analgésicos. Nos casos moderados e graves os sintomas são sistêmicos (náuseas, sudorese, vômitos, taquicardia ou bradicardia, taquipneia, agitação sozinha ou alternada com sonolência, hipertensão arterial, palidez, tremores e espasmos musculares). Pode haver evolução para insuficiência cardíaca, edema agudo do pulmão e choque cardiocirculatório (ESPINO-SOLIS *et al.*, 2009; MINISTÉRIO DA SAÚDE, 2001). Apesar da peçonha não exibir atividade intensa de proteases, alguns poucos pacientes manifestam pancreatite (ORTIZ *et al.*, 2014). O vômito é um sintoma significativo, pois a frequência indica a gravidade do envenenamento (MINISTÉRIO DA SAÚDE, 2001).

A soroterapia é indicada para os casos moderados e graves, em quantidade ajustada à gravidade do envenenamento. Como casos graves evoluem com rapidez, deve ser iniciada o quanto antes. Junto com a soroterapia outras medicações são utilizadas para amenizar os sintomas (MINISTÉRIO DA SAÚDE, 2001). Há casos em que doses de insulina precisam ser administradas, pois a peçonha pode inibir este hormônio, levando as células à depleção de glicose (MURTHY & HASE, 1994).

1.3 História e produção do soro antiescorpiônico

Desde o início do século 20 o soro antiescorpiônico, era utilizado em humanos (CHIPPAUX & GOYFFON, 1997). Porque, só então, foi possível purificar as imunoglobulinas, separando-as de outras proteínas plasmáticas dos equinos, cujos graves efeitos colaterais levavam a morte. Ainda assim o soro provocava efeitos adversos.

O soro antiescorpiônico produzido no Brasil é eficaz, embora apresente necessidade de pequenos ajustes moleculares, possua elevado custo e traga injúrias aos cavalos envolvidos no processo. O soro antiescorpiônico pode promover reações adversas precoces (nas primeiras 24 horas) ou tardias (5 a 24 dias depois da administração do soro). Entre as precoces estão as reações alérgicas de graus variáveis: coceira, vermelhidão na pele, tosse seca, rouquidão, náuseas, vômitos, crise asmática e choque anafilático. Já as reações tardias se caracterizam por febre, coceira, dor nas articulações, aumento dos gânglios e raramente comprometimento neurológico ou renal, estes dois últimos são também conhecidos por “Doença do Soro”. Os efeitos colaterais estão presentes principalmente em pacientes que já tiveram contato com soros de mesma origem (sangue de cavalo) (FUNED, 2017). Ainda assim, países que dispõem do soro antiescorpiônico apresentam um decréscimo expressivo nas taxas de mortalidade (CHIPPAUX & GOYFFON, 2008).

Os cavalos produzem anticorpos contra a peçonha bruta, gerando imunoglobulinas que combatem inúmeros componentes da peçonha, inclusive os não tóxicos ou não ativos em mamíferos, fazendo com que os anticorpos que realmente combatem os efeitos da peçonha estejam diluídos no soro (CHIPPAUX E GOYFFON, 1997; GUATIMOSIM *et al.*, 2000). Além disso, nem todos os cavalos produzem soro com a capacidade neutralizante desejada. Diante disso, estudos no sentido de melhorar a eficácia do soro e propor alternativas à produção clássica deste medicamento vêm sendo feitas (CARMO *et al.*, 2015) e dependem de pesquisas que aprofundem o conhecimento sobre os componentes da peçonha de *Tityus serrulatus*.

Discriminar a sequência de DNA que codifica cada componente da peçonha, sua sequência de aminoácidos e estrutura terciária são de fundamental importância para atingirmos o objetivo de desenvolver um soro mais específico e eficaz contra as moléstias da peçonha e com menos reações adversas para as vítimas de escorpionismo.

1.4 Composição da peçonha

Todos os escorpiões são produtores e inoculadores de peçonha, possuindo a habilidade de controlar a quantidade de peçonha injetada dependendo das características da

vítima (ex: tamanho e mobilidade) (MALLI *et al.*, 1999; CASEWELL *et al.*, 2013; HORTA *et al.*, 2016). A produção da peçonha é uma atividade dispendiosa energeticamente, por isso nem sempre os escorpiões vão utilizá-la, muitas vezes apenas os pedipalpos e as quelíceras são suficientes para render a presa, evitando a necessidade do envenenamento. Inclusive, é comum o escorpião que apresenta pedipalpos robustos, possuir télson reduzido e peçonha fraca, o contrário também é verdadeiro (MEBS, 2002 *apud* HORTA *et al.*, 2016). Escorpiões também são capazes de produzir uma pré-peçonha, de menor gasto metabólico, repleto de sais de potássio e pobre em peptídeos tóxicos, a qual eles utilizam em presas mais indefesas (INCEOGLU *et al.*, 2003; CASEWELL *et al.*, 2013; NISANI & HAYES, 2011).

A peçonha escorpiônica, de modo geral, possui coloração esbranquiçada e aspecto mucoso; contém lipídeos, sais inorgânicos, muco, nucleotídeos, aminas biogênicas, peptídeos, aminoácidos, enzimas, proteínas de alto peso molecular e proteínas de baixa massa com e sem pontes dissulfeto. Os componentes secretados, pelo par de glândulas de *T. serrulatus*, já descritos são: as neurotoxinas, hipotensinas [entre elas os peptídeos potencializadores de bradicinina (BPPs)], hialuronidases, metaloproteases, serinoproteases, peptídeos antimicrobianos (AMPs e ponerinas), peptídeos aniônicos, peptídeos natriuréticos, fosfolipases tipo e inibidores de bradicininas (COLOGNA *et al.*, 2009; ABDEL-RAHMAN *et al.*, 2014; HORTA *et al.*, 2014; CARMO *et al.*, 2014; PUCCA *et al.*, 2015). Modificações pós-traducionais como glicosilação e fosforilação em componentes da peçonha de escorpiões ocorrem, e sua relevância biológica permanece obscura, no entanto a adição de grupos fosfato e glicanos em moléculas tão pequenas provavelmente modificam sua ação (VERANO-BRAGA&DUTRA *et al.*, 2013). Para toxinas de cobras foi sugerido que glicosilações aumentam a estabilidade e difusão destas moléculas (BIRREL *et al.*, 2007).

1.4.1 Toxinas atuantes em canais iônicos

As proteínas de baixo peso molecular (30 – 70 aminoácidos) com três a quatro pontes dissulfeto são neurotoxinas que atuam em canais iônicos (sódio e potássio) de células excitáveis, ou seja, desencadeiam a despolarização das membranas e liberação de catecolaminas e acetilcolina levando aos sintomas clínicos mais graves do escorpionismo (DE LA VEGA & POSSANI, 2005; COLOGNA *et al.*, 2009; BRAZIL&PORTO, 2010; MEBS, 2002 *apud* COLOGNA *et al.*, 2009). Para a peçonha de *T. serrulatus* ainda não foi descrita ação sobre canais iônicos de cloro. As proteínas mais bem estudadas são as de cadeias relativamente longas contendo entre 60-70 aminoácidos entrelaçados por 4 pontes dissulfeto e majoritariamente atuantes em canais de sódio; e as de cadeia curta, compostas por 30-42 aminoácidos com três a quatro pontes

dissulfeto, atuantes principalmente em canais de potássio. As toxinas de cadeia curta usualmente bloqueiam os canais iônicos e as toxinas de cadeia longa geralmente atuam modificando o controle dos canais iônicos, alterando seu potencial de ação (COLOGNA *et al.*, 2009; PUCCA *et al.*, 2015). A equipe do Laboratório de Biotecnologia e Marcadores Moleculares (LBMM) recentemente identificou uma toxina atuante em canais iônicos para cálcio sob a liderança da dra. Bárbara Bruna Ribeiro de Oliveira Mendes. Se trata da CPP-Ts (*Cell-Penetrating Peptides*) que tem potencial biotecnológico e farmacológico para carrear substâncias para a porção intranuclear da célula. A molécula, completa, atua principalmente em células cardíacas despolarizando suas membranas e, alguns minutos após a administração desta molécula no tecido, ela pode ser encontrada no núcleo de células. Se clivada, perde a capacidade tóxica, mas permanece a capacidade de penetrar no núcleo. Um teste já foi realizado e esta CPP-Ts tem preferência por células tumorais.

T. serrulatus apresenta seis toxinas diferentes para canais de sódio: Ts1, Ts2, Ts3, Ts5, Ts17 e Ts18 (COLOGNA *et al.*, 2009; ALVARENGA *et al.*, 2012). Estas são responsáveis pelos efeitos tóxicos mais proeminentes do escorpionismo (COLOGNA *et al.*, 2009). Ts17 e Ts18 foram descritas com base em uma biblioteca de cDNA da glândula de peçonha (ALVARENGA *et al.*, 2012). Já foram descritas, também, sete toxinas para canais de potássio: Ts6, Ts7, Ts8, Ts9, Ts15, Ts16 e Ts19 (COLOGNA *et al.*, 2009; ALVARENGA *et al.*, 2012). Sendo Ts15, Ts16 e Ts19 descritas por técnicas de transcriptômica (ALVARENGA *et al.*, 2012). Possuir toxinas atuantes em canais de potássio é vantajoso evolutivamente, uma vez que esse tipo de canal é provavelmente o tipo mais antigo de canal iônico, estando presente desde organismos procaríotos até os mamíferos (COETZEE *et al.*, 1999). Eles estão distribuídos em, praticamente, todos os tipos de tecidos e células participando de variados processos biológicos (COLOGNA *et al.*, 2009).

1.4.2 Proteínas com pontes dissulfeto

Além das neurotoxinas atuantes em canais de sódio e potássio, há outras proteínas de baixa massa molecular, que também apresentam pares de cisteínas e estão aptas a formar pontes dissulfeto, presentes na peçonha. Ts4 é um exemplo, considerada uma proteína não tóxica, porque não induz os sintomas característicos do envenenamento por escorpiões, embora ela promova reação alérgica, lacrimação e espasmos nas pernas traseiras de camundongos (MARANGONI *et al.*, 1990). Em outro experimento, Ts4 também induziu a produção de anticorpos policlonais que podem neutralizar os efeitos da peçonha, reagindo principalmente com Ts1, Ts2 e Ts3 (CHAVEZ-OLÓRTEGUI *et al.*, 1997; CHAVEZ-OLÓRTEGUI *et al.*, 2002).

Ts11, Ts12 e Ts13 também são proteínas curtas, de 29 aminoácidos, firmemente dobrados entre si por 4 pontes dissulfeto. Estes três peptídeos possuem alta similaridade com uma pequena porção das toxinas para canais de potássio (porção C-terminal da folha β), apesar de seus resíduos de cisteínas não possuírem posição semelhante com as das conhecidas toxinas atuantes em canais de potássio, provavelmente por isso não apresentam toxicidade para camundongos e suas funções biológicas ainda permanecem obscuras (PIMENTA *et al.*, 2003). Ts13 é uma isoforma de Ts12, com uma única modificação na posição 13 (Ala/Pro), que é responsável por um acréscimo de 26Da para Ts13 e a apresentação de um comportamento cromatográfico diferente.

1.4.3 Proteínas sem pontes dissulfeto (NDBPs)

O interesse despertado pelos peptídeos sem pontes dissulfeto (NDBPs – *Non Disulfide Bridge Peptide*) apareceu mais tardiamente. Suas estruturas primárias são altamente diversificadas, bem como suas atividades biológicas (ALMAAYTAH&ALBALAS, 2014). Para *T. serrulatus* já foram comprovadas ações vasculares, diuréticas, antimicrobianas, anticancerígenas e hemolíticas (FERREIRA *et al.*, 1993; ALMEIDA *et al.*, 2002; VERANO-BRAGA&DUTRA *et al.*, 2013; GUO *et al.*, 2013).

1.4.4 Antimicrobianos

Há nove peptídeos antimicrobianos descritos para *T. serrulatus*, todos inicialmente descritos por técnicas de transcriptômica (JK732086, JK732667, JK732409, JK732889, JK732079, JAW07043, A0A218QX07, S6D3A7 e S6CWV8), nem todos são NDBPs. Guo e colaboradores (2013), após obterem duas sequências peptídicas (TsAP-1 e TsAP-2) por meio de uma biblioteca de cDNA, conseguiram caracterizá-las cromatograficamente, sintetizá-las e testá-las com eficácia comprovada. Ambas NDBPs apresentaram atividades antimicrobiana, hemolítica e antiproliferativa em linhagens de células cancerosas humanas. Porém, em todos os ensaios TsAP-2 se mostrou mais potente que TsAP-1.

Oliveira e colaboradores (2017) descreveram uma sequência proteica de 88aa e atribuíram a ela atividade antimicrobiana (A0A218QX07; JAW07022). Porém esta sequência abriga integralmente, com 100% de identidade, um peptídeo menor, de 38aa, potencializador de bradicinina PAPE (P86821) (RATES *et al.*, 2008). Note que uma mesma proteína, se clivada, pode assumir outra função molecular.

1.4.5 Hipotensinas

Hipotensinas são moléculas atuantes em vias metabólicas que controlam a pressão arterial, diminuindo-a. Existem vários tipos de hipotensinas, cada uma delas atua numa via metabólica diferente. Ts10 (Peptídeo T) e Ts14 (hipotensinas I, II, III e IV) são peptídeos potencializadores de bradicinina (BPP), porém Ts10 não apresenta a assinatura de aminoácidos típica de BPPs (um ácido piroglutâmico na porção N-terminal e uma sequência de Ile-Pro-Pro na porção C-terminal) e atua inibindo ACE (Enzima Conversora de Angiotensina), enquanto que as quatro isoformas de Ts14 possuem relativa semelhança com os motivos presentes em BPPs, mas atuam por uma via bioquímica diversa, não inibindo ACE, como a maioria dos BPPs (FERREIRA *et al.*, 1993; VERANO-BRAGA *et al.*, 2010). A TsHpt-I (isoforma de Ts14), por exemplo, possui uma lisina – Lis (hidrofílica e carregada positivamente) no lugar da isoleucina – Ile (hidrofóbica e sem carga) antes do duplete de Prolina (Pro-Pro) e não contém ácido piroglutâmico na porção N-terminal. De forma que ela potencializa o efeito hipotensor da bradicinina, sem inibir ACE, induzindo, também, as células endoteliais a liberar óxido nítrico (NO), um potente vasodilatador (VERANO-BRAGA *et al.*, 2010).

Almaaytah e Albalas (2014), em sua revisão sobre NDBPs não citam Ts14, mas apenas Ts10. Porém, baseada nas características de tamanho de peptídeo, ação bioquímica e fisiológica da molécula e a ausência de pontes dissulfeto, considero Ts14 como pertencente ao grupo dos NDBPs.

1.4.6 Hialuronidases

Hialuronidases são enzimas que catalisam a quebra do ácido hialurônico, um polissacarídeo encontrado principalmente na matriz extracelular de praticamente todos os tecidos de vertebrados (KREIL, 1995). Não são tóxicas por si só, mas são consideradas facilitadoras da dispersão da peçonha nos tecidos da vítima, aumentando a absorção e difusão das toxinas, contribuindo para os sintomas locais e sistêmicos do envenenamento (PESSINI *et al.*, 2001). Horta e colaboradores (2014) descreveram duas isoformas de hialuronidases para *T. serrulatus*, TsHyal-1 e TsHyal-2, com base numa biblioteca de cDNA. A neutralização destas duas hialuronidases reduziu o efeito tóxico da peçonha, reafirmando a importância dos componentes não neurotóxicos no processo de envenenamento. Venâncio e colaboradores (2013) mostram como a atividade das hialuronidases presentes na peçonha de *T. serrulatus* se assemelham com aquelas encontradas em serpentes do gênero *Bothrops*, indicando a potência das hialuronidases do escorpião.

1.4.7 Proteases

Proteases quebram proteínas em suas extremidades (C ou N-terminais) de forma aleatória ou em sítios específicos. Elas são classificadas em 4 grupos: por possuir determinados aminoácidos chave como a serina, cisteína ou ácido aspártico; ou por necessitar de um íon bivalente para realizar suas funções (metaloproteases) (CARMO *et al.*, 2014). Seu papel pode ser o de aumentar a permeabilidade da peçonha nos tecidos ou estar envolvida em processos pós traducionais dos componentes da peçonha, como na remoção de peptídeos sinais ou pro-peptídeos antes da inoculação (COLOGNA *et al.*, 2009; PIMENTA *et al.*, 2003). A quebra de cadeias polipeptídicas parece ser uma modificação pós-traducional comum da peçonha, uma vez que 80% das moléculas identificadas por Verano-Braga&Dutra e colaboradores (2013), num estudo proteômico, eram de fato produtos de proteólise. Eles especulam que a proteólise de toxinas aumenta o repertório molecular da peçonha, apesar de a relevância biológica de algumas modificações pós-traducionais não estarem claras.

O escorpião amarelo possui cerca de 47 proteases descritas e algumas isoformas (Antareases A e B, Metaloserrulases 1-20, Metaloprotease-VAMP (P85842), MQ5 e MQ7 (BERTAZZI, 2007), 16 cisteinoproteases (FUZZITA *et al.*, 2015) e 7 serinoproteases (CARMO *et al.*, 2017). A Metaloserrulase 1 (394aa) (CARMO *et al.*, 2014) abriga integralmente as sequências peptídicas da Antarease, TserMP_A (235aa) (FLETCHER *et al.*, 2010) e sua isoforma TserMP_B (235aa) (ORTIZ *et al.*, 2014) diferindo em apenas um aminoácido. Carmo e colaboradores (2014) optaram por nomear a sequência maior como Metaloserrulase 1 porque à época a sequência depositada por Fletcher e colaboradores (2010) era um pouco distinta da que está depositada na atualidade (P86392).

Apesar de todo este repertório enzimático a ação de proteases na peçonha dos escorpiões é discreta, ao contrário de aranhas e cobras (GWEE *et al.*, 1996; CARMO *et al.*, 2014). Almeida e colaboradores (2002), conseguiram detectar a atividade de proteases serinas na peçonha de *T. serrulatus*. Ortiz e colaboradores (2014) também mostraram que a peçonha deste escorpião pode levar algumas de suas vítimas à pancreatite devido a presença da metaloprotease antarease. A divergência nos resultados pode derivar da diferença nas metodologias empregadas, do uso da peçonha bruta ou purificada ou até de variações intraespecíficas na composição da peçonha. Sabe-se que diferenças na dieta, sexo, origem geográfica e sazonalidade podem influenciar diretamente na composição da peçonha (BRAZIL&PORTO, 2010; DE LA VEGA *et al.*, 2010; DE SOUSA *et al.*, 2010; MA *et al.*, 2012)

Gebel e colaboradores (1979) não haviam encontrado um efeito significativo da peçonha de *Tityus serrulatus* sobre o sistema complemento (SC), mas modificando o método de

análise e aumentando a quantidade de peçonha Bertazzi e colaboradores (2003) conseguiram demonstrar que a peçonha de *T. serrulatus* tem sim um efeito sobre o SC. Bertazzi (2007) mostrou que as metaloproteases MQ-5 e MQ-7 ativam o sistema complemento, reduzindo a atividade lítica das Vias Clássica e Alternativa do soro humano normal. O SC pode amplificar ou desencadear um processo inflamatório, aumentando a permeabilidade capilar, intensificando edemas pulmonar e cerebral.

1.4.8 Evolução dos componentes da peçonha

Há evidências de que a composição da peçonha também sofre pressão seletiva a nível molecular, sua evolução está associada primariamente à predação e posteriormente à defesa. Escorpiões predam diversos invertebrados, portanto evoluíram para a produção de uma maior gama de toxinas de mesma função, que podem atingir os mesmos tipos de receptores num vasto número de invertebrados diferentes (HORTA *et al.*, 2016). Então, como as toxinas escorpiônicas passaram a atingir os canais iônicos em mamíferos? Lourenço (2015) levanta a hipótese de que as neurotoxinas para canais de sódio (Na⁺), específicas para mamíferos, podem ter evoluído durante a aridificação do Paleártico, no período Terciário, quando pequenos mamíferos que vivem em tocas no solo, principalmente roedores, invadiram ambientes áridos competindo por espaço e atuando como predadores noturnos. Tal pressão apresenta uma explicação aceitável para a emergência de toxinas que atuam em mamíferos, que tem como finalidade a defesa e não a caça. Linhagens de Buthidae do Velho Mundo possuem neurotoxinas separadas, especializadas para insetos e para mamíferos, mas os butídeos do Novo Mundo, como os do gênero *Tityus* possuem potentes neurotoxinas que agem em ambos, insetos e mamíferos (LOURENÇO, 2015).

O arsenal molecular presente nas toxinas de animais peçonhentos co-evoluiu por milhares de anos aumentando a especificidade, afinidade e potência dessas moléculas pelos receptores de membrana de suas presas ou agressores, tornando essas substâncias potenciais modelos para fabricação de fármacos (CHUGUNOV A. O. *et al.*, 2013; HORTA *et al.*, 2016). O estudo das toxinas produzidas pelos escorpiões iniciaram, de fato, com a perspectiva de encontrar um antídoto eficaz para os casos de picadas, mas logo se percebeu outras utilidades, como no controle de pragas agrícolas (STEWART *et al.*, 1991; GUREVITZ *et al.*, 2007), contribuição para o entendimento fisiológico e morfológico dos canais iônicos na célula (COLOGNA *et al.*, 2012; DE LA VEGA e POSSANI, 2005) e fins terapêuticos como drogas anticâncer (SALEM *et al.*, 2016), anti-hipertensivos e antitrombogênicos (FERREIRA *et al.*, 1993; HORTA *et al.*, 2016), natriuréticos (ALVES *et al.*, 2013), atuantes no sistema cardiovascular

(ALTAFAJ, 2007; HORTA *et al.*, 2016) e antibióticos (DUEÑAS-CUELLAR *et al.*, 2015). Cientistas de várias áreas como: farmacologia, biologia molecular, terapêutica, bioquímica, química e biofísica têm interesse na ação e elucidação dos componentes da peçonha (GWEE *et al.*, 1996).

Uma revisão bibliográfica e uma pesquisa nos Bancos de Dados do NCBI e UniProt revelam a existência de 242 sequências de proteínas descritas para *Tityus serrulatus*, principalmente componentes da peçonha. Análises proteômicas da peçonha do escorpião amarelo estimam que mais de 300 substâncias diferentes façam parte da peçonha de *T. serrulatus* (NASCIMENTO *et al.*, 2006; VERANO-BRAGA&DUTRA *et al.*, 2013; PIMENTA *et al.*, 2001). Estudos transcriptômicos do tétson também mostram como ainda há muitas moléculas a serem descritas (ALVARENGA *et al.*, 2012; MARTINS, 2016).

1.5 Transcriptoma como estratégia de identificação de genes

O RNA é uma molécula envolvida nos mais diversos processos biológicos, como codificação dos conteúdos genéticos, decodificação, regulação e expressão de genes (YANG&KIM, 2015). Esta molécula polimérica tem meia-vida curta, são degradadas poucas horas após sua síntese em eucariotos. Em procariotos o período é ainda mais curto, poucos minutos (BINNECK, 2004). A brevidade dos transcritos mostra como a composição do transcriptoma muda ao longo do tempo. Por isso, uma análise transcriptômica abrange todos os transcritos de determinadas células, em determinado momento. Uma célula seria capaz de apresentar vários resultados transcriptômicos diferentes, dependendo do seu estágio de desenvolvimento, ou da sua condição fisiológica, por exemplo (YANG&KIM, 2015; PROSDOCIMI *et al.*, 2003).

Em eucariotos, estudos transcriptômicos podem ser a melhor opção quando o interesse é identificar os genes e seus produtos. Porque essa técnica é capaz de sequenciar preferencialmente regiões gênicas, evitando regiões intrônicas que dificultam o entendimento do organismo em estudo. A análise transcriptômica inicia na extração de todos os tipos de RNAs presentes na célula, sucedida por uma purificação de RNA mensageiro (mRNA), quando este é o foco. A partir das sequências nucleotídicas do mRNA o DNA complementar (cDNA) é sintetizado, pela transcriptase reversa, e posteriormente sequenciado. É considerada uma técnica “ômica” porque produz resultados globais dos sistemas biológicos (ESPÍNDOLA *et al.*, 2010). Outros estudos transcriptômicos podem ter interesse em RNAs não codificantes, por exemplo (MARTENS-UZUNOVA *et al.*, 2012), o que muda é a estratégia de purificação do tipo de RNA que se deseja investigar.

A transcriptômica pode auxiliar na resposta a diferentes questionamentos como, por exemplo: identificar os RNAs expressos de um tecido; comparar perfis de expressão gênica em

condições ambientais ou fisiológicas diferentes; detectar a fusão de genes; catalogar os RNAs não codificantes; permitir a observação de variantes de *splicing* ou auxiliar na caracterização de polimorfismos (WANG *et al.*, 2008; ALVARENGA *et al.*, 2012; MAHER *et al.*, 2009; ROYCE *et al.*, 2007; PROSDOCIMI, 2007; BARBAZUK *et al.*, 2007). Existem diversas metodologias empregadas no estudo transcriptômico, cada uma delas atende a um destes questionamentos. Entre elas destaco a produção de *Expressed Sequence Tags* (EST), *Microarray*, *Serial Analysis of Gene Expression* (SAGE) e RNA-Seq (WANG *et al.*, 2009).

1.5.1 Estudo transcriptômico por RNA-Seq

RNA-Seq tem substituído a maioria das outras metodologias disponíveis porque se utiliza de sequenciamento massivo de cDNA (NGS) e por tanto, fornece resultados em larga escala, com alta sensibilidade. Após a conversão do RNA em cDNA, adaptadores são acoplados em uma ou duas extremidades dessa molécula, caso a estratégia seja *single* ou *paired-end sequence*. Após serem sequenciadas por NGS, as *reads* resultantes precisam passar por um controle de qualidade (trimagem), no qual são retirados os adaptadores e as bases com baixa qualidade. Este passo assegura o alinhamento das *reads* na montagem, pois extremidades com erro atrapalham a formação de *contigs* confiáveis. *Reads* trimadas são mapeadas contra um genoma/transcriptoma de referência ou, na ausência de referências, são montadas *de novo*. No presente trabalho, *clusters*, *uniques* e *isotigs* são tratados como sinônimos de *contigs*.

RNA-Seq tem sido a opção de muitos grupos que não trabalham com organismos modelo porque não depende da existência prévia de um genoma da espécie. RNA-Seq possui qualidades como permitir a visualização de onde *reads* começam e terminam com precisão de uma base; poder revelar a existência de *Single Nucleotide Polymorphism* (SNPs); apresentar alto rendimento, sequenciando inclusive transcritos pouco expressos; prover uma quantificação mais precisa dos transcritos e suas isoformas; além dos benefícios de demandar menos amostra de RNA e não possuir etapas de clonagem e amplificação (WANG *et al.*, 2009; BARBAZUK *et al.*, 2007; YANG&KIM, 2015).

Os estudos transcriptômicos são plenos em vantagens, mas qualquer técnica apresenta suas limitações. O transcriptoma não detecta modificações pós-traducionais e nos casos em que é utilizado para anotação de glândulas de peçonha, por exemplo, é difícil dizer se determinadas proteínas são secretadas como componentes da peçonha, ou secretadas para a hemolinfa, ou ainda se são de atuação entre as células da glândula e/ou tecidos adjacentes (OLIVEIRA *et al.*, 2018; XU *et al.*, 2014). Portanto estratégias de sequenciamento de genoma,

transcriptoma e proteoma são complementares e deveriam ser realizadas para o tecido e espécie em questão (PROSDOCIMI, 2007; BINNECK, 2004).

1.5.2 Estudo transcriptômico em glândulas de peçonha escorpiônica

A transcriptômica já contribuiu para um entendimento amplo e geral da complexidade da peçonha de diversos escorpiões (RENDÓN-ANAYA *et al.*, 2012; OLIVEIRA *et al.*, 2018; DENG *et al.*, 2018; ABDEL-RAHMAN *et al.*, 2015). Inclusive o transcriptoma do télson de *Tityus serrulatus* já foi montado e anotado por Alvarenga e colaboradores (2012). Eles obtiveram resultados relevantes como a detecção de novas neurotoxinas putativas, hialuronidases e metaloproteases (ALVARENGA *et al.*, 2012; HORTA *et al.*, 2014; CARMO *et al.*, 2014). Algumas delas foram confirmadas no presente trabalho. No entanto a metodologia utilizada foi a produção de EST, que se comparada a RNA-Seq produz menor quantidade de dados. Ao utilizar EST somente uma porção dos transcritos são analisados e isoformas podem não ser distinguidas limitando os resultados de anotação do transcriptoma (WANG *et al.*, 2009). Não perder isoformas é muito importante no tocante a toxinas escorpiônicas, pois a diferença de um ou dois aminoácidos pode interferir na potência ou na abrangência da atuação de uma toxina (COELHO *et al.*, 2014; informação verbal¹). De forma que realizar uma segunda análise transcriptômica, mas agora por RNA-Seq, do télson de *Tityus serrulatus* pode acrescentar informações importantes no que se refere a descoberta de novos componentes da peçonha e as implicações biotecnológicas que ela acarreta.

1.5.3 A bioinformática nas análises transcriptômicas

Estudos de cunho biotecnológico são direcionados, restritos a pontos muito objetivos de uma determinada área. Eles poupam recursos financeiros e tempo se optarem por análises computacionais de dados obtidos por transcriptômica, por exemplo, selecionando as melhores estratégias experimentais agilizando o processo de descobertas científicas (BINNECK, 2004). Ferramentas bioinformáticas permitem realizar experimentos *in silico*, onde simulações de formas naturais e modificadas são testadas e direcionam estudos clássicos de experimentação *in vitro* e *in vivo* (ESPÍNDOLA *et al.*, 2010).

A bioinformática é uma linha de pesquisa multidisciplinar, iniciou para auxílio na análise de dados genéticos, bioquímicos e moleculares, e expandiu quando sequenciadores em larga escala começaram a produzir milhões de informações genéticas para identificarmos,

¹ Informação obtida em discussão com o professor PhD. Evanguedes Kalapothakis. Contato: kalapothakis@icb.ufmg.br.

entendermos e correlacionarmos (PROSDOCIMI, 2007). Fato que inviabilizou qualquer possibilidade de análises manuais. A bioinformática, por exemplo, é capaz de executar, em tempo hábil, funções de análise da qualidade das centenas de milhões de bases sequenciadas; encontrar determinadas sequências e retirá-las; agrupar *reads* em *contigs* com o objetivo de formar cada um dos genes expressos da célula (YANG&KIM, 2015). Soma-se a isto, as buscas por informações contidas nos inúmeros bancos de dados, que disponibilizam conhecimento sobre função e caracterização de genes, participação em vias metabólicas, predição de estruturas tridimensionais e interações moleculares (ESPÍNDOLA *et al.*, 2010). Enfim, a biologia molecular não estaria tão avançada se não fosse pelo desenvolvimento da informática (PROSDOCIMI, 2007).

Embora os algoritmos sejam habilitados e eficientes para realizar todas estas tarefas, o bioinformata precisa determinar quais parâmetros seguir de acordo com as peculiaridades do organismo em estudo e avaliar os resultados obtidos em cada uma das etapas. As análises computacionais nem sempre geram respostas seguras, principalmente quando tratamos de organismos não-modelo com genomas complexos, como é o caso dos eucariotos. As limitações das ferramentas bioinformáticas são um reflexo do quão complexas são as interações biológicas e também sobre como nosso conhecimento é limitado a respeito do funcionamento das diferentes células a nível molecular.

1.6 Anotação do transcriptoma

Uma etapa comum a projetos genoma e transcriptoma é a de, em posse dos dados de sequenciamento trimados e montados, identificar genes e associá-los a possíveis funções e processos biológicos. A anotação gênica é este processo de relacionar nucleotídeos ao funcionamento do organismo (PROSDOCIMI, 2007).

O processo de anotação pode ser dividido em três etapas (STEIN, 2001). Num primeiro momento os dados são processados em uma série de programas computacionais. Desse modo é possível extrair deles informação suficiente para identificá-los e agrupá-los, preparando-os para as próximas etapas. As orientações para esta fase é que as estratégias de análise sejam voltadas para as peculiaridades do organismo em estudo. Num segundo e terceiro momentos, somente uma equipe multidisciplinar seria competente e possuiria conhecimento suficiente para definir a anotação (STEIN, 2001). Porque, nestas etapas, especialistas selecionam as anotações mais pertinentes e discutem como as informações obtidas nas etapas anteriores se relacionam com a biologia do organismo estudado (STEIN, 2001).

Estudos de anotação de transcriptomas geralmente iniciam suas análises utilizando apenas uma ou algumas das seguintes abordagens: busca por similaridade, busca pela presença de motivos e domínios proteicos, busca pela presença de genes ortólogos ou busca por outras anotações funcionais como participação em vias metabólicas e termos ontológicos (PROSDOCIMI, 2007; CONESA *et al.*, 2016; ESPÍNDOLA *et al.*, 2010; SCHWARTZ *et al.*, 2007; MA *et al.*, 2009; RUIMING *et al.*, 2010; ALVARENGA *et al.*, 2012; DENG *et al.*, 2018; OLIVEIRA *et al.*, 2018). Há ainda outras análises mais restritas que podem ser implementadas como buscas por pontes dissulfeto, por peptídeos sinais, por ORF, por estrutura tridimensional, entre outras (DENG *et al.*, 2018; ALVARENGA *et al.*, 2012; KELLEY *et al.*, 2015).

Anotar transcriptoma realizando buscas por similaridade é a primeira estratégia e a mais utilizada (SCHWARTZ *et al.*, 2007; MA *et al.*, 2009; RUIMING *et al.*, 2010; ALVARENGA *et al.*, 2012; DENG *et al.*, 2018; OLIVEIRA *et al.*, 2018). Buscar por similaridade implica em alinhar sequências global ou localmente. Se as sequências comparadas são alinhadas apenas em regiões mais conservadas, não necessitando estarem alinhadas de uma extremidade a outra, o alinhamento é dito local. Alinhamentos locais retornam geralmente sequências homólogas ou análogas, auxiliando na predição de qual provável função determinada sequência pode apresentar. O programa BLAST (*Basic Local Alignment Search Tool*) é o mais empregado.

Outra estratégia de anotação bastante utilizada é a definição de termos de ontologias (SCHWARTZ *et al.*, 2007; ALVARENGA *et al.*, 2012; DENG *et al.*, 2018; OLIVEIRA *et al.*, 2018). A anotação pelo *Gene Ontology* (GO) caracteriza o gene sob três aspectos: Função Molecular, ou seja, quais atividades moleculares os produtos de genes desempenham; Componente Celular, que aponta aonde, na célula, os produtos de genes estão ativos e Processo Biológico que indica em quais vias metabólicas e processos os produtos de múltiplos genes estão ativos. Os resultados são expressos em códigos (*e.g.* F:GO:0003735; C:GO:0005840; P:GO:0006412) ou forma de árvores (gráficos diretos acíclicos) (PROSDOCIMI, 2007; <http://geneontology.org>).

1.6.1 Bancos de Dados utilizados na anotação

Análises transcriptômicas dependem da completude dos bancos de dados primários e secundários, além dos bancos estruturais e funcionais. Os bancos primários disponibilizam acesso a sequências nucleotídicas ou de aminoácidos sem qualquer processamento, como o NR GenBank NCBI. Os bancos secundários derivam dos primários e agregam, à sequência, informações de função, domínios funcionais, proteínas homólogas entre outras (PROSDOCIMI, 2007). O SwissProt é um exemplo de banco secundário. Por sua vez, os bancos estruturais

contém informações sobre a estrutura química e física das moléculas, como o banco E-MSD (*European Macromolecular Structure Database*), derivado das entradas do Protein Data Bank (PDB) (BOUTSELAKIS *et al.*, 2003). No que concerne aos bancos funcionais, é possível saber qual papel as sequências nucleicas ou proteicas desempenham no metabolismo do organismo em questão. O KEGG é o banco funcional mais utilizado (PROSDOCIMI, 2007).

O *Gene Ontology* (GO) não é um banco de dados, mas reúne e disponibiliza acesso a bancos de dados de ontologias de forma organizada permitindo uma anotação funcional do objeto de estudo. O uso mais comum das anotações via GO é para interpretação de resultados de larga escala, oriundos de experimentos “ômicos” como o transcriptoma (PROSDOCIMI, 2007; <http://geneontology.org>).

O InterPro é um banco de dados integrado de assinaturas proteicas. Os domínios funcionais não são depositados nele, mas o InterPro busca, reúne e disponibiliza de forma sistematizada os dados de domínios e famílias proteicas que foram depositados em bancos diversos como PFAM, SMART, PANTHER, TIGRFAM, SUPERFAMILY entre outros (PROSDOCIMI, 2007; ESPÍNDOLA *et al.*, 2010).

2. JUSTIFICATIVA

Tityus serrulatus é uma espécie de importância médica e epidemiológica (BRASIL, 2009; DATASUS, 2018) e, apesar de sua peçonha ser objeto de estudo há muitos anos, ainda estamos aquém de compreender todo o arsenal molecular tóxico produzido por sua glândula de veneno (DE LA VEGA *et al.*, 2010; VERANO-BRAGA *et al.*, 2013). Nosso grupo de pesquisa, do Laboratório de Biotecnologia e Marcadores Moleculares (LBMM), se dedica há muitos anos ao estudo da peçonha do escorpião *Tityus serrulatus* (MARTINS *et al.*, 2016; FERREIRA *et al.*, 2016; CARMO, 2015; HORTA *et al.*, 2014; CARMO *et al.*, 2014; ALVARENGA *et al.*, 2012; BORGES *et al.*, 2008; KALAPOTHAKIS&CHÁVEZ-OLÓRTEGUI, 1997).

Abordagens bioquímicas e farmacológicas dispendem tempo e recursos que podem ser poupados se análises *in silico* forem primeiro realizadas para posterior confirmação biológica. Pelo fato da bioinformática direcionar estratégias experimentais, ela também acelera a obtenção de resultados e descobertas científicas (BINNECK, 2004).

Alvarenga e colaboradores (2012) já realizaram um estudo transcriptômico para *Tityus serrulatus*, com produção de ESTs. Eles obtiveram resultados relevantes como a detecção de novas neurotoxinas putativas, hialuronidases e metaloproteases (ALVARENGA *et al.*, 2012; HORTA *et al.*, 2014; CARMO *et al.*, 2014). Contudo, o estudo por RNA-Seq abrange mais

transcritos, inclusive os menos expressos e é mais eficiente na distinção de isoformas (WANG *et al.*, 2009), que se fazem presentes entre os componentes da peçonha do escorpião amarelo (HORTA *et al.*, 2014). Por isso, realizar uma segunda análise transcriptômica, mas agora por RNA-Seq, do télson de *Tityus serrulatus*, pode acrescentar informações importantes no que se refere a descoberta de novos componentes da peçonha e as implicações biotecnológicas que este achado acarreta.

Oliveira e colaboradores (2018) também publicaram um estudo do transcriptoma do télson de *Tityus serrulatus* endossado por experimentos proteômicos. Nele, realizam uma análise filogenética das metaloproteases, sugerindo que os genes responsáveis por essas enzimas sofrem expansão intraespecífica, indicando uma pressão evolutiva para a diversificação delas. Eles também relatam encontrar mais de dez famílias de proteínas putativas. No entanto algumas peculiaridades diferem o presente trabalho do de Oliveira e colaboradores (2018) como a quantidade de télsons e o momento de extração do mRNA, além do tipo de sequenciamento. As análises bioinformáticas empregadas para anotação também foram distintas. Acreditamos que essas diferenças podem permitir que o nosso trabalho revele outras características e outras moléculas que não foram relatadas por Oliveira e colaboradores (2018).

A maioria dos estudos de anotação transcriptômica utilizam apenas uma ou duas estratégias de anotação (MA *et al.*, 2009; RUIMING *et al.*, 2010; SCHWARTZ *et al.*, 2007; OLIVEIRA *et al.*, 2018). O presente trabalho se utiliza de seis estratégias diferentes na expectativa que elas sejam complementares e aumentem o número de *contigs* anotados. Uma das estratégias é a busca por similaridade utilizando o tBlastn, uma ferramenta ainda não utilizada na anotação de transcritos.

A produção de um soro antiescorpiônico que provoca menos reações adversas, mais específico e eficaz contra as comorbidades causadas pelas toxinas escorpiônicas, depende de estudos como este, de procura por novos componentes de peçonha. Além disso, propor o uso de várias estratégias de análise do transcriptoma permite que elas sejam utilizadas em diversos outros estudos transcriptômicos, com outros organismos.

3. OBJETIVOS

3.1 Objetivo geral

Anotar o transcriptoma do t lson de *Tityus serrulatus* contido na montagem K24_B145 obtida por RNA-Seq.

3.2 Objetivos espec ficos

- Analisar qu o completa a montagem K24_B145 se apresenta quanto ao seu conte do de genes ort logos.
- Selecionar diferentes estrat gias de anota o para obter o m ximo de informa oes relativas   montagem transcript mica K24_B145.
 - Utilizar ferramenta de alinhamento local para buscar prote nas similares aos *contigs* da montagem K24_B145.
 - Relacionar termos do GO  s ORFs oriundas da tradu o dos *contigs* da montagem K24_B145.
 - Identificar vias metab licas das ORFs oriundas da tradu o dos *contigs* da montagem K24_B145.
- Comparar duas estrat gias de anota o quanto ao retorno de termos do GO distribu dos.
- Quantificar a parcela da montagem K24_B145 que foi anotada e que n o recebeu anota o.
- Verificar se diferentes estrat gias de anota o s o complementares e aumentam as chances de anota o.
- Procurar por pept deos sinais entre os *contigs no match* (n o anotados).
- Relatar quais fam lias, de componentes de pe onha, mais comuns foram encontradas na montagem K24_B145.
- Comparar este trabalho com as anota oes das montagens transcript micas j  realizadas para *Tityus serrulatus*.

4. MATERIAL

4.1 A obten o do transcriptoma

Os dados analisados neste trabalho advieram da montagem do transcriptoma do t lson de *Tityus serrulatus* realizada por Martins (2016). Quarenta indiv duos, coletados em Belo Horizonte (Licen a IBAMA, protocolo 31800-1), foram separados em quatro grupos de 10 indiv duos e submetidos a quatro per odos de extra o da pe onha por estimula o el trica

antes da extração do RNA: 0, 2, 4 e 8 dias. Os télsons dos indivíduos foram então removidos e tratados segundo está descrito em Martins (2016).

Uma biblioteca de cDNA foi construída com o RNA extraído de cada um dos quatro grupos e sequenciada em uma única corrida em sequenciador de segunda geração (NGS) MiSeq (Illumina), identificadas por índices diferentes. A montagem das quatro bibliotecas foi realizada com o Kit TruSeq™ RNA Sample Prep Kit v2 Set A (Illumina) de acordo com as instruções do fabricante. Entre as etapas, destaco: isolamento do mRNA através de esferas magnéticas contendo oligos d(T), fragmentação térmica do mRNA isolado e amplificação dos fragmentos contendo adaptadores. Após estes passos as bibliotecas foram calibradas para apresentarem quantidades iguais de material genético. A biblioteca foi então sequenciada utilizando o MiSeq Reagent Kit v3 600 ciclos e estratégia *paired-end*.

O sequenciamento resultou em 16,7 Gb e 27×10^6 *reads* brutas com tamanho entre 40 e 301 bases. Utilizando o programa CLC Genomics Workbench 9.0 (<https://www.qiagenbioinformatics.com/>) trimou-se as *reads* por qualidade, aplicando o limite de erro de 0,01 (equivalente a Phred 20). Sequências menores do que 70 nucleotídeos e os adaptadores da Illumina também foram removidos pelo mesmo programa (Figura 09). Após a trimagem 24×10^6 *reads* trimadas com tamanho entre 70 e 287 pb prosseguiram nas análises. Outros trabalhos já descreveram novas famílias de toxinas com 3×10^6 *reads* ou menos, 165.646 e 116.027 *reads*, oriundas do sequenciamento do cDNA do télson de escorpiões (RENDÓN-ANAYA *et al.*, 2012; OLIVEIRA *et al.*, 2018; OLIVEIRA *et al.*, 2015).

Não há genoma ou transcriptoma de referência para *Tityus serrulatus*, por isso procedeu-se a montagem *de novo*. O CLC é um *assembler* que utiliza grafos de *Brujin* e emprega dois parâmetros para realizar a montagem: *K-mer* ou *word size* e *bubble size*. Esse programa analisa as *reads* por blocos de tamanhos definidos (*K-mer*), ele se move ao longo das *reads* analisando os blocos e quando encontra diferenças cria um grafo, se no processo de análise houver identidade novamente ele reúne as duas sequências, criando uma bolha (*bubble*). O *bubble size* é definido pelo usuário e representa o tamanho máximo das bolhas, que podem significar SNPs (*Single Nucleotide Polymorphism*), variantes de *splicing* alternativo ou erros de sequenciamento (SCHLIESKY *et al.*, 2012; CLC *user manual*). Os valores de *K-mer* e *bubble size*, da montagem analisada no presente trabalho, foram definidos automaticamente pelo *software* CLC: *K-mer* 24, *bubble size* 145, denominada montagem K24_B145. O número de *contigs* resultantes da montagem foi 26.287, com tamanho médio de 1.637 pb (maior *contig*: 22.487 pb/ menor *contig*: 120 pb) e N50 de 2.037 bases nucleotídicas.

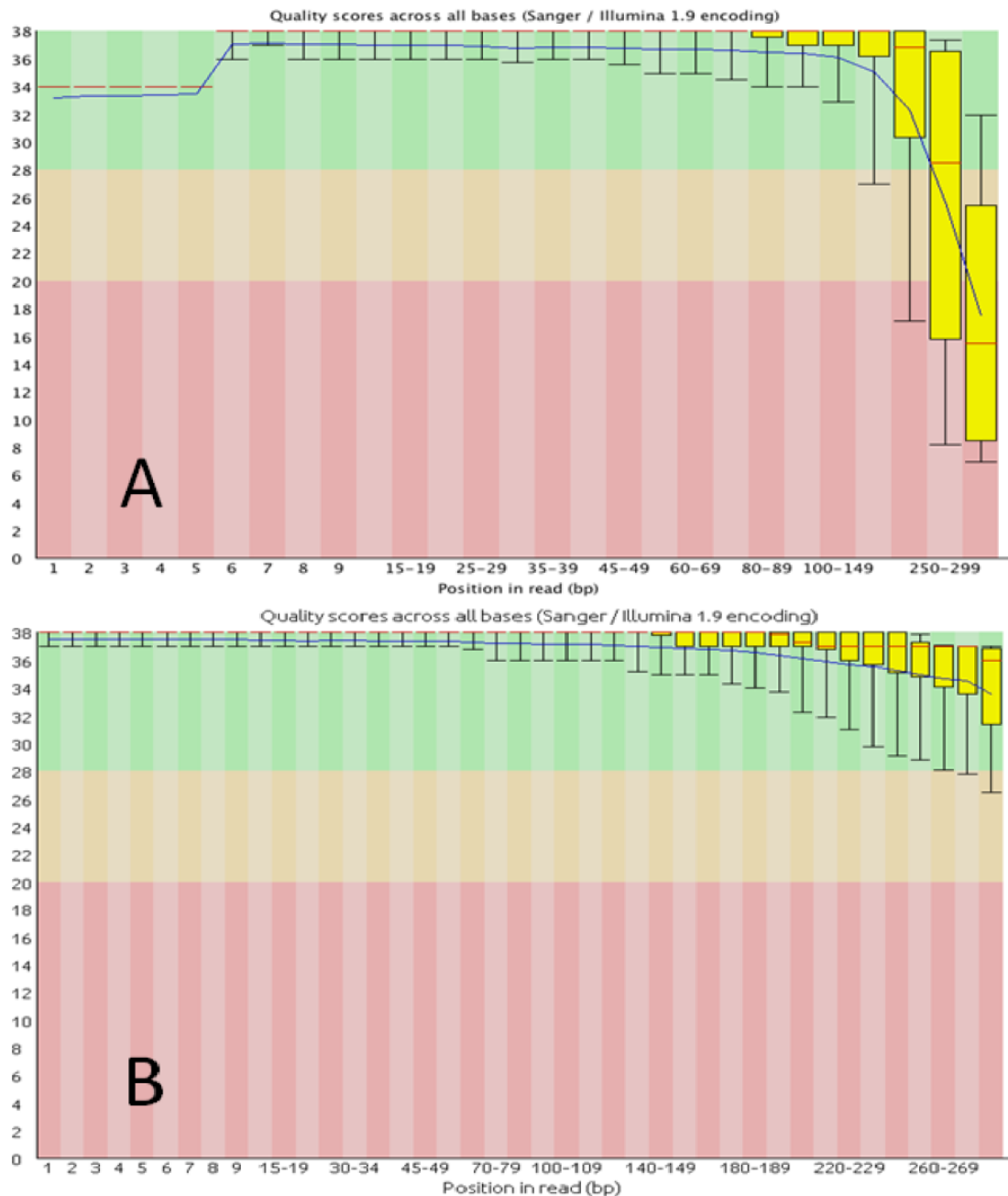


Figura 9 – Gráfico da qualidade geral das reads obtidas por RNA-Seq do télson de *Tityus serrulatus*. Os valores de qualidade Q de Phred são distribuídos de acordo com o tamanho, em número de bases, das reads. A-) Qualidade das reads brutas, antes do tratamento (trimagem) dos dados. Percebe-se como a qualidade cai à medida que a read se estende. B-) Qualidade das reads após trimagem. No gráfico de caixas a linha vermelha indica o valor da mediana, as barras amarelas indicam os quartis 25% e 75%, os traços que limitam inferior e superiormente indicam os 10% e 90% dos dados. A linha azul indica a qualidade média das bases sequenciadas. A faixa verde, laranja e rosa indicam os valores considerados bons, razoáveis e ruins respectivamente. (FastQC 0.10.1 - <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

5. MÉTODO

5.1 Busca por genes ortólogos

Uma montagem pode privilegiar a reconstrução de alguns transcritos em detrimento de outros (CLC *user manual*). Por isso, averiguou-se quão completa a montagem se apresentava em seu conteúdo de genes ortólogos. Genes ortólogos são aqueles que apresentam mesma função e origem ancestral comum, estão associados a eventos de especiação e, geralmente, ligados a funções básicas da célula, por isso são tão conservados dentro do clado. O programa *Benchmarking Universal Single Copy Orthologs v3* (BUSCO v3) (SIMÃO *et al.*, 2015), propõe uma quantificação da completude de uma montagem baseada na expectativa evolutiva que temos a respeito do conteúdo de genes ortólogos num determinado organismo. O BUSCO v3 possui um banco de dados (OrthoDB) dividido em 6 grandes clados filogenéticos, um deles abriga o *Arthropoda abd9* com 1.066 genes ortólogos de 60 espécies de artrópodes que foi utilizado na busca por ortólogos na montagem K24_B145.

5.2 Anotação por buscas de similaridade

5.2.1 Utilização de blastx contra NR e SwissProt

A montagem K24_B145 teve seus *contigs* alinhados localmente por blastx contra os bancos de dados NR e SwissProt. O intuito era conhecer quais proteínas são similares aos *contigs* e obter maiores detalhes desses alinhamentos retornados pelo programa BLAST®. Apresentar similaridade com outras proteínas pode ser um indicativo de qual função determinada molécula desempenha. As análises ocorreram no servidor Sagarana HPC *cluster*, CPAD-ICB-UFMG e os bancos de dados foram formatados pelo programa makeblastdb. Segue um exemplo da linha de comando empregada, com explicações adaptadas do manual do usuário do BLAST® (<https://ncbi.nlm.nih.gov/books/NBK279690/>):

```
blastx1 -query2 Montagem_K24_B145.fasta -db3 Banco_de_Dados.fasta -out4 Result_blastx.out -outfmt5 '6 qseqid qlen qframe sseqid sacc stitle slen qstart qend sstart send ppos pident evalue length nident positive mismatch' -max_target_seqs6 3 -num_threads7 24
```

1 Nome do programa que está sendo requisitado a atuar.

2 Especificar o nome do arquivo que será questionado pelo programa contra banco de dados referência.

3 Especificar o nome do arquivo que contém o banco de dados referência.

4 Especificar o nome do arquivo de saída, contendo os resultados obtidos pelo programa.

5 Especificar o formato do arquivo de saída, com todas as informações que se seguem:

6 este número é o código para formato tabular;

| | |
|----------|--|
| qseqid | identidade da sequência <i>query</i> ; |
| qlen | comprimento da sequência <i>query</i> ; |
| qframe | <i>frame</i> da sequência <i>query</i> na qual o alinhamento ocorreu; |
| sseqid | identidade da sequência <i>subject</i> (banco de dados); |
| sacc | número de acesso da sequência <i>subject</i> ; |
| stitle | título da sequência <i>subject</i> ; |
| slen | comprimento da sequência <i>subject</i> ; |
| qstart | em qual aminoácido a sequência <i>query</i> inicia o alinhamento; |
| qend | em qual aminoácido a sequência <i>query</i> termina o alinhamento; |
| sstart | em qual aminoácido a sequência <i>subject</i> inicia o alinhamento; |
| send | em qual aminoácido a sequência <i>subject</i> termina o alinhamento; |
| ppos | porcentagem de positivos (sinônimo do conceito de similaridade proteica) do alinhamento local das duas sequências <i>query</i> e <i>subject</i> ; |
| pident | porcentagem de identidade do alinhamento local das duas sequências <i>query</i> e <i>subject</i> ; |
| evaluate | valor esperado que calcula a probabilidade de encontrar um alinhamento idêntico entre a <i>query</i> e a <i>subject</i> , com outra sequência do banco de dados; |
| length | comprimento do alinhamento analisado; |
| nident | número de aminoácidos idênticos, na mesma posição dentro do alinhamento; |
| positive | número de aminoácidos similares, na mesma posição dentro do alinhamento |
| mismatch | número de aminoácidos não idênticos, numa determinada posição dentro do alinhamento |

6 Número máximo de sequências similares a serem reportadas por *query* questionada.

7 Número de CPUs que devem ser usados para realizar as análises

5.2.2 Utilização de tblastn contra ArachnoServer e bancos de dados customizados

Os resultados oriundos do blastx contra os bancos públicos NR e SwissProt, compreenderam os mais diversos tipos de proteínas, desde *house keepings* até as neurotoxinas do escorpião amarelo. No entanto, com o intuito de ampliar a anotação e compreender melhor a montagem K24_B145, montou-se sete bancos de dados customizados. Três deles com o objetivo de verificar qual seria o conteúdo relativo aos tecidos presentes no télson de *Tityus serrulatus* (*i.e.* predominantemente células musculares e glandulares) que constam na montagem K24_B145. Outros três bancos para analisar quanto da montagem escolhida se relacionava com o organismo em estudo (*i.e.* escorpião e a espécie *T. serrulatus*). Um último banco de dados foi construído para computar quanto da montagem K24_B145 é relativa a componentes de peçonha.

Juntamente com estes sete bancos customizados utilizou-se também de um banco público chamado ArachnoServer (PINEDA *et al.*, 2018) que se dedica a catalogar apenas componentes de toxinas de aranhas. Todas as entradas deste banco originam do SwissProt, são manualmente curadas por especialistas e depositadas na plataforma, que abriga 1.838 toxinas de 100 espécies diferentes.

Construção dos bancos de dados customizados

Os bancos customizados denominados: “Scorpion”, “Tityus serrulatus”, “Toxin”, “Arachnida Gland”, “Arachnida Muscle” e “Epithelium Glandular” foram originados dos bancos Genbank e UniProt. Para construção dos bancos customizados o procedimento era o seguinte:

- Acesso aos sites: <https://www.ncbi.nlm.nih.gov/protein> e <https://uniprot.org>.
- No campo de busca de cada site as palavras chave, para cada banco específico, eram digitadas:

| Banco Customizado | Palavras de busca lançadas no NCBI e UniProt |
|----------------------|--|
| Scorpion | <i>scorpion</i> |
| Tityus serrulatus | <i>tityus serrulatus</i> |
| Toxin | <i>toxin</i> |
| Arachnida Gland | <i>arachnida and gland</i> |
| Arachnida Muscle | <i>arachnida and muscle</i> |
| Epithelium Glandular | <i>epithelium and glandular</i> |

- *Download* das sequências no formato fasta, originando dois arquivos para cada banco, um com as sequências oriundas do NCBI e outro com as sequências do UniProt.
- Junção dos dois arquivos em um só e retirada da redundância pelo programa CD-HIT (HUANG *et al.*, 2010).

Os parâmetros utilizados para que o programa CD-HIT retirasse as redundâncias foram os seguintes:

| | |
|---|-----|
| -c: sequence identity cut-off | 1.0 |
| -G: use global sequence identity | No |
| -aS: minimal alignment coverage (fraction) for the shorter sequence | 1.0 |

O primeiro parâmetro, -c 1.0, estabelece que apenas sequências com 100% de identidade serão *clusterizadas* (reunidas). O segundo parâmetro, -G No, estipula que o alinhamento entre as sequências não precisa ser global, da ponta carboxi-terminal à amino-terminal, de forma que uma proteína íntegra pode ser comparada a outra fragmentada. O

terceiro parâmetro, -aS 1.0, determina que o alinhamento não global entre duas proteínas só pode ser *clusterizado* se este alinhamento entre a proteína íntegra e a fragmentada possuir 100% de identidade. Os demais parâmetros do CD-HIT foram padrão do programa e a utilização deste software foi *on-line*.

O último banco de dados customizado foi construído de outra forma. Ao lançar as palavras “*tityus serrulatus*” nos campos de busca do NCBI e UniProt percebeu-se que sequências de alguns poucos escorpiões retornavam junto com as do escorpião amarelo. Estas sequências foram retiradas do banco de dados customizado “*Tityus serrulatus*” e agrupadas no sétimo banco intitulado “*Outros Animais*”. Além disso, um estudo de transcriptoma e outro de proteoma de *Tityus serrulatus* já haviam encontrado componentes da peçonha descritas para outras espécies de escorpião (ALVARENGA *et al.*, 2012; VERANO-BRAGA *et al.*, 2013). Estas sequências também foram acrescentadas ao banco “*Outros Animais*”. Este nome foi escolhido porque além das sequências de outros escorpiões, o estudo proteômico também relatou a presença de uma sequência proteica de abelha (*Diploptera punctata*) P82372 – *diuretic hormone class 2* (VERANO-BRAGA *et al.*, 2013).

Escolha do tblastn

O tblastn foi a opção deste trabalho por ter sido observado que os arquivos de saída oriundos das análises com tblastn retornavam proteínas, com valores de similaridade e identidade exigidos neste trabalho, em maior quantidade do que quando obtidas por blastx. A explicação reside no fato de que o blastx compara um banco de dados proteico (referência) com sequências nucleotídicas em questionamento (*query*), traduzindo cada *query* nas seis janelas de leitura e comparando-as contra o banco de dados proteico (GISH&STATES, 1993). Já o tblastn, realiza o trabalho inverso, confrontando sequências proteicas em questionamento (*query*) com um banco de dados nucleotídico traduzido nas seis janelas de leitura (referência) (GERTZ *et al.*, 2006) (Figura 13). Não há problema com *codon-usage* e a opção pelo tblastn leva em conta que componentes da peçonha escorpiônica são pequenos e que os *contigs* da montagem K24_B145 são grandes e podem abrigar vários peptídeos. Seguem maiores detalhes da linha de comando utilizada e do funcionamento do tblastn:

```
tblastn1 -query Banco_de_Dados.fasta -db Montagem_K24_B145.fasta -out Result_tblastn.out  
-outfmt '6 qseqid qacc qtitle qlen sframe sseqid slen qstart qend sstart send ppos pident  
eval length nident positive mismatch' -max_target_seqs2 3 -num_threads 24
```

1 Nome do programa que está sendo requisitado a atuar.

2 Número máximo de sequências similares a serem reportadas por *query* questionada.

Demais parâmetros iguais aos descritos para blastx.

As chances de encontrar sequências similares aumentam em até três vezes pelo fato de estipular o número 3 no parâmetro `max_target_seqs`.

A pesquisa por BLAST® é fragmentada em três partes: “*setup*”, “*preliminary search*” e “*traceback*” (<https://ncbi.nlm.nih.gov/books/NBK153387/>). Resumidamente, na fase “*setup*” o software lê os parâmetros de pesquisa, os arquivos em questionamento (*query*) e o banco de dados referência. Em seguida, baseado na *query*, produz palavras (pequenas sequências de comprimento fixo) que serão utilizadas, na próxima fase, para encontrar similaridades entre a *query* e o banco de dados. A primeira etapa da fase “*preliminary search*” é procurar, em cada sequência do banco de dados, por correspondências com as palavras geradas na fase “*setup*”.

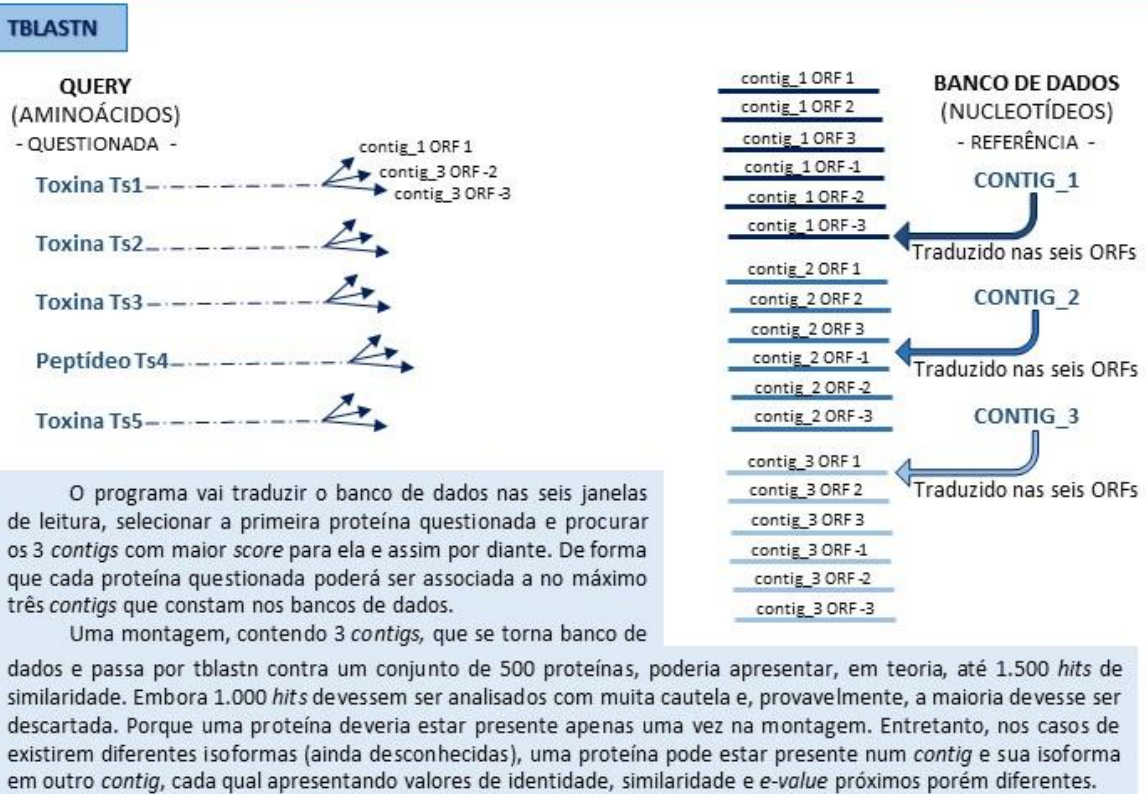
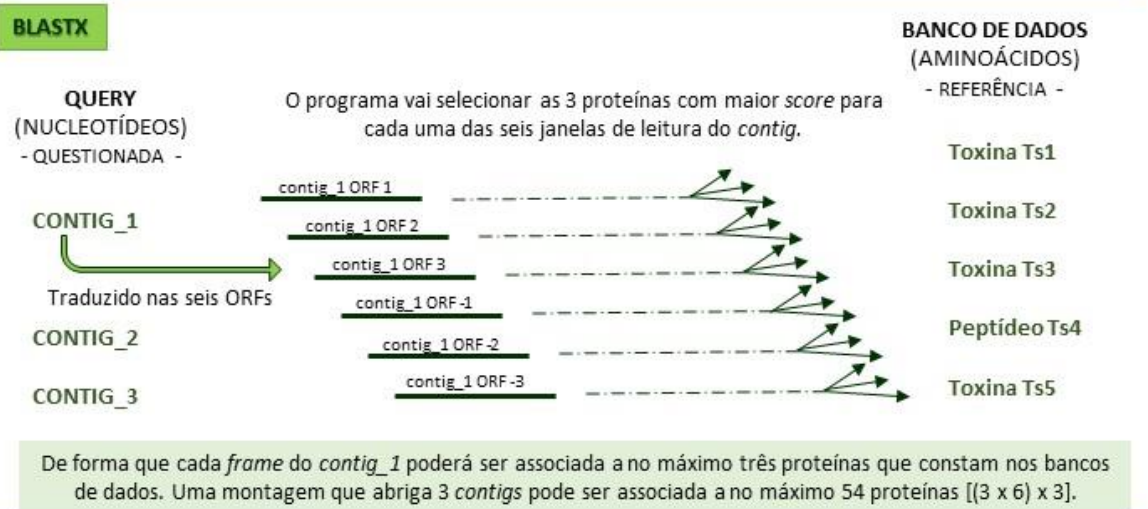


Figura 10 – Metodologias empregadas pelo Blastx e tBlastn de acordo com os parâmetros escolhidos segundo o exemplo da linha de comando anteriormente justaposta à figura.

5.2.3 Escolha dos parâmetros de filtragem dos dados após realização de blastx ou tblastn

Após obter os arquivos de saída dos diversos alinhamentos locais os resultados foram filtrados, primeiro, retirando todos os *hits* com similaridade inferior a 70%. Em seguida, o cálculo da cobertura de cada *hit* foi realizado e os que apresentavam coberturas inferiores a 70% foram retirados. Após estes dois passos um cuidadoso trabalho foi executado, de forma que todos os *hits* que apresentavam cobertura de 70%, obrigatoriamente possuíam 100% de similaridade e os *hits* que detinham 100% de cobertura poderiam conter similaridades entre 70% e 100%. Concluindo, todas as proteínas selecionadas neste trabalho apresentam ao menos 70% de similaridade com proteínas já descritas nos bancos de dados. A identidade mínima que se estipulou aceitar aqui era de 35%, mas todos os *hits* apresentaram valores superiores a esse.

No que se refere ao *e-value* (*Expect value*), foram excluídos todos os valores acima de 5×10^{-42} para os dados obtidos via blastx. Por sua vez, devido às especificidades do algoritmo tblastn, os valores esperados de tblastn receberam outro tratamento. O *e-value* indica o número de *hits*, com mesmo valor de *score*³, que se espera encontrar ao acaso ao pesquisar uma sequência contra um banco de dados de determinado tamanho. Bancos menores favorecem a ocorrência de *e-values* mais altos. Além do tamanho do banco de dados, a extensão da sequência a ser alinhada também influencia o valor esperado. Sequências menores, como é o caso de vários componentes da peçonha de *T. serrulatus*, muitas vezes classificados como peptídeos e não como proteínas, tendem a apresentar *e-value* mais elevados devido à alta probabilidade de estarem presentes ao acaso. Dito isto, os alinhamentos locais que foram realizados por tblastn, utilizando a montagem K24_B145 (26.287 seqs.) como banco de dados, foram filtrados aceitando-se *e-value* máximo de 0,028.

5.3 Utilização do BLAST2GO

Após os alinhamentos locais realizados pelo BLAST®, optou-se por utilizar outros softwares que pudessem trazer novas informações sobre a montagem K24_B145. O programa BLAST2GO (CONESA *et al.*, 2005) foi escolhido porque ele é multifuncional. BLAST2GO anota montagens genômicas e transcriptômicas realizando três abordagens diferentes: busca por

² Este valor de *e-value* foi escolhido arbitrariamente. Ao aplicar os filtros de similaridade, identidade e cobertura o maior valor de *e-value* era 5×10^{-4} e esse valor atendeu nossa expectativa.

³ *Score* é um valor numérico que descreve de maneira geral a qualidade do alinhamento. Quanto maior o *score* maior a similaridade. O *score* é calculado da seguinte maneira: soma-se números positivos que recompensam as compatibilidades parciais e totais com números negativos relacionados às incompatibilidades e os *gaps* que se abrem e/ou se estendem. https://ncbi.nlm.nih.gov/Web/Newsltr/V15N2/BLVie_w.html. Um exemplo gráfico pode ser visto em: <https://pdfs.semanticscholar.org/presentation/7f3b/69237d2259083a2772fa25d4227f3ed604be.pdf>

similaridade executando alinhamento local; buscas de domínios e motivos proteicos por meio do InterProScan e definição de termos de ontologia do GO. O BLAST2GO executa estas funções em quatro etapas chamadas *BLAST*, *InterProScan*, *Mapping* e *Annotation*.

A versão gratuita do BLAST2GO 4.0.7 foi utilizada para realizar as quatro etapas. Os parâmetros foram *padrão* do programa. No primeiro passo, o tipo de alinhamento local escolhido foi o *blastx*, o *e-value* menor ou igual a 1×10^{-5} e o banco de dados utilizado foi Local NR Genbank, cujo *download* foi realizado em janeiro de 2017. Os *contigs* foram traduzidos nas seis janelas de leitura (*frame*) e confrontados com o banco de dados local proteico, quando havia similaridade eles recebiam etiquetas com o nome: *blasted*. Quando não era possível verificar similaridade: *no-blast*.

Na segunda etapa, a busca por domínios e motivos proteicos foi realizada pelo *InterProScan*. Ele seleciona cada *contig*, traduz nas seis janelas de leitura e escolhe a maior delas para enviar ao EBI (*European Bioinformatic Institute*) onde as sequências são analisadas e retornam com as especificações de motivos e domínios proteicos. O *InterProScan* também é capaz de acrescentar a estes dados os termos GO relacionados a cada um dos motivos e domínios proteicos.

Na etapa *Mapping*, o BLAST2GO seleciona todos os *contigs* que apresentaram a etiqueta *blasted*, na primeira etapa, e busca por termos do GO relacionados àquela sequência. Somente os *contigs blasted*, que têm termos do GO atribuídos a eles, recebem a etiqueta: *mapped*.

O último passo que o BLAST2GO realiza é o *Annotation*, que seleciona apenas os *contigs mapped* e aplica a eles o que o programa denomina “regra de anotação”. Esta regra procura pelos termos do GO que foram atribuídos com maior especificidade e confiabilidade e os marca com a etiqueta *annotated*. Os gráficos mais completos gerados pelo BLAST2GO podem ser obtidos a partir deste momento.

5.4 Anotação por termos ontológicos (GO)

Além do BLAST2GO, optou-se por utilizar mais um programa que retornasse termos de ontologia, porém pelo emprego de outro algoritmo. O HMMER 3.1 - HMMER2GO foi, então, selecionado. Este software se utiliza de modelos probabilísticos chamados, na tradução livre, de Modelos Ocultos de Markov (*Hidden Markov Models*). O HMMER2GO seleciona cada *contig*, realiza a busca por ORFs (*Open Reading Frames*), identifica os domínios proteicos a partir do banco de dados Pfam ou outros bancos que participam do InterPro e em seguida retorna os termos associados ao GO. Para interpretar os códigos do GO, liberados na última etapa do

HMMER2GO, e associá-los a uma das três ontologias, foi utilizada a ferramenta *on-line* WEGO 2.0 (*Web Gene Ontology Annotation Plot*) (YE *et al.*, 2006).

Houve um problema com a execução da primeira etapa do HMMER2GO e foi preciso recorrer ao auxílio de outro programa que prediz ORFs, o EMBOSS: GetORF *on-line* (<http://bioinformatics.nl/cgi-bin/emboss/getorf>). Os parâmetros foram *padrão* do programa, exceto por: *Minimum nucleotide size of ORF to report = 27*; *Type of output = Translation of regions between START and STOP codons* e *Number of flanking nucleotides to report = 0*. Assim que o GetORF liberou os resultados de ORFs com no mínimo 9 aminoácidos, este arquivo foi submetido às demais análises performadas pelo HMMER2GO.

5.5 Anotação por vias metabólicas (KAAS-KEGG)

O programa KAAS (*Kegg Automatic Annotation Server*) (MORIYA *et al.*, 2007) versão 2.1 realizou uma anotação funcional dos genes, presentes nos *contigs*, por comparações, via BLAST, contra o banco de dados manualmente curado KEGG (*Kyoto Encyclopedia of Genes and Genomes*). O resultado aponta quais vias metabólicas estão mais enriquecidas mediante a designação de Ontologias KEGG (KO) aos *contigs* (<https://genome.jp/kegg/kaas/>).

Os parâmetros utilizados foram os seguintes: BBH (*Bi-directional Best Hit*); query sequences = nucleotide. Foram selecionados 26 organismos para estarem presentes no banco de dados KEGG para comparação com a montagem K24_B145, entre eles 21 espécies de artrópodes (Anexo A). Quatro espécies foram escolhidas pelo fato de produzirem toxinas, o *Ornithorhynchus anatinus* e três cnidários. Ao todo o banco de dados KEGG ficou composto de 414.407 genes.

5.6 Análises dos *contigs no match*

Após a realização de todas essas estratégias de anotação do transcriptoma, foi contabilizado o número de *contigs* que apresentavam similaridade satisfatória (70%) com outras sequências nos bancos de dados, ou que possuíam domínios e motivos proteicos já conhecidos. Também se contabilizou os *contigs* que estavam presentes nas vias metabólicas do KAAS ou que foram associados a termos do GO. Todos os *contigs* que não entraram nesta contabilização foram denominados *contigs no match*.

Para uma avaliação dos *no match* foi feita uma busca por ORFs e posteriormente para peptídeos sinais. O programa EMBOSS GetORF foi novamente utilizado, os parâmetros foram estipulados de forma que arquivo de saída continha a tradução das regiões entre códon de início e fim, com zero nucleotídeos flanqueadores, número mínimo de 60 nucleotídeos (20

aminoácidos) e os demais parâmetros *padrão*. Para procurar por peptídeos sinais o arquivo de saída do GetORF foi subdividido em vários arquivos menores, contendo até 2.000 ORFs cada um, e submetido ao programa SIGNAL IP 4.1 (PETERSEN *et al.*, 2011). A seleção dos parâmetros, para o SIGNAL IP, foi para grupo de organismos eucariotos, arquivo de saída curto e sem gráficos e os demais parâmetros *padrão*.

5.7 Junção dos resultados

Após todas as análises, foi feita uma contabilização de quantos contigs haviam sido anotados por, pelo menos uma, das estratégias utilizadas. Contabilizou-se primeiro os *contigs* anotados pelo software BLAST2GO. Depois se acrescentou à lista os *contigs* anotados pelo HMMER2GO e a redundância foi retirada. O mesmo procedimento foi adotado para as demais estratégias, de modo que no fim, a ordem foi a seguinte: blastx NR → blastx SwissProt → tblastn “Scorpion” → tblastn “Tityus serrulatus” → tblastn “Toxin” → tblastn ArachnoServer → tblastn “Arachnida Gland” → tblastn “Arachnida Muscle” → tblastn “Epithelium glandular” → tblastn “Outros Animais” → BUSCO → KEGG → HMMER2GO → BLAST2GO.

6. RESULTADOS E DISCUSSÃO

6.1 Busca por sequências ortólogas

A utilização do software BUSCO v3, possibilitou uma visão sobre quão completa a montagem K24_B145 está. Como genes ortólogos são aqueles presentes em inúmeras espécies que compartilham um ancestral comum, o que se espera de um transcriptoma é que apresente, senão todos, ao menos a maioria dos ortólogos para aquele grupo. A base de dados OrthoDB dispõe de 1.066 genes ortólogos catalogados para o filo Arthropoda. Ao submetermos a montagem escolhida ao programa, ele retornou a presença de 871 genes completos no total (81,7%). Ortólogos completos em cópia única são 824 (77,3%), completos em cópia duplicada (parálogos) 47 (4,4%), ortólogos fragmentados 46 (4,3%) e 149 (14%) não foram encontrados (Figura 11). A montagem K24_B145 possui um total de 26.287 *contigs*, desses, 820 *contigs* abrigam os 824 genes ortólogos completos em cópia única.

Num outro trabalho, realizado pelo nosso grupo, com análises do transcriptoma da glândula de peçonha da aranha *Loxocles similis*, 33 montagens diferentes foram testadas quanto ao conteúdo de genes ortólogos de artrópodes. As porcentagens variaram entre 19,4 a 91,8% de presença dos genes ortólogos nas referidas montagens (dados não publicados). Mediante a realização de blastx contra as proteínas já descritas para *Loxocles similis*, a montagem que mais encontrou proteínas de *L. similis*, não foi a que mais apresentou ortólogos.

Como a montagem K24_B145 apresentou um resultado maior do que 70%, a consideramos uma boa montagem para prosseguir nas análises.

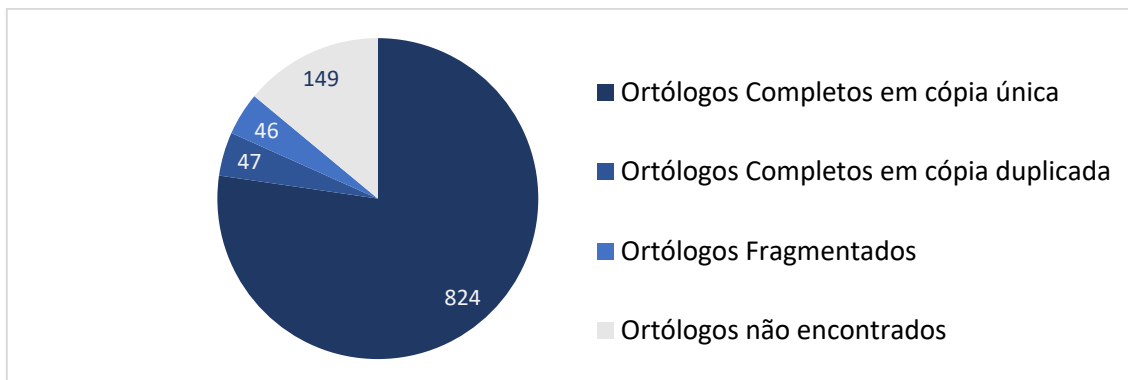


Figura 11 - Número de genes ortólogos catalogados para artrópodes presentes na montagem K24_B145

6.2 Buscas por similaridade

6.2.1 Similaridade por blastx e tblastn

O alinhamento local (blastx) realizado contra os bancos de dados NR e SwissProt, após as filtrações, retornaram 7.620 e 3.261 *hits* de proteínas respectivamente. O alinhamento local (tblastn) realizado contra o banco de dados (BD) ArachnoServer retornou 8 *hits*. Já os alinhamentos locais (tblastn) contra os BD personalizados retornaram: 13.819 *hits* para o BD “Scorpion”; 6.863 *hits* para o BD “Arachnida Gland”; 6.607 *hits* para o BD “Toxin”; 303 *hits* para o BD “Arachnida Muscle”; 210 *hits* para o BD “Tityus serrulatus”; 13 *hits* para o BD “Outros Animais” e 4 *hits* para o BD “Epithelium Glandular”; (Tabela 2).

O número de *contigs* que abrigavam as proteínas alinhadas localmente por blastx contra os bancos de dados NR e SwissProt, após as filtrações, foram 3.480 e 1.333 respectivamente. O número de *contigs* que abrigavam as proteínas alinhadas localmente por tblastn contra o BD Arachnoserver foram de 8. Já os alinhamentos locais (tblastn) realizados contra os BD personalizados retornaram: 7.202 *contigs* para o BD “Scorpion”; 2.696 *contigs* para o BD “Arachnida Gland”; 2.322 *contigs* para o BD “Toxin”; 113 *contigs* para o BD “Tityus serrulatus”; 69 *contigs* para o BD “Arachnida Muscle”; 13 *contigs* para o BD “Outros Animais” e 2 *contigs* para o BD “Epithelium Glandular” (Tabela 2).

Tabela 2 - Valores apresentados por cada um dos bancos de dados personalizados, ou não, após filtragem. | *A. Gland* – *Arachnida Gland*; *A. Muscle* – *Arachnida Muscle*; *Epithelium G.* – *Epithelium glandular*; *Outros* – *Outros Animais*.

| | <i>SCORPION</i> | <i>TOXIN</i> | <i>ARACHNOSERVER</i> | <i>A. GLAND</i> | <i>A. MUSCLE</i> | <i>EPITHELIUM G.</i> | <i>OUTROS</i> | <i>T. serrulatus</i> | <i>NR</i> | <i>SwissProt</i> |
|---|-----------------|--------------|----------------------|-----------------|------------------|----------------------|---------------|----------------------|-----------|------------------|
| | tBlastn | tBlastn | tBlastn | tBlastn | tBlastn | tBlastn | tBlastn | tBlastn | Blastx | Blastx |
| Similaridade mínima | 70 | 70 | 74,86 | 70 | 70 | 73,14 | 71,83 | 70 | 70 | 70 |
| Identidade mínima | 40,2 | 44,87 | 43,75 | 37,23 | 49,76 | 53 | 57,75 | 48,48 | 41,61 | 44,37 |
| <i>e-value</i> máximo | 0,009 | 1,00E-03 | 4,00E-04 | 0,003 | 1,00E-14 | 8,00E-34 | 5,00E-12 | 0,028 | 6,00E-05 | 3,00E-04 |
| cobertura mínima | 70 | 70 | 91 | 70 | 74 | 81 | 91 | 70 | 70 | 73 |
| Nº de proteínas | 13.819 | 6.607 | 8 | 6.863 | 303 | 4 | 13 | 210 | 7.620 | 3.261 |
| % de proteínas (<i>hits</i>) | 38,04% | 1,90% | 0,44% | 10,12% | 41,28% | 1,82% | 56,52% | 71,92% | ----- | ----- |
| Nº de seqüências do BD | 36.330 | 347.630 | 1.838 | 67.827 | 734 | 220 | 23 | 292 | ----- | ----- |
| Nº de <i>contigs</i> com <i>hits</i> | 7202 | 2322 | 8 | 2696 | 69 | 2 | 13 | 113 | 3480 | 1333 |
| % de <i>contigs</i> com <i>hits</i> | 27,40% | 8,83% | 0,03% | 10,26% | 0,26% | 0,01% | 0,05% | 0,43% | 13,24% | 5,07% |

O comprimento das quatro sequências de proteínas resultantes deste blast varia de 129 aa a 283 aa, por isso o *e-value* tão alto.

Uma possível crítica, aqui, poderia ser o valor de *e-value* mais alto do que o comum para um Blastx contra um banco de dados tão extenso quanto o NR. Se eu baixasse o valor de *e-value* para $1,0 \times 10^{-5}$ (valor mais comumente aceito) eu perderia apenas 9 sequências proteicas, das 7.620 aqui encontradas. Porém estas 9 proteínas possuem entre 18 aa e 75 aa e todas são toxinas escorpiônicas (*Tityus sp.* e *Androctonus sp.*). Por isso aceito o *e-value* mais alto do que o comum.

Tabela 2 - Valores apresentados por cada um dos bancos de dados personalizados, ou não, após filtragem. | A. Gland – Arachnida Gland; A. Muscle – Arachnida Muscle; Epithelium G. – Epithelium glandular; Outros – Outros Animais.

| | SCORPION | TOXIN | ARACHNOSERVER | A. GLAND | A. MUSCLE | EPITHELIUM G. | OUTROS | T. serrulatus | NR | SwissProt |
|------------------------|----------|----------|---------------|----------|-----------|---------------|----------|---------------|----------|-----------|
| | tBlastn | tBlastn | tBlastn | tBlastn | tBlastn | tBlastn | tBlastn | tBlastn | Blastx | Blastx |
| Similaridade mínima | 70 | 70 | 74,86 | 70 | 70 | 73,14 | 71,83 | 70 | 70 | 70 |
| Identidade mínima | 40,2 | 44,87 | 43,75 | 37,23 | 49,76 | 53 | 57,75 | 48,48 | 41,61 | 44,37 |
| <i>e-value</i> máximo | 0,009 | 1,00E-03 | 4,00E-04 | 0,003 | 1,00E-14 | 8,00E-34 | 5,00E-12 | 0,028 | 6,00E-05 | 3,00E-04 |
| cobertura mínima | 70 | 70 | 91 | 70 | 74 | 81 | 91 | 70 | 70 | 73 |
| Nº de proteínas | 13.819 | 6.607 | 8 | 6.863 | 303 | 4 | 13 | 210 | 7.620 | 3.261 |
| % de hits | 38,04% | 1,90% | 0,44% | 10,12% | 41,28% | 1,82% | 56,52% | 71,92% | ----- | ----- |
| Nº de sequências do BD | 36.330 | 347.630 | 1.838 | 67.827 | 734 | 220 | 23 | 292 | ----- | ----- |
| Nº de contigs com hits | 7202 | 2322 | 8 | 2696 | 69 | 2 | 13 | 113 | 3480 | 1333 |
| % de contigs com hits | 27,40% | 8,83% | 0,03% | 10,26% | 0,26% | 0,01% | 0,05% | 0,43% | 13,24% | 5,07% |

A proteína que apresenta este *e-value* alto é a MQ7 com apenas 14 aminoácidos. Os outros *e-values* apresentam valores $\leq 4,0 \times 10^{-5}$.

Se eu baixar o valor de *e-value* para $1,0 \times 10^{-5}$ também perco apenas 9 sequências de proteínas já depositadas no SwissProt similares à ORFs da montagem K24_B145.

O banco de dados Arachnoserver recuperou 8 proteínas similares às ORFs contidas na montagem K24_B145. Por isso, pareceu oportuno listar quais são e a porcentagem de similaridade entre elas e os *contigs*:

- P0C2K0 - Phospholipase D(LbSicTox-alpha1B1b)[Loxosceles boneti] 80%
- Q7M485 - Phospholipase D(LrSicTox1) [Loxosceles reclusa] 78%
- P0C2K1 - Phospholipase D(LdSicTox1) [Loxosceles deserta] 78%
- M5AYF1 - CRISP-1 Cysteine Rich Secretory Protein [Grammostola rosea] 82%
- U1-filistatoxin-Kh2a|1212 Insecticidal toxin [Kukulcania hibernalis] 79%
- D5J6Y7 - EF-hand protein [Pelinobius muticus] 78%
- G3LU44 - TCTP1-1 Translationally-controlled tumor protein [L. intermedia] 87%
- M5B4R7 - TCTP1-1 Translationally-controlled tumor protein [G. rosea] 88%

Os nomes das proteínas do banco de dados customizado “*Outros Animais*” estão listados na tabela 4.

Os bancos de dados personalizados que retornaram maior proporção de *contigs* com *hit* foram os BD “*Scorpion*” (27%), NR (13%), “*Arachnida Gland*” (10%), “*Toxin*” (9%) e SwissProt (5%) (Figura 12). Por sua vez, os bancos de dados personalizados que, proporcionalmente aos seus respectivos tamanhos, apresentaram mais *hits* com a montagem K24_B145 foram os BD “*Tityus serrulatus*” (72%), “*Outros Animais*” (57%), “*Arachnida Muscle*” (41%), “*Scorpion*” (38%) e “*Arachnida Gland*” (10%) (Figura 13).

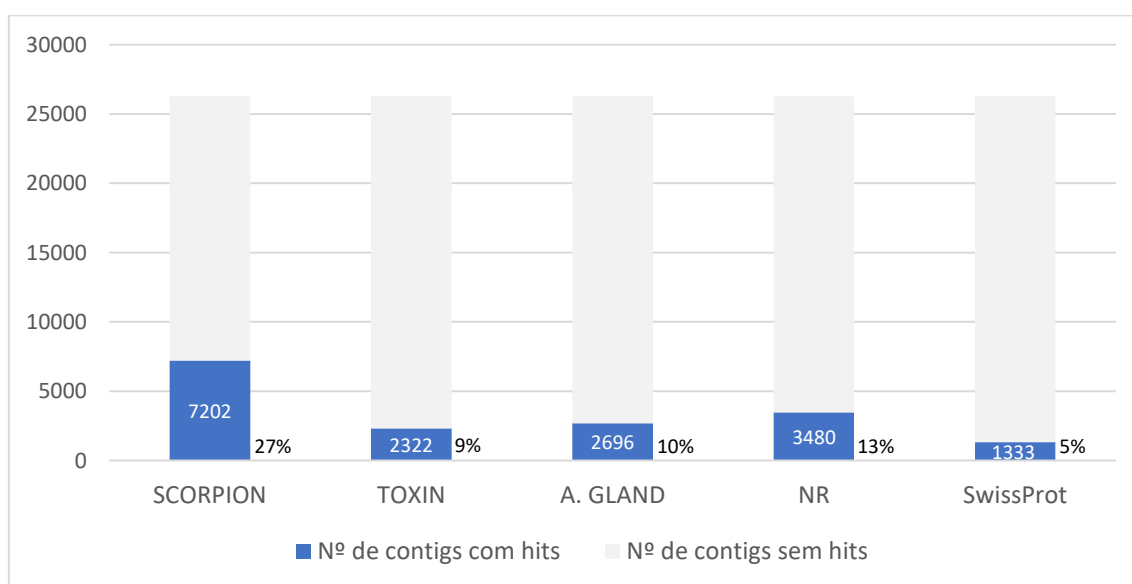


Figura 12 - Número de *contigs* que apresentaram *hit* com os bancos de dados em maior quantidade. As barras cinzas representam os 26.287 *contigs* totais da montagem K24_B145 e as barras azuis quantos *contigs* que cada banco de dados conseguiu anotar.

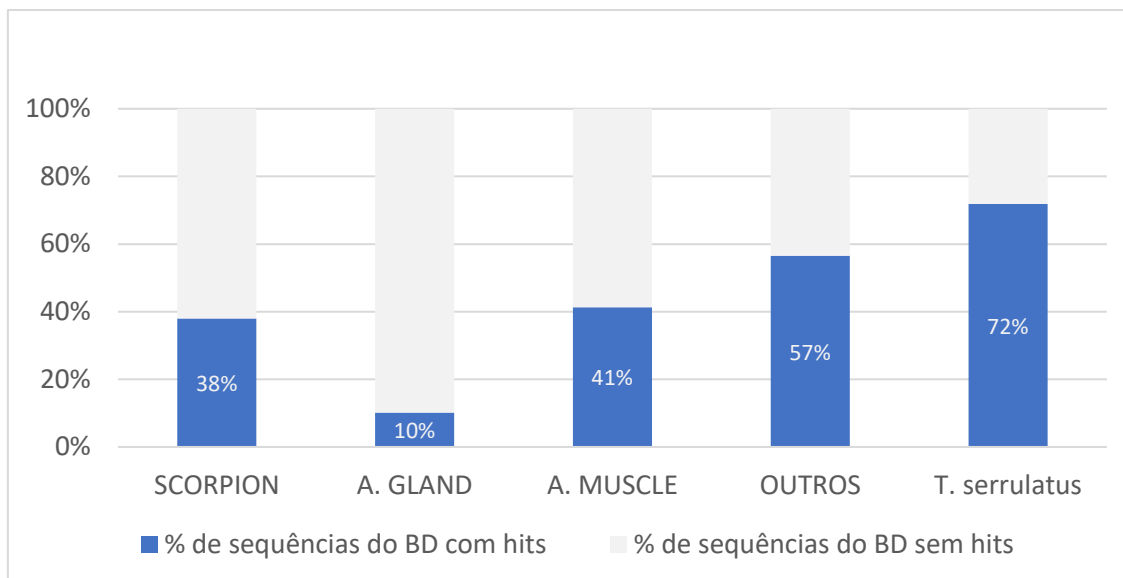


Figura 13 - Proporção de sequências dos bancos de dados que apresentaram *hit* com a montagem K24_B145. As barras cinza representam 100% de cada banco de dados e as barras azuis representam qual a porcentagem destes bancos estão presentes na montagem. Este gráfico também serve para validação da montagem.

Alguns bancos utilizados para anotação manual retornam as mesmas proteínas, alinhadas aos mesmos *contigs*. Um exemplo é a metaloprotease antarease, que aparece nos bancos SwissProt, “*Arachnida Gland*” e “*Toxin*”, com os mesmos valores de similaridade, identidade e cobertura, sempre alinhada ao *contig_252*. Por isso foi primordial a retirada das proteínas redundantes do resultado final da estratégia de anotação com uso do BLAST. Juntos, retirando as redundâncias, os alinhamentos locais apresentaram 35.302 proteínas distintas, com no mínimo 70% de similaridade com as ORFs encontradas nos *contigs* da montagem K24_B145. É possível anotar um maior número de proteínas do que o número total de *contigs* devido ao fato de existirem muitos *contigs* abrigando mais de uma proteína. O blastx contra o BD NR retornou 7.620 proteínas, já o BD SwissProt acrescentou a este valor mais 3.185 proteínas. O tblastn realizado contra o BD “*Scorpion*” acrescentou 13.554 proteínas, o BD “*T. serrulatus*” acrescentou 15, o BD “*Toxin*” acrescentou 4.978, o BD Arachnoserver mais 3, “*Arachnida Gland*” 5.737, o BD “*Epithelium Glandular*” 3 e “*Arachnida Muscle*” foi o último banco a acrescentar proteínas 207 (Figura 14).

A maior parte das moléculas similares à montagem K24_B145 estão relacionadas ao banco de dados personalizado “*Scorpion*” (38%). Os bancos “*Arachnida Gland*” (16%) e “*Toxin*” (14%) também figuram entre os que mais contribuiriam com sequências similares à montagem.

Foram disponibilizados no Apêndice D a lista das 1.896 sequências que apresentaram 100% de similaridade com as ORFs dos *contigs* da montagem K24_B145.

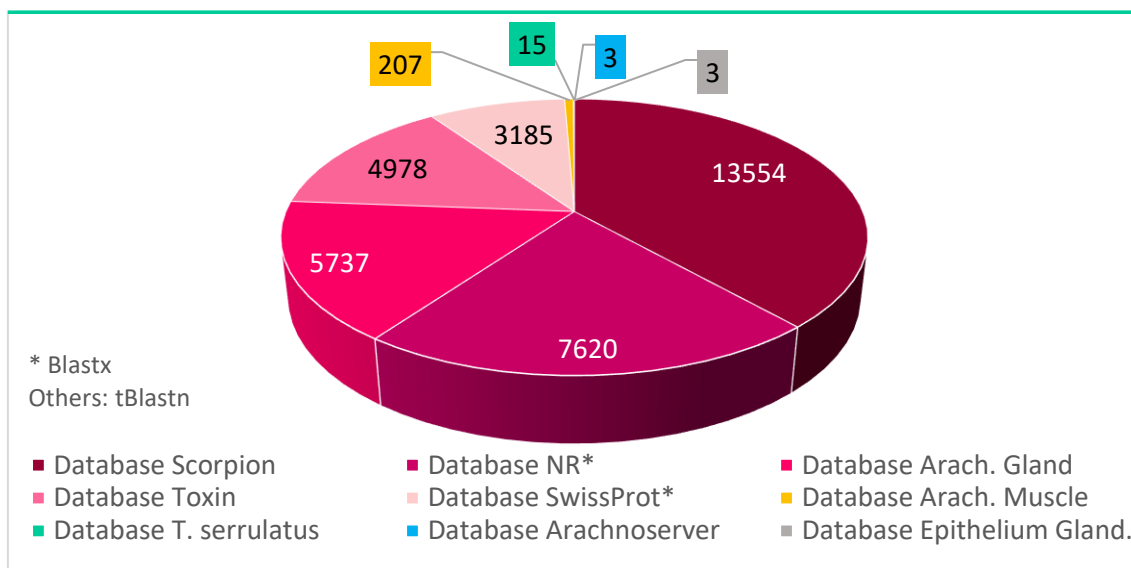


Figura 14 – Número total de proteínas com ao menos 70% de similaridade com os *contigs* da montagem K24_B145. Quantidade de proteínas, não redundantes, que cada banco acrescenta à anotação final, por alinhamento local.

Hits de componentes de peçonhas presentes na montagem K24_B145

Uma busca manual foi realizada entre as 35.302 proteínas para uma listagem das famílias dos componentes da peçonha de *T. serrulatus* já conhecidos (Tabela 3).

Tabela 3 - Número de proteínas pertencentes a determinadas famílias de proteínas na montagem K24_B145.

| Categoria | Nº | Categoria | Nº |
|--|-----|--|-------|
| Toxina - Total | 165 | EC - Enzima conversora de endotelina | 24 |
| Tx. Canal de potássio | 60 | Peptídeo Aniônico | 8 |
| Tx. Canal de sódio | 94 | Proteínas ricas em cisteínas | 60 |
| Protease - Total | 115 | Anticoagulante | 5 |
| Metaloprotease | 40 | Fosfolipase | 66 |
| Serinoprotease | 15 | Hialuronidase | 2 |
| Catepsinas | 50 | Fator de crescimento <i>Insulin-like</i> | 13 |
| Quimotripsina (serinoprotease) | 6 | Proteína hipotética conservada | 15 |
| Peptídeo Antimicrobiano | 23 | Proteínas putativas | 3.201 |
| Hipotensina | 8 | Proteínas preditas | 138 |
| ACE - Ezima conversora de Angiotensina | 11 | Proteínas prováveis | 250 |
| Alérgenos | 15 | Proteínas hipotéticas | 593 |

Não são as 2 hialuronidases encontradas por Horta *et al.* (2014)

É importante salientar que a tabela 3 não representa as proporções reais da presença dos componentes na peçonha. Essas famílias estão presentes na montagem, mas somente uma análise por RPKM (*Reads Per Kilobase of exon model per Million mapped reads*) e uma anotação manual cuidadosa poderia afirmar as proporções de cada uma. Por exemplo, algumas das 60 toxinas atuantes em canais de potássio estão alinhadas no mesmo *contig*, na mesma posição. De forma que uma delas é que deve ser escolhida, manualmente, para representar a anotação daquele transcrito. É óbvio que o transcriptoma da glândula não engendra 35.302 transcritos simultaneamente. Reforço que o objetivo foi encontrar o máximo de moléculas similares à montagem, para posterior análise manual e escolha dos melhores resultados da anotação do transcriptoma.

Proteínas hipotéticas, putativas, preditas e prováveis perfazem 11,8% do total das 35.302 proteínas. Hipotéticas são aquelas preditas por programas de identificação de genes, durante uma análise genômica. O programa identifica uma ORF, mas aquela sequência não alinha com nenhuma outra proteína já conhecida nos bancos de dados. As putativas são identificadas da mesma maneira que as proteínas hipotéticas, mas elas apresentam similaridade com sequências já conhecidas nos bancos de dados. Já as proteínas preditas e prováveis parecem ser variações dos nomes hipotética e putativa, respectivamente.

Hits dos componentes de peçonha anteriormente descritos para Tityus serrulatus

O banco de dados customizado “*Tityus serrulatus*” contém 292 proteínas já descritas para a espécie. Com o uso da estratégia de anotação manual, 210 proteínas foram encontradas (Tabela 2). Entre elas, vários componentes da peçonha. A tabela 4 lista cada um dos componentes da peçonha de *Tityus serrulatus* aqui encontrados com no mínimo 70% de similaridade. A quantidade de componentes indica a qualidade da montagem K24_B145. Notar que a tabela 4 possui 110 componentes de peçonha, as outras 100 proteínas descritas para *T. serrulatus* possuem outras funções que não toxinas.

Tabela 4 - Componentes da peçonha de *Tityus serrulatus*, anteriormente descritos, agora encontrados neste estudo transcriptômico. *(c200)= Esta sequência já havia sido descrita para *T. serrulatus* na tese de Martins, 2016; mas foi depositada no NCBI em 2017 por outro grupo que também a encontrou em estudos transcriptômicos. Martins, nomeia a sequência de c200. | Alvarenga *et al.*, 2012= no seu estudo transcriptômico de télson de *T. serrulatus* descreve estas sequências. | ^A= Alvarenga *et al.*, 2012 encontra componentes de peçonha anteriormente descritas para outras espécies de escorpiões. | ^B= Componentes da peçonha de outros escorpiões que o SwissProt retorna quando realizamos a busca apenas para a espécie *Tityus serrulatus*. | ^C= Verano-Braga e colab. (2013) em seu estudo de proteoma do *Tityus serrulatus* encontraram proteínas de outros animais entre os componentes da peçonha. | MESEU= *Mesobuthus eupeus* (escorpião); HOTJU= *Hottentotta judaicus* (escorpião), pttv.= putative; TITCL= *Tityus clathratus*; TITBA= *Tityus bahiensis*; TITCO= *Tityus costatus* e DIPPU= *Diploptera punctata* (abelha)

| Categoria | Nº | Categoria | Nº |
|--------------------|--------------------------------|-------------------------------|------------|
| Potassium channel | ----- | Metaloserrulase | ----- |
| Ts7 | P46114.1 | 19 | AMO02561.1 |
| Ts8 | P69940.2 | 20 | AMO02562.1 |
| Ts9 | P56219.2 | Alérgeno - Ts4 | O77463.1 |
| Ts11 | POC174 | Hipotensina | ----- |
| Ts12 | P0C175.1 | Ts 10 - BPP | Q9TWR4.1 |
| Ts13 | P0C176.1 | Ts14 (TsHpt-I) | P84189 |
| Ts16 | P86271 | Ts14 (TsHpt-II) | P84190 |
| Ts19 | P86822.2 | Ts14 (TsHpt-III) | P84191 |
| Similar Ts12 | Alvarenga <i>et al.</i> , 2012 | Ts14 (TsHpt-IV) | P84192 |
| Putative | JAW06974.1 | BPP PAPE | P86821 |
| Putative | JAW07004.1 | Hipotensina | JAW07002.1 |
| Putative | JAW07007.1 | Antimicrobial peptide | JAW07044.1 |
| Putative | JAW07008.1 | Antimicrobial peptide | JAW07043.1 |
| Putative | JAW07010.1 | Antimicrobial peptide | JAW06999.1 |
| Putative | JAW07012.1 | Antimicrobial peptide | JAW07022.1 |
| Putative | JAW07037.1 | Cysteine rich protein | JAW07036.1 |
| Putative | JAW06990.1 | Cysteine rich protein | JAW07034.1 |
| Sodium channel | ----- | Cysteine rich protein | JAW07032.1 |
| Ts1 | P15226.2 | Cysteine rich protein | JAW07031.1 |
| Ts2 | P68410 | Cysteine rich protein | JAW07021.1 |
| Ts3 | P01496.4 | Cysteine rich protein | JAW06995.1 |
| Ts5 | P46115.1 | Cysteine rich protein | JAW06986.1 |
| Ts17 | Alvarenga <i>et al.</i> , 2012 | Angiotensin Converting Enzyme | JAW07033.1 |
| Ts18 | Alvarenga <i>et al.</i> , 2012 | Anionic peptide | JAW07020.1 |
| Putative / (c200)* | JAW07009.1 | Chitinase | JAW07015.1 |
| Putative | JAW07016.1 | Anticoagulant peptide | JAW07014.1 |
| Putative | JAW06976.1 | Putative venom toxin | JAW07011.1 |
| Putative | JAW06971.1 | Metalloproteinase | JAW07039.1 |
| Metaloserrulase | ----- | Metalloproteinase | JAW06985.1 |
| 1 | AIJ02109.2 | Metalloproteinase | JAW06982.1 |
| 2 | AIJ02110.2 | Metalloproteinase | JAW06981.1 |
| 3 | AIJ02111.3 | Metalloproteinase | JAW06972.1 |
| 4 | AIJ02112.2 | Peptidase | ----- |
| 5 | AIJ02113.2 | Serine peptidase | CDJ26713.1 |
| 6 | AIJ02114.2 | ASTL metallopeptidase | CDJ26721.1 |
| 7 | AIJ02115.2 | ASTL metallopeptidase | CDJ26720.1 |
| 8 | AIJ02116.2 | ASTL metallopeptidase | CDJ26710.1 |
| 9 | AIJ02117.2 | ASTL metallopeptidase | CDJ26709.1 |
| 12 | AMO02554.1 | ASTL metallopeptidase | AMO02539.1 |
| 13 | AMO02555.1 | ASTL metallopeptidase | AMO02538.1 |
| 15 | AMO02557.1 | ASTL metallopeptidase | AMO02537.1 |
| 17 | AMO02559.1 | Phospholipase A2 | JAW07019.1 |

Continuação da Tabela 4

| Categoria | Nº | Categoria | Nº |
|------------------|------------|---|------------|
| Phospholipase A2 | JAW07023.1 | Protease | ----- |
| Hyaluronidase | AIX87602 | Cathepsin | CDJ26703.1 |
| Hyaluronidase | POC8X3 | Cathepsin | JAW07018.1 |
| Protease | ----- | ^A MESEU Venom toxin-like peptide | E4VP01 |
| Serine Protease | AMO02568.1 | ^A MESEU Venom toxin-like peptide | E4VP08 |
| Serine Protease | AMO02564.1 | ^A MESEU Venom protein-3 | E4VP25 |
| Metalloprotease | JAW07028.1 | ^A MESEU Venom protein-4 | E4VP36 |
| Metalloprotease | JAW07027.1 | ^A MESEU Venom protein-5 | E4VP54 |
| Cathepsin | CDJ26735.1 | ^A MESEU Venom neuropeptide-3 | E4VP55 |
| Cathepsin | CDJ26733.1 | ^A HOTJU Pttv. cystein-rich secretory peptide | F1CJ75 |
| Cathepsin | CDJ26731.1 | ^B TITCL Neurotoxin TcI2 | J9PIJ6 |
| Cathepsin | CDJ26730.1 | ^B TITCL Neurotoxin TcI1 | J9PJ66 |
| Cathepsin | CDJ26715.1 | ^B TITBA Toxin TbTx5 | POC5K8 |
| Cathepsin | CDJ26714.1 | ^C TITBA Toxin Tb1 | P56611 |
| Cathepsin | CDJ26705.1 | ^C DIPPU Diuretic hormone class 2 | P82372 |
| Cathepsin | CDJ26704.1 | ^C TITCO Scorpine-like peptide Tco 41.46-2 | Q5G8A6 |

A peçonha de *Tityus spp.* exibe evidente atividade enzimática de hialuronidases. Venâncio e colaboradores (2013) além de afirmarem a ação das hialuronidases presentes na peçonha, relatam não verificar ação de fosfolipases. Porém, nenhuma das hialuronidases já descritas para a espécie foram encontradas nesta abordagem transcriptômica, ao passo que duas fosfolipases foram. Isso pode ser explicado pelo fato de que apenas uma montagem não revela todos os transcritos de um tecido, as montagens favorecem alguns transcritos em detrimento de outros dependendo dos parâmetros utilizados. Sobre as fosfolipases, nosso grupo de pesquisa analisou uma dessas sequências e encontrou uma mutação, de um aminoácido, no sítio ativo dessa enzima que caracterizava a perda de atividade. Isso pode explicar porque a sequência é encontrada, mas não apresenta atividade⁴.

Há só que se reparar que, apesar das hialuronidases específicas do escorpião amarelo não terem sido encontradas, duas hialuronidases, uma de *Tityus stigmurus* e outra de *Androctonus bicolor*, foram encontradas. A de *T. stigmurus* apresenta 100% de cobertura, 91,7% de identidade e *e-value* 2×10^{-8} . Já a hialuronidase de *A. bicolor* possui 98% de cobertura, 79,8% de similaridade, 71,3% de identidade e *e-value* 2×10^{-43} . Esta ocorrência pode ser um indicativo de que existem outras hialuronidases ainda não descritas para *Tityus serrulatus*.

⁴ Informação verbal: dra. Bárbara Bruna Ribeiro de Oliveira Mendes (barbarab.mendes@gmail.com)

Alvarenga e colaboradores (2012); em seu estudo sobre transcriptoma do télson de *Tityus serrulatus*, por ESTs; não anotam a neurotoxina Ts5. Ao invés disso eles relatam a presença de uma molécula 86% similar a Ts5, com ponto isoelétrico e peso molecular muito próximos aos apresentados por Ts5. No presente estudo encontramos o *contig_369* abrigando Ts5 e Ts17, praticamente na mesma posição. Onde Ts17 apresenta 100% de cobertura, 100% de identidade e *e-value* 3×10^{-52} . Por sua vez, Ts5 exibe 86,1% de similaridade, 84,6% de identidade e *e-value* 3×10^{-18} . Alvarenga (2012) também mostra como Ts17 é 95% idêntica a uma toxina já conhecida de *Tityus bahiensis*, a Tbt5, atuante em canal iônico para sódio. No presente trabalho Tbt5 foi alinhada em toda sua extensão, também com o *contig_369*, na mesma posição que Ts17 e Ts5, apresentando 97,67% de similaridade e 95,35% de identidade (Figura 15). Os valores de similaridade e identidade para Ts5 e Tbt5 são idênticos aos do estudo transcriptômico de Alvarenga e colaboradores (2012). Ts5 é composto por 64 aminoácidos, Tbt5 e Ts17 por 86 aa.

Proteínas putativas, provenientes de outros estudos transcriptômicos, que se supõe serem produzidas pelas glândulas de peçonha de *T. serrulatus* também foram encontradas no presente trabalho, reforçando a importância de serem investigadas por métodos experimentais, principalmente as de atividade tóxica.

Diferenças de resultados entre o transcriptoma da glândula de peçonha e os experimentos clássicos de descrição de toxinas para *Tityus serrulatus* são esperados. Um dos motivos é que um estudo proteômico relatou que 80% dos compostos proteicos da peçonha do escorpião amarelo são resultado da ação de proteólise. A quebra de cadeias polipeptídicas parece ser uma modificação pós-traducional comum para este gênero (Verano-Braga *et al.*, 2013).

6.2.2 Escolha dos parâmetros para filtrar os alinhamentos locais

Quando duas sequências proteicas são comparadas, a similaridade entre elas pode ser verificada pela presença de aminoácidos diferentes que possuem as mesmas propriedades físico-químicas, nas mesmas posições. Proteínas que exercem mesma função podem possuir identidades baixas, contanto que o sítio catalítico e a conformação geral delas sejam mantidas. No alinhamento entre neurotoxinas atuantes em canais iônicos, por exemplo, as identidades podem ser mais baixas do que 30%. A manutenção da conformação pode ser dar pela existência da similaridade entre os aminoácidos: se são neutros ou carregados, se são polares ou apolares. Assim sendo, a opção por permitir baixa proporção de identidade ($\geq 35\%$), mas porcentagens de similaridade mais elevadas ($\geq 70\%$) se deve ao fato de que, quando se trata de proteínas, a similaridade tem um papel relevante.

Como as estratégias de anotação, aqui realizadas, abrangem os vários tecidos do tétson de *T. serrulatus*, a escolha de ao menos 70% na similaridade dificulta a anotação indiscriminada de falsos positivos. Uma futura curadoria de anotação, dos dados aqui obtidos, com enfoque em determinada família de proteína, poderia aceitar valores inferiores de identidade e similaridade, com a condição de procurar por motivos e domínios conservados, por exemplo.

A cobertura de um alinhamento revela qual a proporção, de uma sequência, que foi alinhada contra uma outra sequência. Se a cobertura de um alinhamento é baixa, provavelmente o alinhamento entre as sequências não indica função, nem uma real similaridade. A opção por excluir sequências que apresentavam cobertura inferior a 70% se deve ao fato de que se a maior parte de uma proteína não está presente no alinhamento com o *contig*, então há grandes chances de que aquele alinhamento não possua real significado e a anotação não deva ser feita.

Como o *e-value* é influenciado pelo tamanho do banco de dados e pela extensão das sequência alinhadas, os alinhamentos locais que foram realizados por *tblastn*, utilizando a montagem K24_B145 (26.287 seqs.) como banco de dados, foram filtrados aceitando-se *e-value* não usual, $> 10^{-4}$. O valor esperado máximo, de 0,028, foi encontrado em um dos alinhamentos locais contra o banco "*Tityus serrulatus*", se trata da metaloprotease MQ7 que contém apenas 14 aminoácidos (Tabela 2).

É necessário esclarecer que muitos *contigs* abrigam sequências que possuem *e-value* zero ou próximos de zero, mas coberturas, identidades e similaridades ínfimas. A maioria dos trabalhos publicados relatam apenas os valores esperados (*e-value*), como indicativo de

qualidade dos seus achados, mas dados de similaridade, identidade e cobertura devem acompanhar a informação do *e-value*.

Diante do exposto, a opção por procurar por sequências similares aos *contigs* da montagem K24_B145, retornou 8.118 *contigs* anotados.

6.3 Análises do BLAST2GO

Os resultados das análises realizadas pelo BLAST2GO são bastante integrados e auxiliam numa maior compreensão da montagem K24_B145. O resultado geral da análise da montagem K24_B145 contabiliza 14.269 *contigs* que receberam a etiqueta *blasted*, indicando que todos eles apresentam proteínas similares no banco de dados NR. O BLAST2GO também encontrou domínios e motivos proteicos, com o auxílio do InterProScan, para 19.692 *contigs*. Dos 14.269 *contigs blasted*, 10.829 foram associados a termos do GO e por isso receberam a etiqueta *mapped*. Dentre os *contigs mapped*, 6.570 receberam a etiqueta *annotated*. Esta etiqueta indica que esses *contigs* foram associados a termos do GO com alto grau de especificidade e confiabilidade de acordo com a “regra de anotação” do software. Os *contigs* que não alinharam localmente (*no-blast*) e não continham domínios nem motivos proteicos (*no-interpro*) foram 12.018 e 6.595 respectivamente (Figura 16).

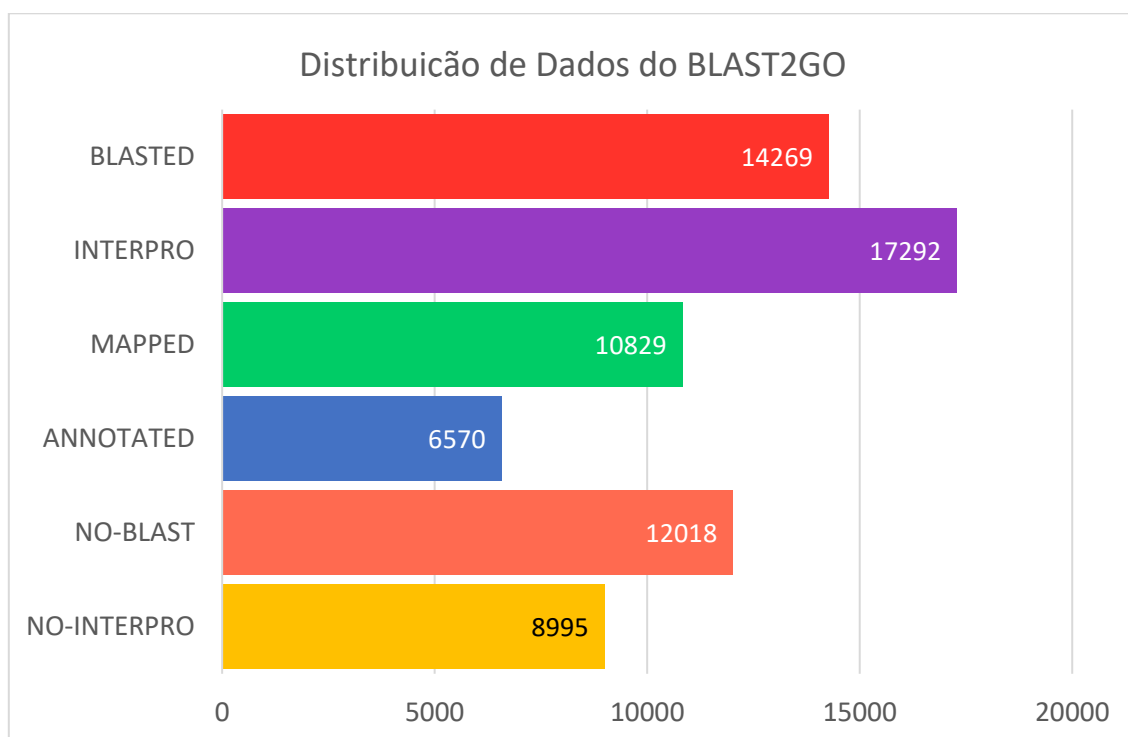


Figura 16 - Resultado geral da análise bruta da montagem K24_B145 (26.287 *contigs* totais) pelo programa BLAST2GO. No eixo x - o número de *contigs*. No eixo y - as etiquetas que o programa confere a cada *contig* dependendo do processo de anotação ao qual se enquadram.

Assim que o BLAST2GO disponibiliza quais *contigs* apresentam a etiqueta *blasted*, ele fornece, também, uma figura das espécies que mais aparecem nos resultados de similaridade (*Top-Hit*) (Figura 17). As cinco primeiras espécies, as mais representativas, são artrópodos quelicerados, em sua maioria aracnídeos. Entre as 29 espécies, há três escorpiões (*Tityus obscurus*, *Isometrus maculatus* e *Tityus serrulatus*), duas aranhas (*parasteatoda tepidariorum* e *stegodyphus mimosarum*) e quatro membros da sub-classe Acari (*Amblyomma aureolatum*, *Ixodes scapularis*, *Tetranychus urticae* e *Metaseiulus occidentalis*).

Top-Hit Species Distribution [tserrulatus_clc_k24_b145_recover]

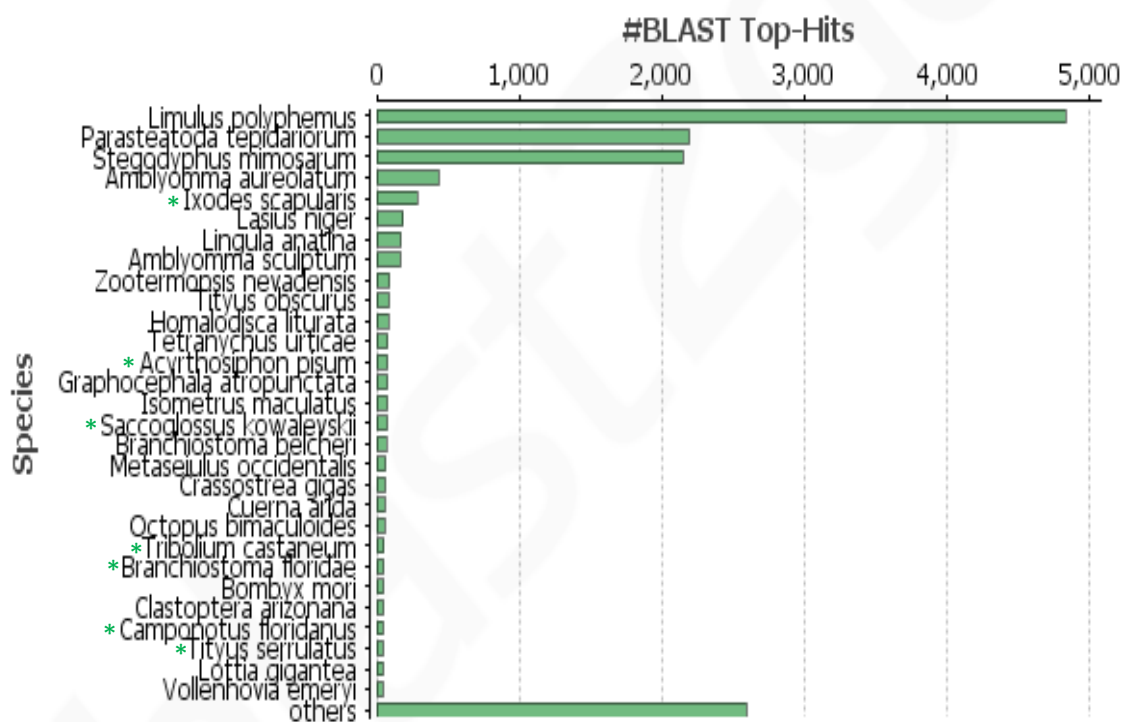


Figura 17 - As 29 espécies que mais aparecem entre os resultados de BLAST da montagem K24_B145. No eixo x - quantidade de contigs. No eixo y - nome das espécies. *Espécies também presentes entre as *Top-Hit* do estudo transcriptômico para *Tityus serrulatus* de Alvarenga e colaboradores (2012).

Os dois trabalhos transcriptômicos, anteriores a este, também fizeram uso do programa BLAST2GO. No entanto, nenhum deles especifica as quantidades de *contigs* que receberam a etiqueta *blasted*, ou foram relacionados a termos do GO ou ainda, que possuíam domínios e motivos proteicos. Apenas Alvarenga e colaboradores (2012) realizam uma análise sobre o conteúdo das três ontologias (Componente Celular, Processo Biológico e Função Molecular).

Comparando o presente trabalho com o de Alvarenga *et. al.* (2012), os resultados para Componente Celular (CC) se assemelham por conter o termo “*cell*” como o mais abundante, já

o termo “*organelle*” é o segundo mais abundante em Alvarenga *et al.* (2012), mas é o quarto mais presente aqui. Os outros termos de CC diferem completamente ressaltando a importância do presente trabalho, que acrescentou novas informações sobre componentes celulares presentes no transcriptoma do tétson do *T. serrulatus*. Os resultados para Processo Biológico (PB) se assemelham nos três termos mais abundantes: “*cellular process*”, “*metabolic process*” e “*biological regulation*” respectivamente. Ainda dentro de PB, os termos “*response to stimulus*” (6,5%) e “*localization*” (7,3%) possuem porcentagens muito próximas às encontradas aqui, 7% e 7% respectivamente. Os demais termos são completamente diferentes, ou seja, são informações novas. Os resultados para Função Molecular (FM) obtidos por Alvarenga *et al.* (2012) são mais diversificados dos encontrados aqui, mas os dois termos mais abundantes coincidem nos dois trabalhos: “*binding*” e “*catalytic activity*” (Figura 18).

Mais discussões sobre os resultados do BLAST2GO, referentes aos termos do GO associados aos *contigs*, serão abordadas numa próxima seção, onde os resultados do HMMER2GO serão comparados aos do BLAST2GO (ver item 6.5).

Alvarenga *et al.* (2012) também relata as espécies “*Top Hit*”, elencadas pelo BLAST2GO, para os resultados do blastx. Entre as 29 espécies que mais aparecem nos resultados do alinhamento local, apenas sete também são elencadas no presente trabalho (Figura 17). Explicações para este fato podem estar relacionadas com o fato de que desde 2012, muitos avanços no sequenciamento de outros organismos ocorreram, aumentando o repertório do banco de dados. Além disso, todas as divergências apresentadas entre o presente trabalho e o estudo de Alvarenga *et al.*, 2012 são influenciadas pelo tipo de obtenção e sequenciamento de cDNA e os diferentes valores de *e-value*, que cada trabalho estipulou (ver item 6.9).

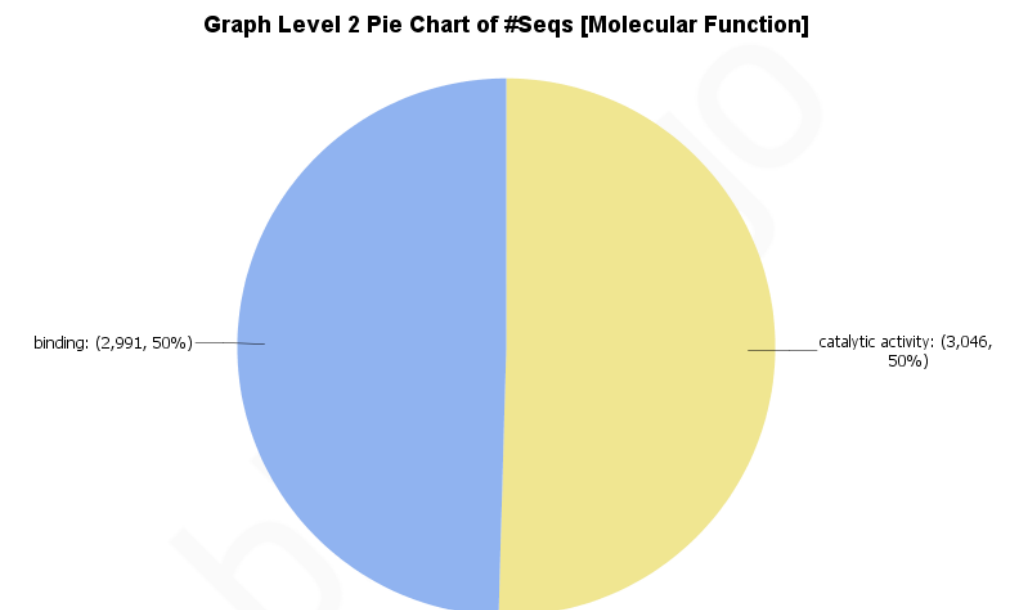
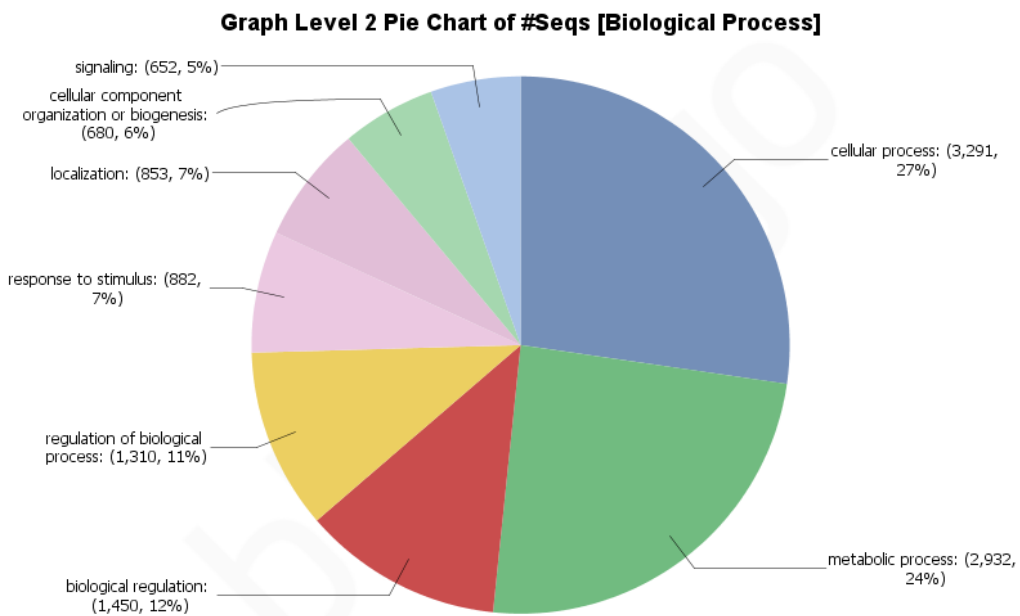
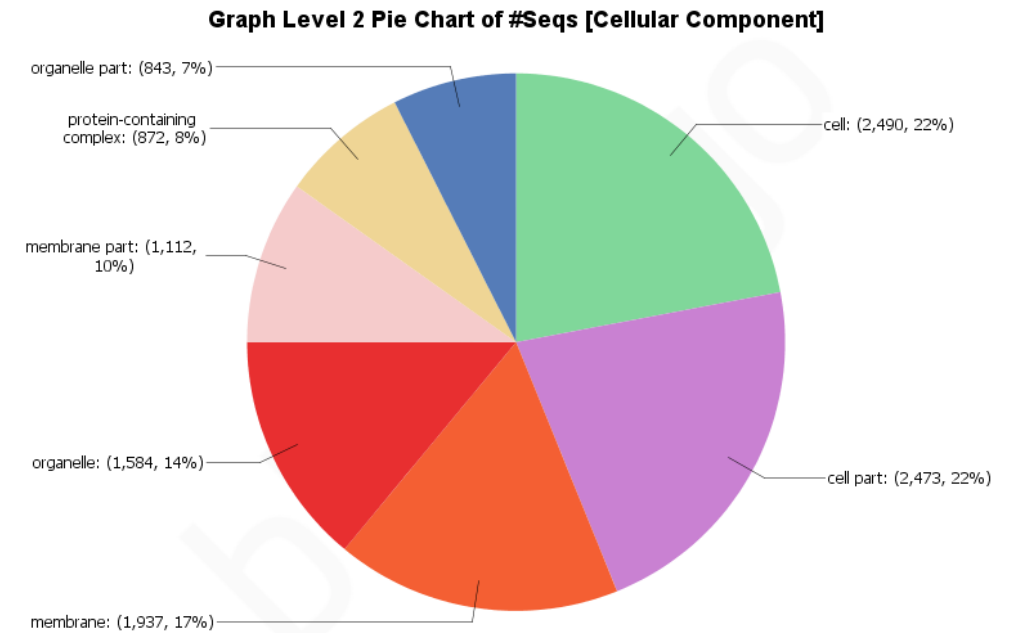


Figura 18 - Proporção de termos do segundo nível do GO para cada uma das três ontologias (Componente Celular, Processo Biológico e Função Molecular). Fonte: BLAST2GO

No presente estudo foram considerados como anotados, pelo programa BLAST2GO, todos os *contigs* que apresentaram qualquer domínio e/ou motivo proteicos. No entanto, apesar de 828 *contigs* não apresentarem domínios nem motivos proteicos eles ostentavam 70%, ou mais, de similaridade no alinhamento local, por blastx, com proteínas já descritas no banco de dados NR. Tendo em vista esse fato, acrescentamos à categoria dos anotados pelo BLAST2GO esses *contigs blasted*. No total, foram anotados pelo BLAST2GO 18.120 *contigs*, sendo 17.292 *contigs* recuperados pelo InterProScan e 828 *contigs* pelo alinhamento local (Figura 19). De fato muitos desses *contigs* anotados apenas por possuírem domínios e/ou motivos podem representar uma anotação que contenha muitos falsos positivos, porque domínios e motivos são sequências pequenas, passíveis de serem montadas ao acaso.

Os termos do GO que foram associados a motivos e domínios proteicos encontrados pelo InterProScan não são contabilizados na categoria *mapped* do BLAST2GO. Esta categoria só se refere aos termos do GO que foram distribuídos a sequências *blasted*. De modo que, dos 26.287 *contigs* totais da montagem K24_B145, 11.599 *contigs* foram associados a termos do GO: 10.829 *mapped* e 770 InterPro (Figura 19).

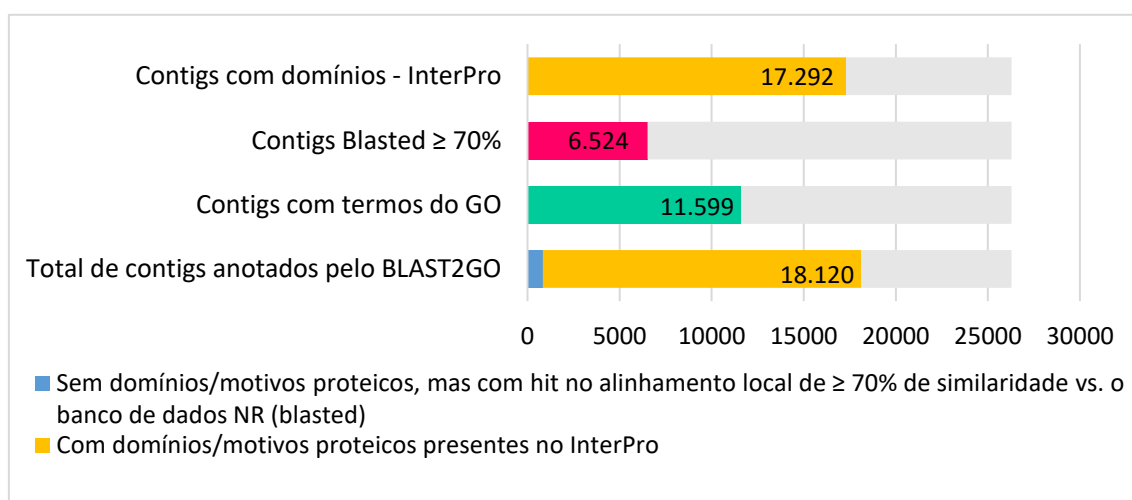


Figura 19 - Resumo da anotação funcional realizada pelo BLAST2GO para a montagem K24_B145. No eixo x – o número de contigs. No eixo y – as categorias a que os contigs pertencem

6.4 Análises do HMMER2GO

O programa GetORF, utilizado para retornar todas as ORFs entre códons de início e fim, iniciadas com metionina e possuindo no mínimo 9 aminoácidos, apresentou 685.850 ORFs como resultado. Ao usar estas ORFs como entrada no programa HMMER2GO, foram obtidos um total de 6.816 *contigs* que abrigavam 12.083 termos do GO. O software *on-line* WEGO 2.0 permitiu a visualização gráfica dos resultados do HMMER2GO e a identificação de quantos

termos cada ontologia possuía. Por fim, o HMMER2GO anotou 6.816 contigs (Figura 19). Apesar de o número ser inferior ao encontrado pelo BLAST2GO (15.039), este resultado acrescentou à anotação final da montagem K24_B145, mais 1.049 *contigs* que não haviam sido contemplados pelo programa BLAST2GO (ver Tabela 6).

6.5 Análise comparativa dos programas BLAST2GO vs. HMMER2GO

Os números e comparações dos resultados das anotações funcionais do BLAST2GO e HMMER2GO estão representados em gráficos para facilitar o entendimento (Figuras 20-25).

Ao observar a figura 20 percebe-se que, para ambos programas, os termos do GO associados a Função Molecular estão presentes em maior número, seguido dos Processos Biológicos e Componentes Celulares. Quando todos os termos GO, de cada uma das três categorias, foram somados percebeu-se uma diferença razoavelmente pequena no total: BLAST2GO encontrou 890 termos a mais que o HMMER2GO. Resultado este que não tira o mérito do segundo programa, pois nas figuras 22, 23 e 24 fica claro que, para 10 classes, HMMER2GO recuperou mais termos do GO (Figura 25). A divergência nos resultados apresentados pelos dois programas advém das diferenças nos cálculos e nas considerações dos algoritmos de cada *software*.

Deng e colaboradores (2018) também anotaram o transcriptoma do escorpião *Heterometrus spinifer* utilizando Blast2GO e o WEGO 2.0. Os termos de processo biológico foram mais abundantes no experimento deles, a disparidade pode se dar pelo fato de ser uma espécie diferente e/ou eles terem extraído RNA de glândulas repletas de peçonha, enquanto nossa extração se deu com as glândulas em fase de abastecimento, pois haviam sido drenadas alguns dias antes. Ao todo, o estudo de Deng e colaboradores (2018) encontrou 10.952 termos do GO: 1.177 para componentes celulares, 2.498 para função molecular e 7.277 para processos biológicos.

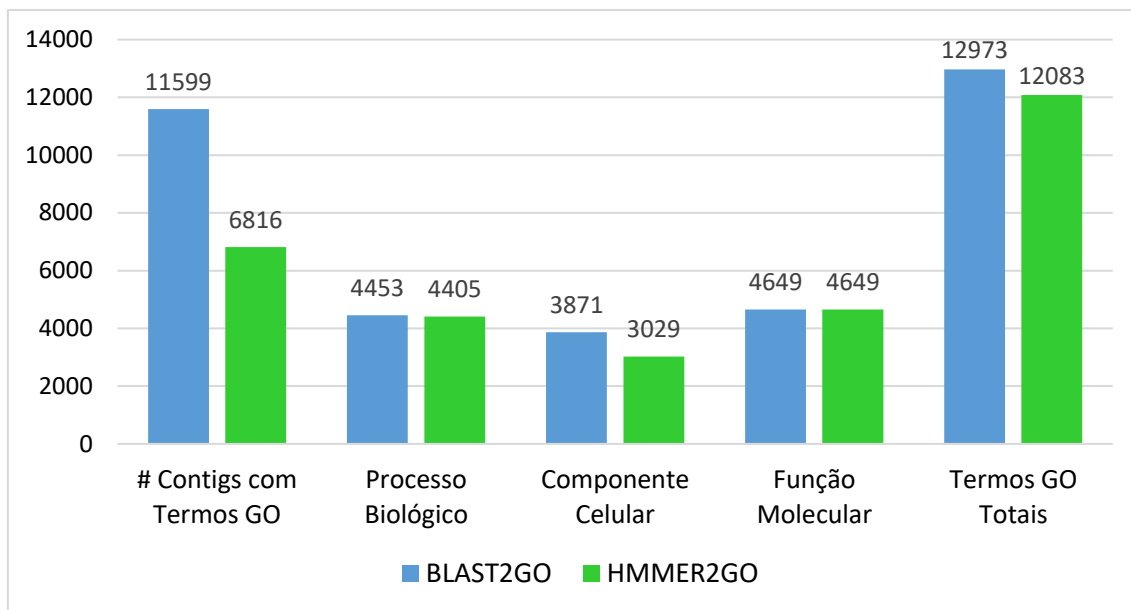


Figura 20 - Comparação da anotação funcional relativa a ontologias do GO realizadas pelos programas BLAST2GO e HMMER2GO. No eixo y – a primeira comparação, da esquerda para a direita, é referente ao número de *contigs*. Nas demais comparações os números são relativos à quantidade de termos do GO distribuídos.

Entre os Componentes Celulares anotados por GO, a presença de transcritos relacionados desde a membrana celular à organelas e outras partes intracelulares indicam a abrangência da montagem (Figura 21). Dentre as Funções Moleculares, os GO relacionados a atividade catalítica ocorrem expressivamente. Este fato reafirma o conhecimento prévio de que a peçonha de *T. serrulatus* contém diversas proteases, proteinases, hialuronidases e algumas fosfolipases. A classe “*drug binding*”, que aparece no gráfico de Funções Moleculares, possui vários termos do GO (Figura 22). Dado que corrobora a afirmação de que a peçonha do escorpião amarelo possui potencial biotecnológico e farmacêutico. Já entre os *contigs* anotados com GO para Processos Biológicos, os processos de biossíntese de moléculas e processos metabólicos estão em maior quantidade, o que condiz com nossa expectativa, afinal os RNAs das glândulas de peçonha foram extraídos em momento de reabastecimento (Figuras 23 e 24).

A complementariedade entre BLAST2GO e HMMER2GO vai além desses programas anotarem alguns *contigs* que o outro não anota. Ao pesquisar sobre quais identificadores GO foram associados a um determinado *contig* observou-se que haviam casos em que o HMMER2GO anota GOs que o BLAST2GO não relata e o contrário também é verdadeiro. Um exemplo é o *contig_10368* que pelo BLAST2GO recebe os identificadores GO:0003676 e GO:0046872, enquanto o HMMER2GO designa para o mesmo *contig_10368* os identificadores GO:0008270 e GO:0003677.

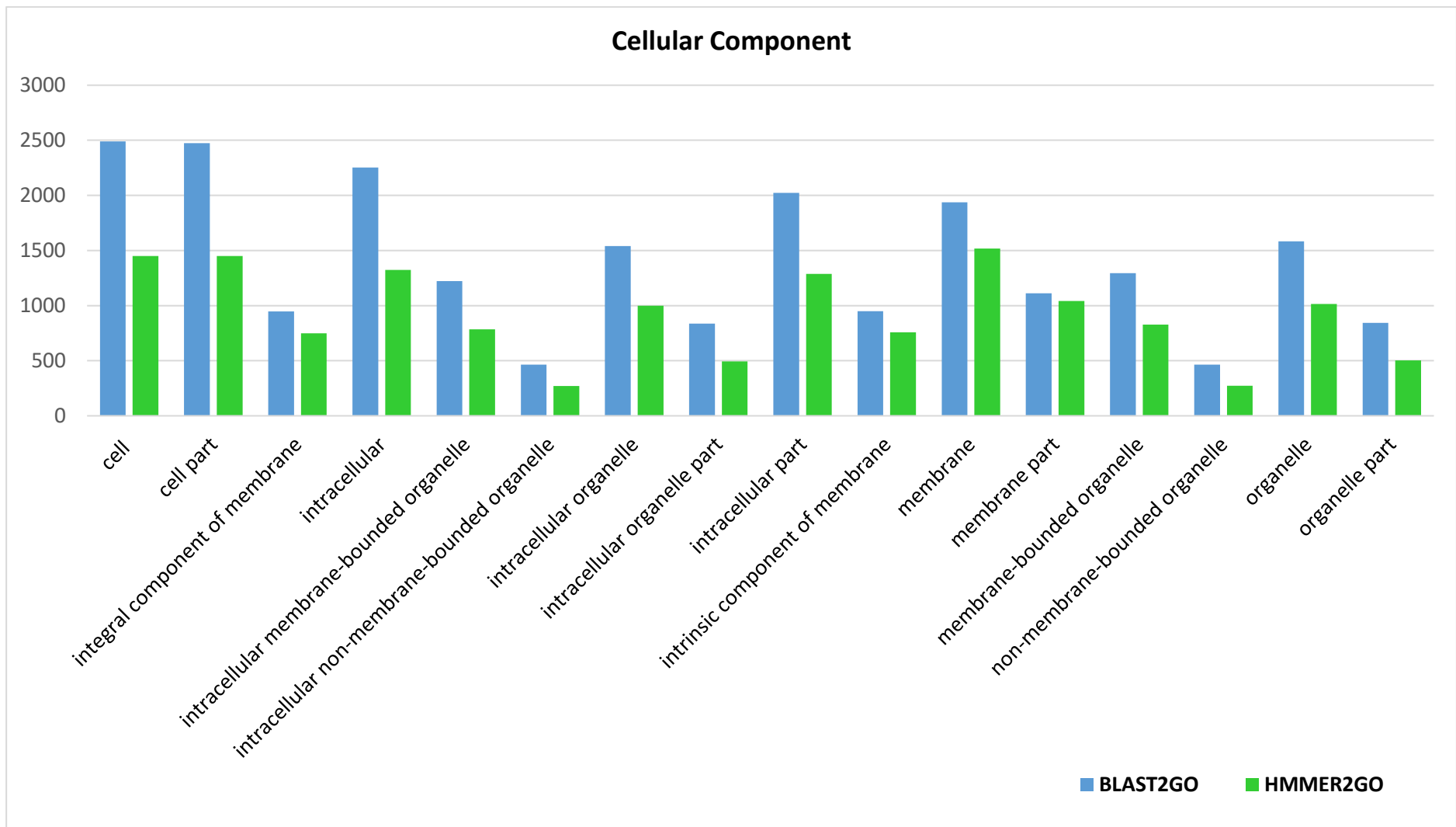


Figura 21 - Comparação entre os resultados de anotação funcional, dos programas BLAST2GO e HMMER2GO, para os termos GO relacionados a Comp. Celulares. Eixo y – Nº de termos do GO.

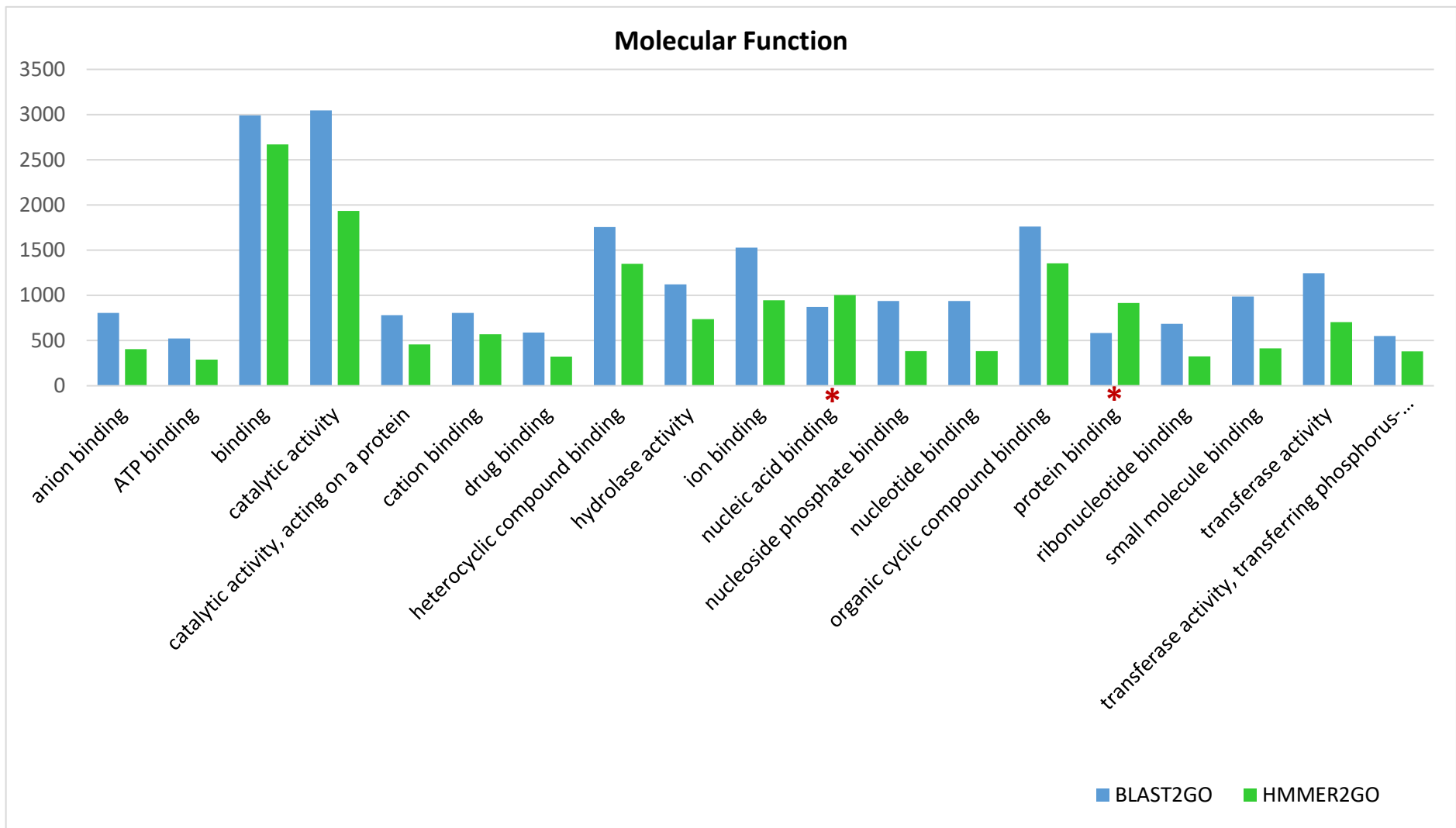


Figura 22 - Comparação entre os resultados de anotação funcional, dos programas BLAST2GO e HMMER2GO, para os termos do GO relacionados a F. Molecular. Eixo y – Nº de termos do GO.

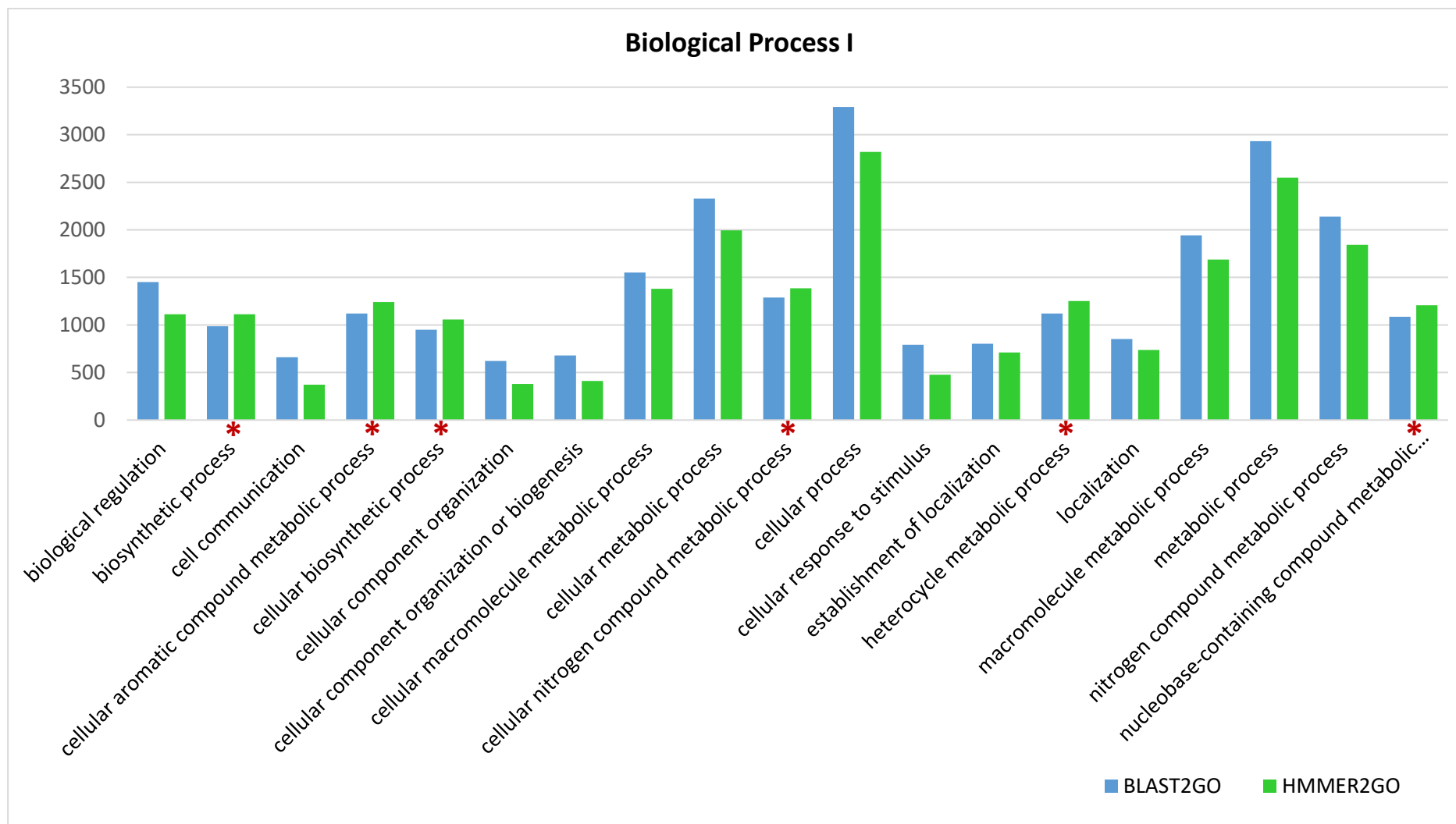


Figura 23 - Comparação entre os resultados de anotação funcional, dos programas BLAST2GO e HMMER2GO, para os termos GO relacionados a P. Celular. Eixo y – Nº de termos GO (parte 1)

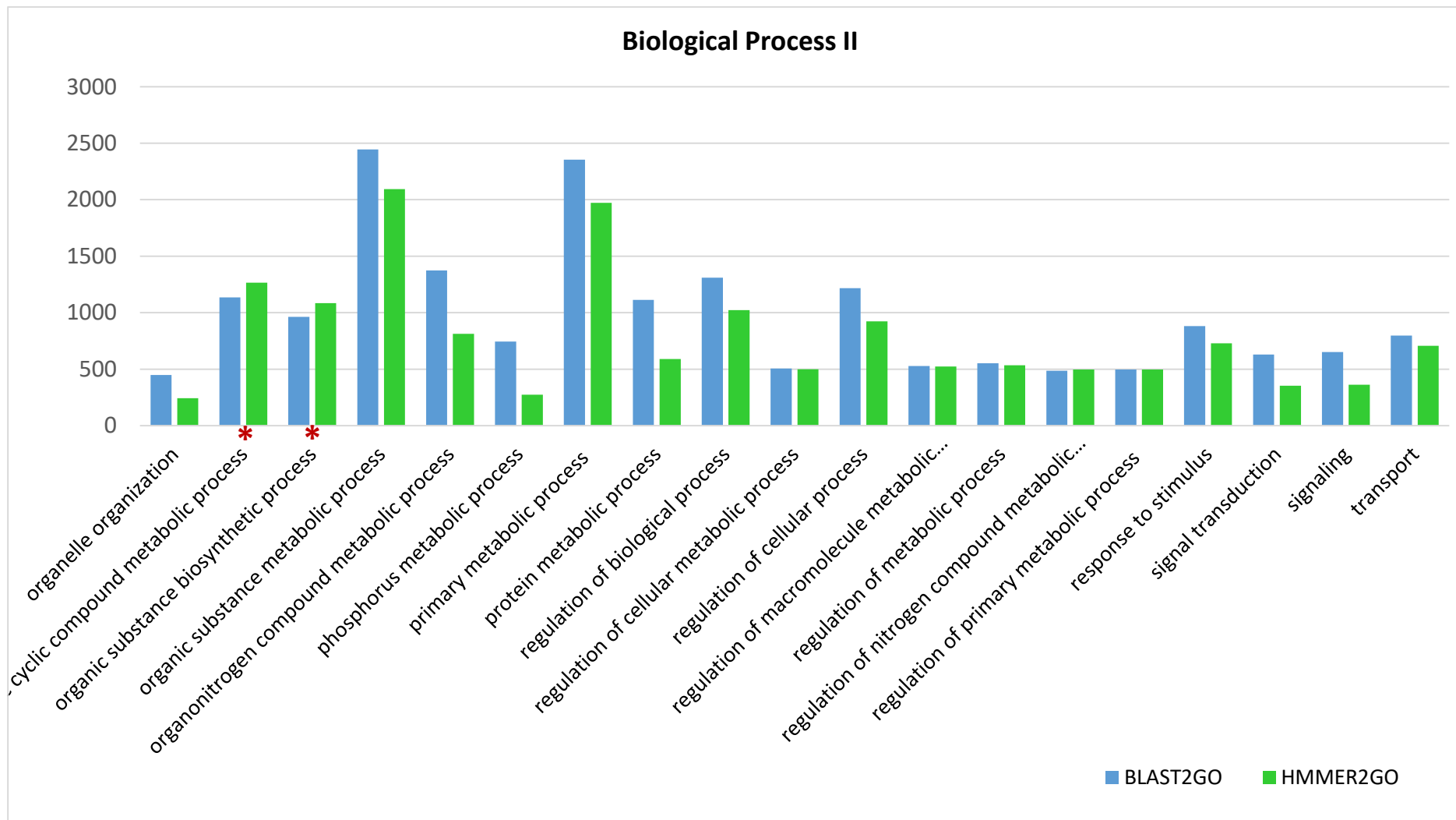


Figura 24 - Comparação entre os resultados de anotação funcional, dos programas BLAST2GO e HMMER2GO, para os termos GO relacionados a P. Celular. Eixo y – Nº de termos GO (parte 2)

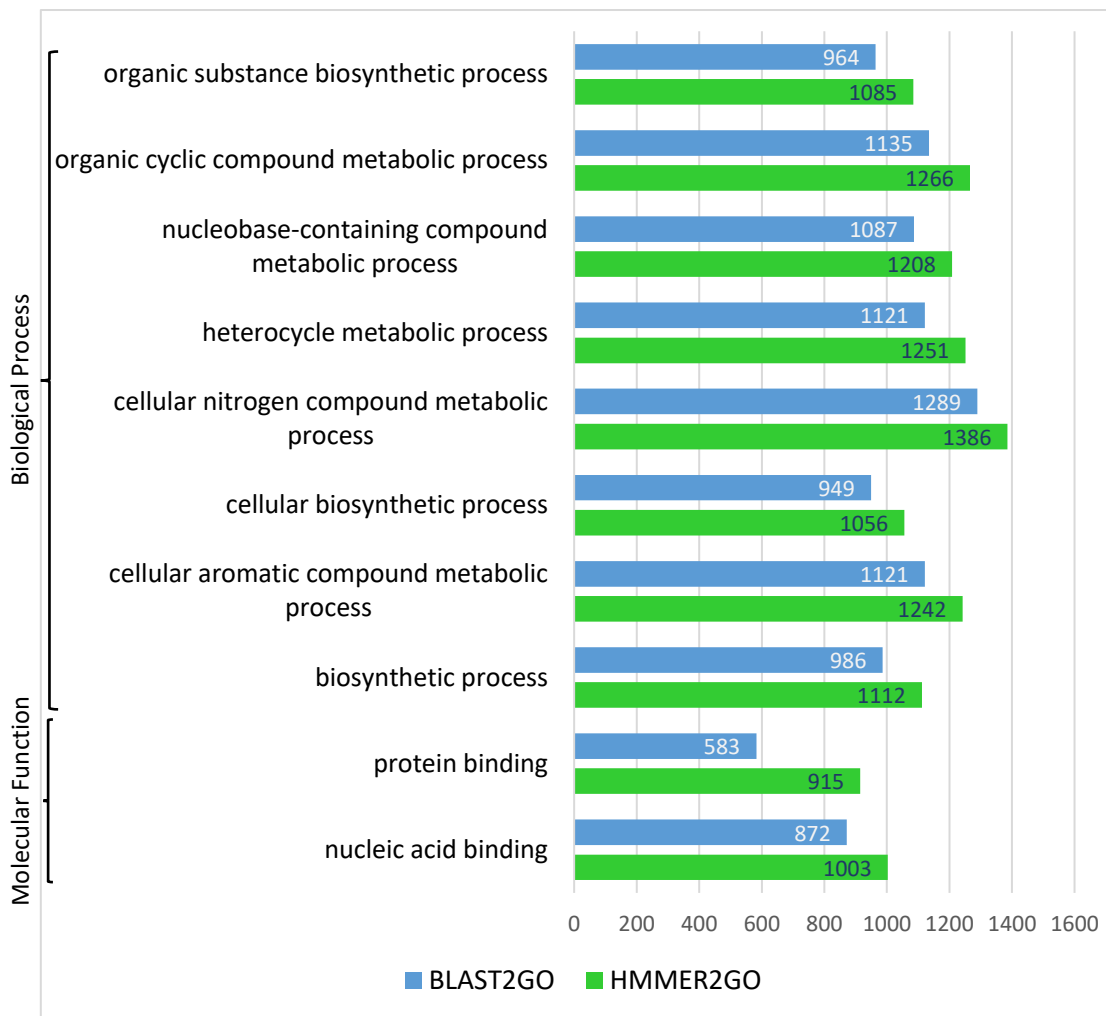


Figura 25 – Número de termos do GO anotados em maior quantidade pelo HMMER2GO comparado ao BLAST2GO.

6.6 Estratégia de anotação da montagem por vias metabólicas KAAS-KEGG

Uma outra forma de obter informação do transcriptoma montado é procurar pelas vias metabólicas mais representadas. A anotação automática pelo KEGG retornou 5.979 *contigs*. Neles foram identificadas vias que pertencem às 6 classificações principais do KEGG: “Processamento de informação genética” 4.935 (82,53%); “Doenças humanas” 3.551 (59,39%); “Sistemas do organismo” 3.085 (51,59%); “Metabolismo” 2.870 (48,0%); “Processamento de informação ambiental” 2.755 (46,07%) e “Processos celulares” 2.206 (36,89%). Os 3,72% dos *contigs* restantes ficaram sem uma classificação clara e definida e o programa os agrupou em “Unclassified” (Figura 26).

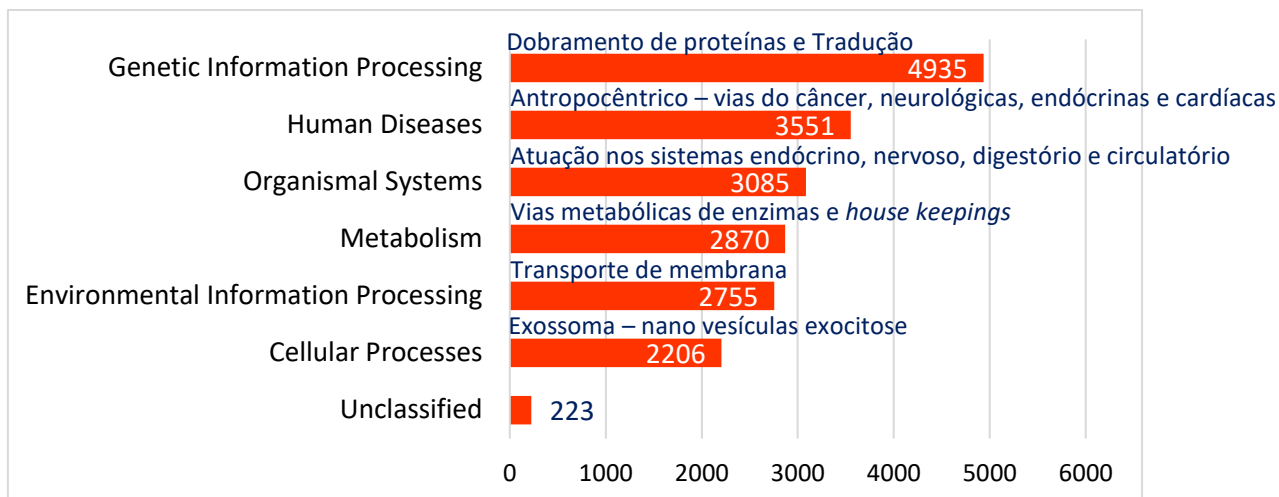


Figura 26 - Número de *contigs* dentro de cada uma das principais categorias do KEGG.

Uma tabela mais detalhada com os resultados fornecidos pelo KEGG segue na próxima página (Tabela 5). A categoria de “Processamento de informação genética” abriga subcategorias de “dobramento de proteínas” e “tradução”, estas são as que mais anotam *contigs*. Este é um resultado esperado diante da nossa hipótese de que as glândulas estão em atividade metabólica intensa para abastecimento das mesmas. “RNA Family” se apresenta zerado, esta subcategoria abriga os RNA’s não codificantes. Já a categoria “Doenças humanas” nos leva a pensar que assim como o GO, as vias metabólicas KEGG retornam anotações enviesadas para organismos modelo, predominantemente vertebrados e plantas. O campo da toxonologia é uma área à parte, que ainda cresce discretamente no que concerne a bancos de dados e programas de anotação específicos, devido às peculiaridades dos componentes da peçonha. Dentro da categoria “Sistemas do organismo”, no nível de terceiro grau (Apêndice A), percebemos vários KO’s⁵ que representam vias onde as toxinas escorpiônicas atuam, causando os sintomas típicos do envenenamento. Exemplos são KOs relacionados aos sistemas endócrino, nervoso, digestório e circulatório. Por sua vez a categoria “Metabolismo” além de evidenciar o funcionamento de vários genes *house keepings*, também salienta a grande presença de enzimas, o que corrobora com o conhecimento que se tem sobre as várias proteases, proteinases, hialuronidasas e fosfolipases presentes na peçonha de *Tityus serrulatus*. Em “Processos celulares” a subcategoria que mais anotou *contigs* foi “transporte e catabolismo”, dentro dela é possível observar que Exossoma (nano-vesículas de exocitose) são as mais enriquecidas. Resultado que confirma a expectativa em se tratando de um transcriptoma de células glandulares.

⁵ KO = Kegg Orthology. São sequências de números que identificam vias metabólicas específicas.

Tabela 5 - Quantidade de *contigs* pertencentes às categorias KEGG de primeiro (cinza) e segundo nível (branco).

| KEGG Categories | Nº of <i>contigs</i> | KEGG Categories | Nº of <i>contigs</i> |
|---|----------------------|---|----------------------|
| 09120 Genetic Information Processing | 4935 | 09100 Metabolism | 2870 |
| 09123 Folding, sorting and degradation | 1652 | 09112 Enzyme families | 777 |
| 09122 Translation | 1340 | 09107 Glycan biosynthesis and metabolism | 408 |
| 09124 Replication and repair | 1128 | 09101 Carbohydrate metabolism | 361 |
| 09121 Transcription | 815 | 09105 Amino acid metabolism | 337 |
| 09125 RNA family | 0 | 09103 Lipid metabolism | 303 |
| 09160 Human Diseases | 3551 | 09104 Nucleotide metabolism | 178 |
| 09161 Cancers | 1313 | 09108 Metabolism of cofactors and vitamins | 121 |
| 09167 Infectious diseases | 1178 | 09102 Energy metabolism | 120 |
| 09163 Neurodegenerative diseases | 286 | 09106 Metabolism of other amino acids | 85 |
| 09166 Endocrine and metabolic diseases | 281 | 09111 Xenobiotics biodegradation and metabolism | 78 |
| 09165 Cardiovascular diseases | 157 | 09109 Metabolism of terpenoids and polyketides | 66 |
| 09168 Drug resistance | 144 | 09110 Biosynthesis of other secondary metabolites | 36 |
| 09164 Substance dependence | 142 | 09130 Environmental Information Processing | 2755 |
| 09162 Immune diseases | 50 | 09132 Signal transduction | 1824 |
| 09150 Organismal Systems | 3085 | 09133 Signaling molecules and interaction | 569 |
| 09152 Endocrine system | 918 | 09131 Membrane transport | 362 |
| 09151 Immune system | 642 | 09140 Cellular Processes | 2206 |
| 09156 Nervous system | 427 | 09141 Transport and catabolism | 968 |
| 09154 Digestive system | 264 | 09143 Cell growth and death | 582 |
| 09159 Environmental adaptation | 195 | 09144 Cellular community - eukaryotes | 336 |
| 09158 Development | 156 | 09142 Cell motility | 312 |
| 09149 Aging | 149 | 09145 Cellular community - prokaryotes | 8 |
| 09153 Circulatory system | 143 | 09190 Unclassified | 223 |
| 09157 Sensory system | 103 | 09193 Cellular processes and signaling | 81 |
| 09155 Excretory system | 88 | 09191 Metabolism | 78 |
| | | 09192 Genetic information processing | 44 |
| | | 09194 Poorly characterized | 20 |
| | | 09196 Viral protein family | 0 |

O estudo transcriptômico de *Tityus serrulatus* realizado por Alvarenga e colaboradores (2012) também usou KEGG para anotar as vias metabólicas. Porém, das 25 vias mais relevantes para o trabalho deles, apenas três (*Protein processing in endoplasmic reticulum*, *RNA transport* e *Spliceosome*) também se encontram entre as 25 vias mais relevantes para o presente trabalho (APÊNDICE B). A divergência pode ser explicada pelas diferenças na obtenção e sequenciamento do cDNA entre os dois trabalhos (ver ítem 6.9).

6.7 Resultados gerais

Ao utilizar as diversas estratégias de anotação, mencionadas anteriormente, auferimos que 73,35% dos *contigs*, da montagem escolhida, possuíam *match* com proteínas já descritas nos bancos de dados pesquisados ou possuía domínios e motivos proteicos. Isto é, dos 26.287 *contigs* presentes na montagem K24_B145, 19.281 *contigs* apresentam proteínas com domínios/motivos conhecidos, ou sequências altamente similares com outras já descritas (Figura 27). Este é um bom resultado quando comparamos com outros trabalhos de anotação do transcriptoma das glândulas de peçonha de escorpiões: 47,1%, e-value < 10⁻⁵ (SCHWARTZ *et al.*, 2007); 49,7%, e-value < 10⁻⁴ (RUIMING *et al.*, 2010); 54,6%, e-value < 10⁻¹⁰ (ALVARENGA *et al.*, 2012); 70,9%, e-value < 10⁻⁴ (MA *et al.*, 2009) e 78,01% e-value < 10⁻⁶ (OLIVEIRA *et al.*, 2018). Contudo, as diferenças nas porcentagens de anotação dos *contigs* devem ser interpretadas com cautela pois todos os trabalhos, exceto o último, anotaram transcriptomas por meio de ESTs (*Expressed Sequence Tags*); cada trabalho adotou um limite de *e-value* próprio e não se sabe qual a similaridade, identidade e cobertura dos transcritos anotados. Além disso, o trabalho de Oliveira e colaboradores (2018) anotou um total de 4.120 *contigs*, pois sua montagem possuía inicialmente 5.282 *contigs*.

O fato de 26,65% dos *contigs* não apresentarem *match* pode ser interpretado de duas maneiras: parte deles representa um potencial científico grande, pois indicam ser moléculas existentes que ainda não foram descobertas e depositadas nos bancos de dados públicos, podendo ser específicas a *Tityus serrulatus* ou não; a outra fração é de fato erro de montagem. Os programas de montagem disponibilizam inúmeros parâmetros que o algoritmo leva em consideração ao agrupar as *reads* em *contigs*. O mesmo conjunto de *reads* pode originar montagens bastante diversas, apesar de parcialmente sobrepostas. De forma que uma parcela de *contigs* sempre conterà algum erro como, por exemplo, a ausência de ORFs (*Open Reading Frames*). Dos 7.006 *contigs no matches*, apenas 21 não apresentam nenhuma ORF, os outros 6.985 *contigs* apresentaram 59.923 ORFs com no mínimo 20 aminoácidos. Dentre elas 988 ORFs

contém peptídeos sinal canônicos, indicando secreção pela glândula (Figura 27). Outros erros de montagem se fazem presentes quando, durante a montagem, ocorre a fragmentação de proteínas, inviabilizando seu encontro pelas ferramentas bioinformáticas.

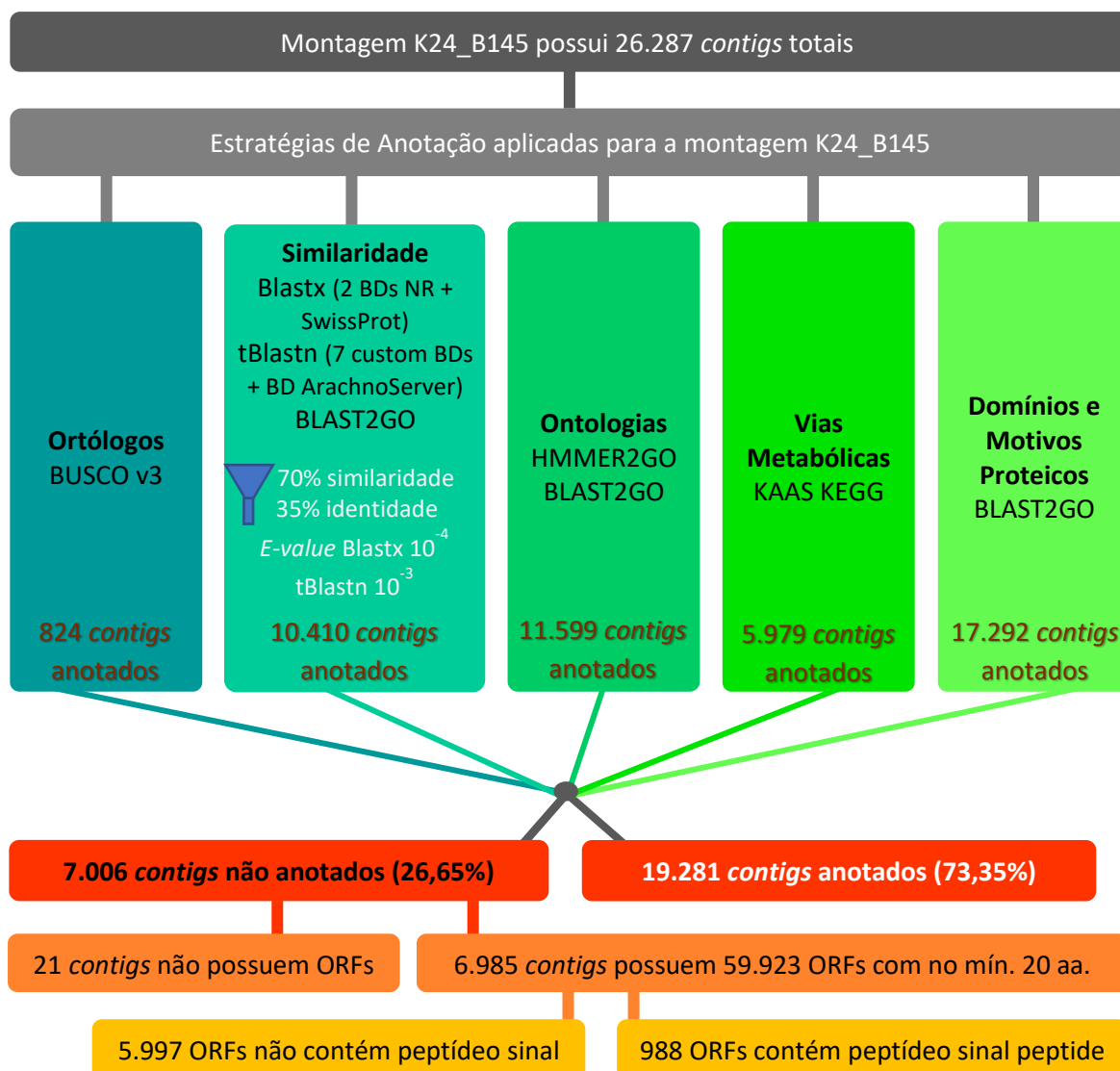


Figura 27 - Fluxo da análise dos *contigs* da montagem K24_B145 para anotação. Os números de *contigs* anotados por cada estratégia estão em marrom, dentro de seus respectivos quadros. Os 10.410 *contigs* anotados por similaridade são a soma dos 8.118 *contigs* dos alinhamentos locais realizados por blastx e tblastn mais 2.292 *contigs* acrescentados pelo BLAST2GO na sua busca por similaridade. Os 11.599 *contigs* anotados por busca por ontologias são a soma dos 3.361 *contigs* anotados via HMMER2GO mais os 8.238 *contigs* anotados via BLAST2GO. *Contigs* que são anotados por uma estratégia também podem ser anotados por outra estratégia, por isso, após a retirada das redundâncias, o número total de *contigs* anotados é de apenas 19.281.

6.8 Junção dos resultados das diferentes estratégias de anotação

As diversas estratégias de anotação se mostraram complementares tanto na abrangência de *contigs* passíveis de anotação, quanto na informação que cada uma delas traz. A realização da estratégia de anotação por alinhamento local utilizando blastx contra o banco de dados NR poderia ser alvo de críticas, pois o BLAST2GO também realiza esta tarefa. No entanto, além de obtermos resultados de alinhamento mais detalhados, esta estratégia

retornou 3.480 *contigs* para o saldo final de *contigs* com *match*. O blastx contra o banco SwissProt recuperou 20 *contigs*. Ao utilizar tblastn e o banco de dados público “ArachnoServer” um *contig* foi recuperado para a anotação final. Já a estratégia por tblastn com a utilização dos bancos de dados customizados recuperaram os seguintes números de *contigs*: “Scorpion” mais 4.042 *contigs* foram incluídos na categoria “*match*”; “Tityus serrulatus” acrescentou mais quatro *contigs*; “Toxin” mais 430 *contigs*; “Arachnida Gland” mais 135 *contigs* e “Arachnida Muscle” acrescentou mais seis *contigs* na contabilização final de anotados. Os demais bancos personalizados não acrescentaram outros *contigs* além dos já contabilizados (Figura 27 e Tabela 6). A ferramenta BUSCO acrescentou 83 *contigs* à anotação, KEGG mais 1.210 *contigs*, HMMER2GO acrescentou mais 1.840⁶ *contigs* à anotação e BLAS2GO acrescentou mais 8.030 *contigs* (Figura 27 e Tabela 6).

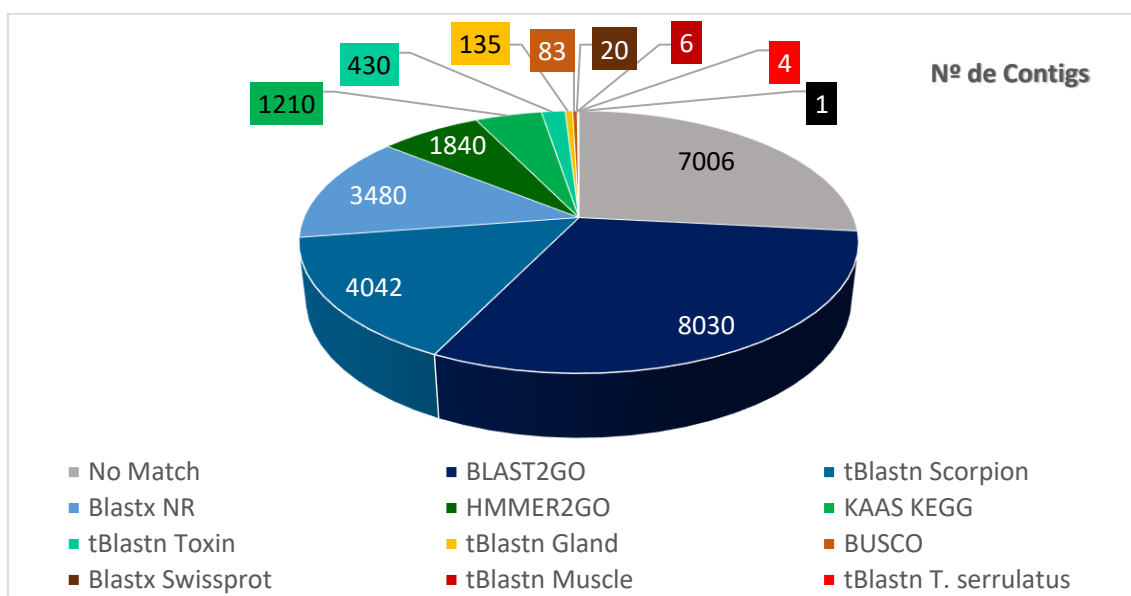


Figura 28 - Visualização da anotação por *contigs*. Número de *contigs* que foram anotados por estratégia empregada.

⁶ A diferença no número de *contigs* anotados por HMMER2GO em relação ao valor que está descrito na página 69, item 6.4, se deve ao fato de qual programa é avaliado primeiro, se o BLAST2GO ou HMMER2GO. Na página 69 eu considerei primeiro os *contigs* do BLAST2GO para depois retirar as redundâncias dos *contigs* anotados pelo HMMER2GO, apenas com efeito comparativo, por isso o número é inferior, apenas 1.049 *contigs*.

Tabela 6 - Resultados das estratégias de anotação.

| | |
|---|---------------|
| BLAST2GO | Nº |
| <i>Nº de contigs</i> anotados por blastx vs. NR | 14.269 |
| <i>Nº de contigs</i> anotados por GO | 11.599 |
| <i>Nº de contigs</i> anotados por InterPro | 17.292 |
| TOTAL de <i>contigs</i> anotados | 18.120 |
| Acrescentou quantos <i>contigs</i> à anotação | 8.030 |
| HMMER2GO | |
| <i>Nº de contigs</i> anotados por GO | 6.816 |
| Acrescentou quantos <i>contigs</i> à anotação | 1.840 |
| BUSCO | |
| Nº de genes ortólogos recuperados | 871 |
| Nº de <i>contigs</i> anotados por ortologia | 824 |
| Acrescentou quantos <i>contigs</i> à anotação | 83 |
| KEGG | |
| Nº de <i>contigs</i> anotados por vias metabólicas | 5.979 |
| Acrescentou quantos <i>contigs</i> à anotação | 1.210 |
| BLASTx vs. NR | |
| Nº de proteínas recuperadas | 7.620 |
| <i>Nº de contigs</i> com <i>match</i> (70% sim.) | 3.480 |
| Acrescentou quantos <i>contigs</i> à anotação | 3.480 |
| BLASTx vs. SwissProt | |
| Nº de proteínas recuperadas | 3.261 |
| Nº de proteínas acrescentadas | 3.185 |
| <i>Nº de contigs</i> com <i>match</i> (70% sim.) | 1.333 |
| Acrescentou quantos <i>contigs</i> à anotação | 20 |
| tBLASTn vs. banco personalizado <i>Scorpion</i> | |
| Nº de proteínas recuperadas | 13.819 |
| Nº de proteínas acrescentadas | 13.554 |
| <i>Nº de contigs</i> com <i>match</i> (70% sim.) | 7.202 |
| Acrescentou quantos <i>contigs</i> à anotação | 4.042 |
| tBLASTn vs. banco personalizado <i>Tityus Serrulatus</i> | |
| Nº de proteínas recuperadas | 210 |
| Nº de proteínas acrescentadas | 15 |
| <i>Nº de contigs</i> com <i>match</i> (70% sim.) | 113 |
| Acrescentou quantos <i>contigs</i> à anotação | 4 |
| tBLASTn vs. banco personalizado <i>Toxin</i> | |
| Nº de proteínas recuperadas | 6.607 |
| Nº de proteínas acrescentadas | 4.978 |
| <i>Nº de contigs</i> com <i>match</i> (70% sim.) | 2.322 |
| Acrescentou quantos <i>contigs</i> à anotação | 430 |
| tBLASTn vs. banco de toxinas <i>ArachnoServer</i> | |
| Nº de proteínas recuperadas | 8 |
| Nº de proteínas acrescentadas | 3 |
| <i>Nº de contigs</i> com <i>match</i> (70% sim.) | 8 |
| Acrescentou quantos <i>contigs</i> à anotação | 1 |

| tBLASTn vs. banco personalizado <i>Arachnida Gland</i> | |
|--|---------------|
| Nº de proteínas recuperadas | 6.863 |
| Nº de proteínas acrescentadas | 5.737 |
| Nº de <i>contigs</i> com <i>match</i> (70% sim.) | 2.696 |
| Acrescentou quantos <i>contigs</i> à anotação | 135 |
| tBLASTn vs. banco personalizado <i>Epithelium glandular</i> | |
| Nº de proteínas recuperadas | 4 |
| Nº de proteínas acrescentadas | 3 |
| Nº de <i>contigs</i> com <i>match</i> (70% sim.) | 2 |
| Acrescentou quantos <i>contigs</i> à anotação | 0 |
| tBLASTn vs. banco personalizado <i>Arachnida Muscle</i> | |
| Nº de proteínas recuperadas | 303 |
| Nº de proteínas acrescentadas | 207 |
| Nº de <i>contigs</i> com <i>match</i> (70% sim.) | 69 |
| Acrescentou quantos <i>contigs</i> à anotação | 6 |
| tBLASTn vs. banco personalizado <i>Outros Animais</i> | |
| Nº de proteínas recuperadas | 13 |
| Nº de proteínas acrescentadas | 0 |
| Nº de <i>contigs</i> com <i>match</i> (70% sim.) | 13 |
| Acrescentou quantos <i>contigs</i> à anotação | 0 |
| Total de <i>contigs</i> com <i>match</i> | 19.281 |
| Total de proteínas 70% similares às ORFs da montagem K24_B145 | 35.302 |

Pela tabela 6 é possível notar que o uso do banco de dados “*Epithelium Glandular*” não acrescenta *contigs* novos ao grupo “*match*”, mas acrescenta proteínas a serem investigadas sob sua existência na montagem. Também é perceptível como a utilização das diversas estratégias de anotação se complementam aumentando o número de *contigs* anotados.

Os bancos de dados customizados foram montados com base em proteínas existentes nos bancos NR e SwissProt. Então, como conseguimos novos *hits* depois de já ter realizado o alinhamento local contra esses bancos públicos? Com o crescimento exponencial do conteúdo dos bancos de dados os programas BLAST perdem um pouco sua sensibilidade (PROSDOCIMI, 2007). Estas ferramentas foram formuladas para lidar heurísticamente com os dados processados, de modo a aumentar a velocidade da busca por similaridade (ALTSCHUL *et al.*, 1997)

6.9 Breve comparação entre os estudos transcriptômicos de *Tityus serrulatus*

Diferenças na condição fisiológica das células, cujos mRNAs serão extraídos, permitem a obtenção de diferentes transcriptomas. Os estudos de Alvarenga e colaboradores (2012) e Oliveira e colaboradores (2018) extraíram mRNAs em um único período, 48 horas após a

drenagem das glândulas de peçonha. Aqui, os mRNAs foram extraídos 0, 2, 4 e 8 dias após a drenagem das glândulas de peçonha. Apesar de os indivíduos da população de *T. serrulatus* serem oriundos de partenogênese elas não são clones, assim, o número de indivíduos também afeta a quantidade de isoformas proteicas presentes na peçonha e por consequência afetam as montagens transcriptômicas (KALAPOTHAKIS&CHÁVEZ-OLÓRTEGUI, 1997). Oliveira *et al.* (2018) extraíram mRNA de 15 télsons, Alvarenga *et al.* extraíram de 60 télsons e Martins (2016) extraiu de 40 télsons, utilizados no presente estudo (Tabela 7).

Outro componente que influencia nos resultados adquiridos por um estudo transcriptômico é o modo de sequenciamento do cDNA. Em 2012 Alvarenga *et al.* sequenciaram pelo método de Sanger (ABI 3130 Genetic Analyzer), que possui um rendimento menor frente às técnicas de sequenciamento massivo de DNA (NGS). Oliveira *et al.* (2018) realizaram pirosequenciamento, por 200 ciclos, num GS Junior 454 Sequencing System (Roche Diagnostics). Nosso trabalho lida com mRNAs sequenciados com a estratégia *pair-end* para sequenciamento em MiSeq (Illumina). No pirosequenciamento as *reads* são mais longas (± 500 pb), facilitando as montagens dos *contigs* (KIM, 2013). O MiSeq, no presente estudo, gerou *reads* entre 70 a 287 pb, mas seu rendimento, em geral, é muito maior (16.7 Gb/run) em comparação com o sequenciamento 454 (0,04 Gb/run) (LOMAN *et al.*, 2012; KIM, 2013) (Tabela 7).

As análises bioinformáticas para anotação dos transcriptomas, que cada trabalho adotou, abordaram aspectos relativamente diferentes e, por vezes, parâmetros diferentes, o que prejudicou a comparação direta dos resultados obtidos por cada um. Alvarenga *et al.* (2012) produziram 643 *contigs* e anotaram pelo uso de blastx 54,6% deles, utilizando como corte de qualidade o *e-value* 1×10^{-10} . Oliveira *et al.* (2018) produziram 5.282 *contigs* e anotaram 78% do transcriptoma, também pelo uso de blastx, utilizando *e-value* menor que 1×10^{-6} como corte de qualidade. O presente trabalho produziu 26.287 *contigs* e anotou 73,35% pelo uso de diversas estratégias de anotação. Os alinhamentos locais realizados por blastx tiveram seu *e-value* limitado a valores $\leq 3 \times 10^{-4}$, mas outros parâmetros importantes como valores de cobertura e similaridade foram levados em conta (Tabela 7).

Acreditamos que essas diferenças foram responsáveis por permitir que o nosso trabalho revelasse outras características do transcriptoma de *Tityus serrulatus* não encontradas pelos dois estudos transcriptômicos anteriores. Também é possível que, diante de uma cuidadosa anotação manual, o presente trabalho revele outras moléculas que não foram relatadas nem por Oliveira *et al.* (2018) nem por Alvarenga *et al.* (2012).

Tabela 7 - Comparação entre os estudos transcriptômicos já realizados para *Tityus serrulatus*.

| | Alvarenga <i>et al.</i> , 2012 | Oliveira <i>et al.</i> , 2018 | Presente estudo |
|---|---------------------------------------|--|---|
| Extração de RNA | 48h após drenagem das glândulas | 48h após drenagem das glândulas | 0, 2, 4, 8 dias após a drenagem das glândulas |
| Nº de indivíduos | 60 télsons | 15 télsons | 40 télsons |
| Sequenciamento | Sanger ABI 3130 – Genetic Analyzer | [NGS] GS Junior 454 Sequencing System – Roche Diagnostics 200 ciclos <i>Reads</i> longas ± 500pb 0,04 – 0,7 Gb/run | [NGS] MiSeq – Illumina 600 ciclos <i>paired-end</i> <i>Reads</i> menores 70-287pb 1 – 16,7 Gb/run |
| Montagem | 1.629 ESTs 643 <i>uniques</i> | 5.282 <i>contigs</i> | 26.287 <i>contigs</i> |
| Anotou | 54,6% | 78% | 73,35% |
| E-value Blastx | ≤ 1x10 ⁻¹⁰ | ≤ 1x10 ⁻⁶ | ≤ 3x10 ⁻⁴ |
| Análises bioinformáticas de anotação | Blastx, BLAST2GO e KEGG | Blastx e BLAST2GO | Blastx, tblastn com uso de BD Customizados, BLAST2GO, HMMER2GO, KEGG, Genes Ortólogos |

7. CONCLUSÃO

O presente trabalho permitiu observar que a utilização de diferentes estratégias de anotação, em estudos transcriptômicos, se complementam aumentando o número de *contigs* anotados. Ao todo, 73,35% da montagem K24_B145 (19.281 *contigs*) foi anotada por similaridade com alguma proteína já existente, ou por possuir domínios proteicos já descritos, ou por parecer com genes ortólogos, ou ainda, por ter sido associada a termos do GO e/ou se fazer presente em alguma via metabólica conhecida. Os trabalhos de anotação de transcriptomas costumam utilizar apenas uma ou duas estratégias de anotação, este trabalho comprova a importância de se utilizar as várias ferramentas de anotação disponíveis para uma anotação mais completa.

As estratégias de busca por similaridade adicionaram milhares de proteínas, descritas pela primeira vez para o transcriptoma de *Tityus serrulatus*. Elas devem ser investigadas sobre a sua real existência, principalmente as que se supõe potencial biotecnológico.

Os programas que relacionam as ORFs dos *contigs* a termos do GO também se mostraram complementares, indicando que se há o real interesse em obter o máximo de informações relacionadas às ontologias, deve-se empregar mais de um programa para tal tarefa.

O uso de bancos de dados menores e customizados se mostrou de extrema importância na tarefa de anotação do transcriptoma. Duzentos e cinquenta e nove *contigs* só foram anotados por causa do emprego dessa técnica e todas as famílias proteicas de componentes de peçonha já descritos para *Tityus serrulatus* foram encontradas.

Até onde se sabe, o uso de tblastn como ferramenta de anotação de transcriptomas associado a bancos de dados customizados é uma metodologia nova e pelo desempenho no presente trabalho é bem recomendada. De fato, a opção por tblastn acarretou na contabilização de falsos positivos proteicos, mas se o objetivo é obter o máximo de informação possível sobre uma determinada sequência e posteriormente submetê-la a um processo de curadoria cuidadoso, a opção é bastante válida.

O presente estudo transcriptômico revelou outras características do transcriptoma de *Tityus serrulatus* que ainda não haviam sido encontradas pelos dois estudos anteriores. Por isso, é possível que, diante de uma cuidadosa anotação manual, os dados aqui produzidos revelem outras moléculas não relatadas por Oliveira *et al.* (2018) e por Alvarenga *et al.* (2012). Várias proteínas putativas, provenientes desses estudos transcriptômicos, também foram encontradas no presente trabalho, reforçando a importância de serem investigadas por métodos experimentais, principalmente as de atividade tóxica.

Os 7.006 *contigs* (26,65% da montagem K24_B145) não anotados, apresentam 988 ORFs contendo peptídeos sinais canônicos, sugerindo sua secreção pela glândula. Este dado reforça aquilo que já é consenso entre os toxinologistas: que ainda faltam componentes da peçonha de *Tityus serrulatus* a serem descritos.

8. PERSPECTIVAS

Muitas análises podem derivar dos resultados obtidos neste trabalho ou complementá-los. Entre as possibilidades destaco:

- Realizar as mesmas estratégias de anotação para outras montagens deste transcriptoma.

- Realizar um alinhamento local dos contigs das montagens contra todas as ESTs depositadas no NCBI para *Tityus serrulatus*.

- Selecionar determinadas famílias de proteínas e realizar uma curadoria para uma anotação detalhada, classificando-as de forma sistemática e retirando os falsos positivos. Além disso, determinar a porcentagem de ocorrência delas na montagem e discriminar seus níveis de expressão.

- Realizar buscas específicas por RNAs não traduzidos. Apesar da biblioteca de cDNA, utilizada para realizar este estudo transcriptômico, ter sido enriquecida para mRNA, há alguma possibilidade de termos sequenciado RNA não traduzido.

- Realizar ensaios biológicos para confirmação da presença e atuação de determinados transcritos.

REFERÊNCIAS

- ABDEL-RAHMAN, M. A.; HARRISON, P. L. and STRONG, P. N. (2015) Snapshots of scorpion venomics. *Journal of Arid Environments*. Elsevier Ltd 112: 170–176.
- ALMAAYTAHA, A.; ALBALASB, Q. Scorpion venom peptides with no disulfide bridges: A review. *Peptides* 51 (2014) 35– 45
- ALMEIDA, F. M.; PIMENTA, A. M.; DE FIGUEIREDO, S. G.; SANTORO, M. M.; MARTIN-EAUCLAIRE, M. F.; DINIZ, C. R.; DE LIMA, M. E. Enzymes with gelatinolytic activity can be found in *Tityus bahiensis* and *Tityus serrulatus* venoms. *Toxicon*. 2002. Jul;40(7):1041-5.
- ALTAFAJ, X.; FRANCE, J.; ALMASSY, J.; JONA, I.; ROSSI, D.; SORRENTINO, V.; MABROUK, K.; DE WAARD, M.; RONJAT, M. Maurocalcine interacts with the cardiac ryanodine receptor without inducing channel modification. *Biochem. J.*, 2007, 406(2), 309-315.
- ALTSCHUL, S. F.; MADDEN T. L.; SCHAFFER A. A.; ZHANG J.; ZHANG Z.; MILLER W., AND LIPMAN D. J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", *Nucleic Acids Res.* 25:3389-3402.
- ALVARENGA É. R.; MENDES T. M.; MAGALHÃES B. F.; SIQUEIRA F. F.; DANTAS A. E.; BARROCA T. M.; HORTA C. C. AND KALAPOTHAKIS E. (2012) Transcriptome analysis of the *Tityus serrulatus* scorpion venom gland. *Open Journal of Genetics* 2(December): 210–220.
- ALVES, R.S.; XIMENES, R.M.; JORGE, A.R.; NASCIMENTO, N.R.; MARTINS, R.D.; RABELLO, M.M.; HERNANDES, M.Z.; TOYAMA, D.O.; TOYAMA, M.H.; MARTINS, A.M.; HAVT, A.; MONTEIRO, H.S. Isolation, homology modeling and renal effects of a C-type natriuretic peptide from the venom of the Brazilian yellow scorpion (*Tityus serrulatus*). *Toxicon*, 2013, 74, 19-26.
- AMNH – AMERICAN MUSEUM OF NATURAL HISTORY. <http://scorpion.amnh.org/index.html> (acessado em 04 de agosto de 2018)
- BARBAZUK, W. B., EMRICH, S. J., CHEN, H. D., LI, L. & SCHNABLE, P. S. SNP discovery via 454 transcriptome sequencing. *Plant J.* 51, 910–918 (2007).
- BECCALONI, J. Arachnids. Natural History Museum, London. 2009. ISBN 978-0-656-09220-7. P. 6, 239, 249, 266
- BERTAZZI, D. T.; ASSIS-PANDOCHIA, A. I.; AZZOLINIA, A. E. C. S.; TALHAFERROA, V. L.; LAZZARINI, M.; ARANTES, E. C. Effect of *Tityus serrulatus* scorpion venom and its major toxin, TsTX-I, on the complement system in vivo. *Toxicon* 41 (2003) 501–508
- BERTAZZI, D. T. Isolamento e caracterização bioquímica de componentes do veneno de *Tityus serrulatus* com ação sobre o sistema complemento. Universidade de São Paulo. Faculdade de Ciências Farmacêuticas de Ribeirão Preto (2007) Tese doutorado.
- BINNECK, E. As ômicas: integrando a bioinformação. *Biotec Ci & Des.* 2004, 32: 28-37. http://www.biotecnologia.com.br/revista/bio32/omicas_32.pdf
- BIRRELL, G. W.; EARL, S.; MASCI, P. P.; DE JERSEY, J.; WALLIS, T. P.; GORMAN, J. J.; LAVIN, M. F. Molecular diversity in venom from the Australian Brown snake, *Pseudonaja textilis*. *Mol. Cell. Proteomics* 2006, 5 (2), 379–89.

BORGES, A ; SOUZA, L. ; ESPINOZA, J. ; Santos RG ; KALAPOTHAKIS, E. ; VALADARES, D. ; CHÁVEZ-OLÓRTEGUI, C. . Characterization of Tityus scorpion venoms using synaptosome binding assays and reactivity towards Venezuelan and Brazilian Antivenoms?. *Toxicon* (Oxford), v. 51, p. 66-79, **2008**.

BORTOLUZZI, L. R.; QUEROL, M. V. M.; QUEROL, E. Notas sobre a ocorrência de *Tityus serrulatus* Lutz & Mello, 1922 (Scorpiones, Buthidae) no oeste do Rio Grande do Sul, Brasil. *Biota Neo trop.* **2007**. vol.7, n.3, pp. 357-359.

BOUTSELAKIS, H.; DIMITROPOULOS, D.; FILLON, J.; GOLOVIN, A.; HENRICK, K.; HUSSAIN, A.; IONIDES, J.; JOHN, M.; KELLER, P. A.; KRISSEL, E.; MCNEIL, P.; NAIM, A.; NEWMAN, R.; OLDFIELD, T.; PINEDA, J.; RACHEDI, A.; COPELAND, J.; SITNOV, A.; SOBHANY, S.; SUAREZ-URUENA, A.; SWAMINATHAN, J.; TAGARI, M.; TATE, J.; TROMM, S.; VELANKAR, S. AND VRANKEN, W. E-MSD: the European Bioinformatics Institute Macromolecular Structure Database. *Nucleic Acids Research*, **2003**, Vol. 31, No. 1

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Manual de controle e manejo de escorpiões. Brasília, **2009**.

BRAZIL, T. K.; PORTO, T. J. **2010**. Os escorpiões. Editora Universidade Federal da Bahia, Salvador. ISBN 978-85-232-0729-8.

BROWNELL, P. & G.A. POLIS. **1990**. Scorpions Biology and Research. Oxford University Press. 544p.

CANDIDO, D. M. *et al.* Uma nova espécie de *Tityus* C. L. Koch, 1836 (Scorpiones, Buthidae) do Estado da Bahia, Brasil. *Biota Neotropica*, v. 5, n. 1, p. 193-200, **2005**.

CARMO, A. O.; OLIVEIRA-MENDES, B. B. R.; HORTA, C. C. R.; MAGALHÃES, B. F.; DANTAS, A. E.; CHAVES, L. M.; CHÁVEZ-OLÓRTEGUI, C. and KALAPOTHAKIS, E. (**2014**) Molecular and functional characterization of metallo-serrulases, new metalloproteases from the *Tityus serrulatus* venom gland. *Toxicon* 90: 45–55.

CARMO, A. O.; CHATZAKI, M.; HORTA, C. C. R.; MAGALHÃES, B. F.; OLIVEIRA-MENDES, B. B. R.; CHÁVEZ-OLÓRTEGUI, C. and KALAPOTHAKIS, E. (**2015**) Evolution of alternative methodologies of scorpion antivenoms production. *Toxicon* 97: 64–74.

CARMO, A. O. Busca por novos antígenos para a produção de soro antiescorpiônico: Construção de antígenos quiméricos e caracterização de proteases da peçonha de *T. serrulatus*. Tese de doutorado. **2015**
(http://www.pggenetica.icb.ufmg.br/tese_defesas_detalhes.php?aluno=152&tipo=D)

CARMO A.O., MARTINS A.P.V., OLIVEIRA-MENDES B.B.R., HORTA C.C.R., DANTAS A.E., KALAPOTHAKIS E.; Proteases from *Tityus serrulatus* venom gland: venom proteases and peptide maturation. **2017** (<https://uniprot.org/uniprot/A0A1S5QN69.txt?version=5>) (A0A1S5QN68; A0A1S5QN66; A0A1S5QN64; A0A1S5QN71; A0A1S5QN83; A0A1S5QNT6)

CASEWELL, N.R.; WUSTER, W.; VONK, F.J.; HARRISON, R.A.; FRY, B.G. Complex cocktails: the evolutionary novelty of venoms. *Trends Ecol. Evol.*, **2013**, 28(4), 219-229.

CHÁVEZ-OLÓRTEGUI, C.; KALAPOTHAKIS, E.; FERREIRA, A.M.; FERREIRA, A.P.; DINIZ, C. R. Neutralizing capacity of antibodies elicited by a non-toxic protein purified from the venom of the scorpion *Tityus serrulatus*. *Toxicon*. **1997**. Feb;35(2):213-21.

CHAVEZ-OLORTEGUI, C.; MOLINA, F.; GRANIER, C. Molecular basis for the cross-reactivity of antibodies elicited by a natural anatoxin with α - and β -toxins from the venom of *Tityus serrulatus* scorpion. *Molecular Immunology* 38 (**2001**) 867–876

CHIPPAUX, J. P. and GOYFFON, M. (**1997**) Venoms, antivenoms and immunotherapy. *Toxicon*, 823–846.

CHIPPAUX, J. P. & GOYFFON, M. Epidemiology of scorpionism: a global appraisal. *Acta. Trop.* **2008**. Aug;107(2):71-9.

CHUGUNOV, A. O.; KOROMYSLOVA, A. D.; BERKUT, A. A.; PEIGNEUR, S.; TYTGAT, J.; POLYANSKY, A. A.; EFREMOV, R. G. Modular organization of α -toxins from scorpion venom mirrors domain structure of their targets, sodium channels. *J. Biol. Chem.*, **2013**, 288 (26), 19014-19027.

CLOUDSLEY-THOMPSON, J. L. Scorpions in mythology, folklore, and history. In: POLIS, G. A. (Org.). *The Biology of Scorpions*. Stanford: Stanford University Press, **1990**. p. 462-485.

COELHO, V. A.; CREMONEZ, C. M.; ANJOLETTE, F. A. P.; AGUIAR, J. F.; VARANDA, W. A.; ARANTES, E.C. Functional and structural study comparing the C-terminal amidated β -neurotoxin Ts1 with its isoform Ts1-G isolated from *Tityus serrulatus* venom. *Toxicon* 83 (**2014**) 15–21.

COETZEE, W. A.; AMARILLO, Y.; CHIU, J.; CHOW, A.; LAU, D.; MCCORMACK, T.; MORENO, H.; NADAL, M. S.; OZAITA, A.; POUNTNEY, D.; SAGANICH, M.; MIERA, E. V. AND RUDY, B. Molecular Diversity of K⁺ Channels, *ANNALS NEW YORK ACADEMY OF SCIENCES* **1999**.

COLOGNA, C. T.; MARCUSSI, S.; GIGLIO, J. R.; SOARES, A. M.; ARANTES, E. C. *Tityus serrulatus* scorpion venom and toxins: an overview. *Protein Pept.Lett.***2009**. 16(8):920-32.

COLOGNA, C. T.; PEIGNEUR, S.; RUSTIGUEL, J. K.; NONATO, M. C.; TYTGAT, J. and ARANTES, E. C. (**2012**) Investigation of the relationship between the structure and function of Ts2, a neurotoxin from *Tityus serrulatus* venom. *FEBS Journal* 279: 1495–1504.

CONESA, A.; GÖTZ, S.; GARCÍA-GÓMEZ, J. M.; TEROL, J.; TALÓN, M.; ROBLES, M. (**2005**). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, (18):3674-6.

CONESA, A.; MADRIGAL, P.; TARAZONA, S.; GOMEZ-CABRERO, D.; CERVERA, A.; MCPHERSON, A.; SZCZEŚNIAK, M. W.; GAFFNEY, D. J.; ELO, L. L.; ZHANG, X. AND MORTAZAV, A. A survey of best practices for RNA-seq data analysis. *Genome Biology* (**2016**) 17:13

CRUZ, E. F. S. **1994a**. *Biologia dos Escorpiões*. In Barraviera, B. *Venenos Animais - Uma Visão Integrada*. Rio de Janeiro: Editora de Publicações Científicas Ltda. p. 135-150.

CUELLAR, O. (**1994**) Biogeography of parthenogenetic animals. *Compte rendu des Séances de la Société de Biogeographie* 70: 1–13.

CUPO, P.; AZEVEDO-MARQUES, M. M.; HERING, S. E. Escorpionismo. In: CARDOSO, J. L. C. *et al.* Animais peçonhentos no Brasil: biologia, clínica e terapêutica dos acidentes. 2. ed. São Paulo: Sarvier; Fapesp, **2009**. p. 214-224.

DATASUS (**2017**) www.tabnet.datasus.gov.br (acessado em março de 2017)

DE LA VEGA, R.C. RODRÍGUEZ; POSSANI, L.D. Current views on scorpion toxins specific for K⁺-channels. *Toxicon*. **2004**. May 18;43. p865-75.

DE LA VEGA, R.C. RODRÍGUEZ; POSSANI, L.D. Overview of scorpion toxins specific for Na⁺ channels and related peptides: biodiversity, structure-function relationships and evolution. *Toxicon*. **2005**. Dec 15;46(8):831-44.

DE LA VEGA, R. R.C.; SCHWARTZ, E.F.; POSSANI, L.D. Mining on scorpion venom biodiversity. *Toxicon*. **2010**. Dec 15;56(7):1155-61.

DE SOUZA, C. A. R.; CANDIDO, D. M.; LUCAS, S. M. & BRESCOVIT, A. D. On the *Tityus stigmurus* complex (Scorpiones, Buthidae). *Zootaxa* 1987: 1–38 (**2009**)

DENG, Y.; GU, J.; YAN, Z.; WANG, M.; MA, C.; ZHANG, J.; JIANG, G.; GE, M.; XU, S.; XU, Z.; XIAO, L. De novo transcriptomic analysis of the venomous glands from the scorpion *Heterometrus spinifer* revealed unique and extremely high diversity of the venom peptides. *Toxicon* 143 (**2018**) 1e19

DUEÑAS-CUELLAR, R. A.; KUSHMERICK, C.; NAVES, L. A.; BATISTA, I. F. C.; GUERRERO-VARGAS, J. A.; PIRES, J.; FONTES WOR and CASTRO, M. S. (**2015**) Cm38 : A New Antimicrobial Peptide Active Against *Klebsiella pneumoniae* is Homologous to Cn11. *Protein & Peptide Letters* 22: 164–172.

ESPINDOLA, F. S.; CALÁBRIA, L. K; REZENDE, A. A. A; PEREIRA, B. B.; SANTANA, F. A.; AMARAL, I. M. R.; LOBATO, J.; FRANÇA, J. L; MARIO, J. L.; FIGUEIREDO, L. B.; SANTOS-LOPES, L. P.; GOUVEIA, N. M.; NASCIMENTO, R.; TEIXEIRA, R. R.; REIS, T. A.; ARAÚJO, T. G. Recursos De Bioinformática Aplicados Às Ciências Ômicas Como Genômica, Transcriptômica, Proteômica, Interatômica E Metabolômica. *Biosci. J., Uberlândia*, v. 26, n. 3, p. 463-477, May/June **2010**

ESPINO-SOLIS, G. P.; RIAÑO-UMBARILA, L.; BECERRIL, B.; POSSANI, L. D. Antidotes against venomous animals: state of the art and prospectives. *J Proteomics*. **2009**. Mar 6;72(2):183-99.

FERREIRA, L. A.; ALVES, E. W.; HENRIQUES, O. B. Peptide T, a novel bradykinin potentiador isolated from *Tityus serrulatus* scorpion venom. *Toxicon*, **1993**, 31(8), 941-947.

FERREIRA, M. G.; DUARTE, C. G.; OLIVEIRA, M. S.; CASTRO, K. L. P.; TEIXEIRA, M. S.; REIS, L. P. G.; ZAMBRANO, J. A.; KALAPOTHAKIS, E.; MICHEL, A. F. R. M.; SOTO-BLANCO, B.; Chávez-Olórtegui, C.; MELO, M. M. Toxicity of crude and detoxified *Tityus serrulatus* venom in anti-venom-producing sheep. *Journal of Veterinary Science (Seoul. Print)*, v. 17, p. 467, **2016**

FET, V. *et al.* Catalog of the scorpions of the world (1758-1998). New York: New York Entomological Society, **2000**. 690 p.

FLETCHER, P. L. JR.; FLETCHER, M. D.; WENINGER, K.; ANDERSON, T. E.; MARTIN, B. M. Vesicle-associated membrane protein (VAMP) cleavage by a new metalloprotease from the Brazilian scorpion *Tityus serrulatus*. *J Biol. Chem.* **2010**. Mar 5;285(10):7405-16.

FUNED (2017) - http://www.funed.mg.gov.br/wp-content/uploads/2016/02/BULFJ-0049-REV02_Profissional-de-Sa%C3%BAde.pdf (Acessado em setembro de 2017)

FUZITA, F. J.; PINKSE, M. W. H.; PATANE, J. S. L.; JULIANO, M. A.; VERHAERT, P. D. E. M.; LOPES, A. R. (2015) Biochemical, Transcriptomic and Proteomic Analyses of Digestion in the Scorpion *Tityus serrulatus*: Insights into Function and Evolution of Digestion in an Ancient Arthropod. *PLoS ONE* 10(4): e0123841. doi:10.1371/journal.pone.0123841

GERTZ, E. M.; YU, YI-KUO; AGARWALA, R.; SCHÄFFER, A. A and ALTSCHUL, S. F. Composition-based statistics and translated nucleotide searches: Improving the TBLASTN module of BLAST. *BMC Biology* 2006, 4:41 doi:10.1186/1741-7007-4-41

GISH, W.; STATES, D. J. Identification of protein coding regions by database similarity search. *Nat Genet* 1993, 3:266-272.

GLESENER, R. R.; TILMAN, D. (1978) Sexuality and the components of environmental uncertainty: clues from geographic parthenogenesis in terrestrial animals. *American Naturalist* 112(986): 659–673. doi: 10.1086/283308

GUATIMOSIM, S. C. F.; KALAPOTHAKIS, E.; DINIZ, C. R. and CHÁVEZ-OLÓRTEGUI, C. (2000) Induction of neutralizing antibodies against *Tityus serrulatus* toxins by immunization with a recombinant nontoxic protein. *Toxicon* 38: 113–121.

GUIA DE VIGILÂNCIA EPIDEMIOLÓGICA, 7ª ed. Brasília - Distrito Federal, 2009. Página 793 (http://bvsmis.saude.gov.br/bvs/publicacoes/guia_vigilancia_epidemiologica_7ed.pdf) acessado em junho de 2017

GUO, X.; MA, C.; DU, Q.; WEI, R.; WANG, L.; ZHOU, M.; CHEN, T.; SHAW, C. Two peptides, TsAP-1 and TsAP-2, from the venom of the Brazilian yellow scorpion, *Tityus serrulatus*: Evaluation of their antimicrobial and anticancer activities. *Biochimie* 95 (2013) 1784e1794

GEBEL, H. M.; FINKE, J. H.; ELGERT, K. D.; CAMBELL, B. J.; BARRETT, J. T. Inactivation of complement by *Loxosceles reclusa* spider venom. *Am J Trop Med Hyg*, Baltimore, v. 28, p. 756-62, 1979.

GUREVITZ, M.; KARBAT, I.; COHEN, L.; ILAN, N.; KAHN, R.; TURKOV, M., STANKIEWICZ, M.; STÜHMER, W.; DONG, K. and GORDON, D. (2007) The insecticidal potential of scorpion β -toxins. *Toxicon* 49(4): 473–489.

GWEE, M. C. E.; GOPALAKRISHNAKONE, P.; CHEAH, L. S.; WONG, P. T. H.; GONG, J. P.; KINI, R. M. Studies on venom from the Black scorpion *Heterometrus longimanus* and some other scorpion species. *J. Toxicol., Toxin Rev.* 1996.15, 37– 57.

HORTA, C. C. R.; MAGALHÃES, B. D. E. F.; OLIVEIRA-MENDES B. B. R.; CARMO, A. O. DO; DUARTE, C. G.; FELICORI, L. F.; MACHADO-DE-ÁVILA, R. A.; CHÁVEZ-OLÓRTEGUI, C. and KALAPOTHAKIS, E. (2014) Molecular, Immunological, and Biological Characterization of *Tityus serrulatus* Venom Hyaluronidase: New Insights into Its Role in Envenomation. *PLoS Neglected Tropical Diseases* 8(2).

HORTA, C. C. R. & MAGALHAES, B. F.; CHATZAKI, M.; REZENDE, B. A.; DUARTE, C. G.; FELICORI, L. F.; OLIVEIRA-MENDES, B. B. R.; CARMO, A. O.; CHAVEZ-OLORTEGUI, C. and KALAPOTHAKIS,

E. Cardiovascular-Active Venom Toxins: An Overview. *Current Medicinal Chemistry*, **2016**, *23*, 603-622

HOSHINO, K.; MOURA, A. T. V.; DE PAULA, H. M. G. Selection of environmental temperature by the yellow scorpion *Tityus serrulatus* Lutz & Mello, 1922 (Scorpiones, Buthidae). *Journal of Venom Animals toxins including Tropical diseases*. Vol 12, n1, p.59-66, **2006**.

HUANG, Y.; NIU, B.; GAO, Y.; FU, L.; LI, W. CD-HIT Suite: a web server for clustering and 1244 comparing biological sequences. *Bioinformatics*. 26:680, **2010**.

KREIL, G. (1995) Hyaluronidases - a group of neglected enzymes. *Protein Sci* 4(9): 1666–1669.

ICMBio. Ministério do Meio Ambiente. Instituto Chico Mendes de Conservação da Biodiversidade. Sumário Executivo – Livro Vermelho da Fauna Brasileira Ameaçada de Extinção. Brasília, **2016**

INCEOGLU, B.; LANGO, J.; JING, J.; CHEN, L.; DOYMAZ, F.; PESSAH, I. N.; HAMMOCK, B. D. One scorpion, two venoms: pre venom of *Parabuthus transvaalicus* acts as an alternative type of venom with distinct mechanism of action. *Proc. Natl. Acad. Sci. U.S.A.*, **2003**, *100*, 922-927.

JEYAPRAKASH, J.; HOY, M. A. First divergence time estimate of spiders, scorpions, mites and ticks (subphylum: Chelicerata) inferred from mitochondrial phylogeny. *Exp Appl Acarol* **2009**; *47*:1–18.

KALAPOTHAKIS, E. & CHÁVEZ-OLÓRTEGUI, C. Venom variability among several *Tityus serrulatus* specimens. *Toxicon*. **1997** Oct;35(10):1523-9.

KELLEY, L. A. *et al.* The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10**, 845-858 (**2015**)

KIM, MIN SOO. (2013) <https://pt.slideshare.net/mkim8/a-comparison-of-ngs-platforms> (Acessado em julho, 2018)

KNAPP, O.; MCARTHUR, J. R.; ADAMS, D. J. Conotoxins targeting neuronal voltage-gated sodium channel subtypes: potential analgesics? *Toxins*, **2012**, *4*(11), 1236-1260.

LOMAN, N. J.; MISRA, R. V.; DALLMAN, T. J.; CONSTANTINIDOU, C.; GHARBIA, S. E.; WAIN, J.; PALLEN, M. J. Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology*, **2012**, *30*, 434. Macmillan Publishers Limited. (<http://dx.doi.org/10.1038/nbt.2198>)

LOURENÇO, W.R.; CUELLAR, O. (1995) Scorpions, scorpionism, life history strategies and parthenogenesis. *Journal of Venomous Animals and Toxins* 1: 50–64. doi: 10.1590/S0104-79301995000200002 (disponível on-line: http://www.scielo.br/scielo.php?pid=S0104-79301995000200002&script=sci_arttext)

LOURENÇO, W. R.; CLOUDSLEY-THOMPSON, J. L. (1999) Discovery of a sexual population of *Tityus serrulatus*, one of the morphs within the complex *Tityus stigmurus* (Scorpiones, Buthidae). *Journal of Arachnology* 27(1): 154–158.

LOURENÇO, W. R. Scorpiones. In: ADIS, J. (Org.). *Amazonian arachnida and myriapoda*:

identification keys to all classes, orders, families, some genera and lists of known terrestrial species. Moscow: Pensoft Publishes, **2002a**. p. 399-438.

LOURENÇO, W. R. Reproduction in scorpions, with special reference to parthenogenesis. In: TOFT, S.; SCHARFF, N. (Org.). European Arachnology 2000. Aarhus: Aarhus University Press, **2002b**. p. 71-85.

LOURENÇO, W. R. Scorpions of Brazil. Paris: Les Editions de l'IF, **2002c**. 320p.

LOURENÇO, W.R. (**2006**) Nouvelle proposition de découpage sous-générique du genre *Tityus* C. L. Koch, 1836(Scorpiones, Buthidae). Boletín Sociedad Entomológica Aragonesa, 39, 55–67.

LOURENÇO, W. R. Parthenogenesis in scorpions: some history – new data. Journal of Venomous Animals and Toxins including Tropical Diseases, Botucatu, v. 14, n. 1, p.19-44, **2008**.

LOURENÇO, W. R.; EICKSTEDT, V. R. Escorpiões de Importância Médica. In: CARDOSO, J. L. C. *et al.* Animais peçonhentos no Brasil: biologia, clínica e terapêutica dos acidentes. São Paulo: Sarvier, **2009**. p. 198-213.

LOURENÇO, W. R. What do we know about some of the most conspicuous scorpion species of the genus *Tityus*? A historical approach. Journal of Venomous Animals and Toxins including Tropical Diseases (**2015**) 21:20 DOI 10.1186/s40409-015-0016-9

LYNCH, M. (**1984**) Destabilizing hybridization, general purpose genotypes and geographic parthenogenesis. Quarterly Review of Biology 59(3): 257–290. doi: 10.1086/413902

MA, Y.; RUIMING, Z.; HE, Y.; LI, S.; LIU, J.; WU, Y.; CAO, Z.; and WENXIN, LI. Transcriptome analysis of the venom gland of the scorpion *Scorpiops jendeki*: implication for the evolution of the scorpion venom arsenal. *BMC Genomics* **2009**, 10:290 doi:10.1186/1471-2164-10-290

MA, Y.; HE, Y.; ZHAO, R.; WU, Y.; LI, W.; CAO, Z. (**2011**). Extreme diversity of scorpion venom peptides and proteins revealed by transcriptomic analysis: Implication for proteome evolution of scorpion venom arsenal. *J Proteomics* 75(5):1563-76.

MAHER, C. A.; KUMAR-SINHA, C.; CAO, X.; KALYANA-SUNDARAM, S.; HAN, B.; JING, X.; *et al.* Transcriptome sequencing to detect gene fusions in cancer. *Nature* **2009**; 458:97-101.

MALLI, H.; KUHN-NENTWIG, L.; IMBODEN, H.; NENTWING, W. Effects of size, motility and paralysation time of prey on the quantity of venom injected by the hunting spider *Cupiennius salei*. *J. Exp. Biol.*, **1999**, 202, 2083-2089.

MARTINS A. P. V., Novas toxinas de canais para sódio de *Tityus serrulatus* (Lutz e Mello, 1922) através de RNA-Seq e o primeiro mitogenoma completo do gênero *Tityus*. Tese de doutorado. **2016** (http://www.pggenetica.icb.ufmg.br/tese_defesas_detalhes.php?aluno=283&tipo=D)

MARTINS, A. P. V.; CARMO, A. O. DO; MESQUITA, F. O.; PIMENTA, R. J. G.; CHAGAS, A. T. DE A.; KALAPOTHAKIS, E. Many unique characteristics revealed by the complete mitochondrial genome of the scorpion (Lutz e Mello 1922) (Chelicerata; Arachnida). *Mitochondrial DNA*, v. 1, p. 1-2, **2016**.

MARTENS-UZUNOVA; JALAVA, S. E.; DITS, N. F.; LEENDERS, G. V.; MULLER, S.; TRAPMAN,

J.; BANGMA, C. H.; LITMAN, T.; VISAKORPI, T. and JENSTER, G. Diagnostic and prognostic signatures from the small non-coding RNA transcriptome in prostate cancer. *Oncogene* (2012) 31, 978–991

MATTHIESEN, F. A. (1962) Parthenogenesis in Scorpions. *Evolution* 16(2): 255–256. doi: 10.2307/2406202

MATTHIESEN, F. A. On the sexual behaviour of some Brazilian scorpions. *Revista Brasileira de Pesquisas Médicas e Biológicas*, v. 1, p. 93-96, 1968.

MAURANO, H. R. Do escorpionismo. 1915. Tese de Doutorado - Faculdade de Medicina do Rio de Janeiro, Rio de Janeiro.

MEBS, D. *Venomous and Poisonous Animals*; Medpharm Scientific Publishers, CRC Press: Stuttgart, 2002.

MINEO, M. F.; FRANCO-ASSIS, G. A.; DEL-CLARO, K. Repertório comportamental do escorpião amarelo *Tityus serrulatus* Lutz & Mello 1922 (Scorpiones, Buthidae) em cativeiro. *Revista Brasileira de Zoociências*, Juiz de Fora, v.5, n.1, p. 23-31, 2003

MORIYA, Y.; ITOH, M.; OKUDA, S.; YOSHIZAWA, A.C.; KANEHISA, M. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, Volume 35, Issue suppl_2, Pages W182–W185, 2007.

NASCIMENTO, D. G.; RATES, B.; SANTOS, D. M.; VERANO-BRAGA, T.; BARBOSA-SILVA, A.; DUTRA, A. A.; BIONDI, I.; MARTIN-EAUCLAIRE, M. F.; DE LIMA, M. E.; PIMENTA, A. M. (2006). Moving pieces in a taxonomic puzzle: venom 2D-LC/MS and data clustering analyses to infer phylogenetic relationships in some scorpions from the Buthidae family (Scorpiones). *Toxicon*, 47:628-639.

NISANI, Z. and HAYES, W.K. (2011) Defensive stinging by *Parabuthus transvaalicus* scorpions: risk assessment and venom metering. *Anim. Behav.* 81, 627–633

OLIVEIRA, U. C.; NISHIYAMA, M. Y. JR.; DOS SANTOS, M. B. V.; SANTOS-DA-SILVA, A. D. P.; CHALKIDIS, H. D. M.; SOUZA-IMBERG A. *et al.* (2018) Proteomic endorsed transcriptomic profiles of venom glands from *Tityus obscurus* and *T. serrulatus* scorpions. *PLoS ONE* 13(3): e0193739. <https://doi.org/10.1371/journal.pone.0193739>

OLIVEIRA, U. C.; CANDIDO, D. M.; DORCE, V. A. C. and JUNQUEIRA-DE-AZEVEDO, I. D. L. M. (2015) The transcriptome recipe for the venom cocktail of *Tityus bahiensis* scorpion. *Toxicon* 95: 52–61.

OLIVEIRA, U. C.; NISHIYAMA, M. Y. JR.; CANDIDO, D. M.; YAMANOUYE, N.; DORCE, V. A. and JUNQUEIRA-DE-AZEVEDO, I.L. Comparative transcriptomes and proteomes of Brazilian medically important scorpions *Tityus serrulatus* from Parana and Sao Paulo. 2017 – Unpublished

ORTIZ, E.; RENDÓN-ANAYA, M.; REGO, S. C.; SCHWARTZ, E. F.; POSSANI, L. D. 2014 Jun. Antirease-like Zn-metalloproteases are ubiquitous in the venom of different scorpion genera. *Biochim. Biophys. Acta* 1840 (6), 1738e1746.

PESSINI, A. C.; TAKAO, T. T.; CAVALHEIRO, E. C.; VICHNEWSKI, W.; SAMPAIO, S. V.; GIGLIO, J. R.; ARANTES, E. C. A hyaluronidase from *Tityus serrulatus* scorpion venom: isolation, characterization and inhibition by flavonoids. *Toxicon*. **2001**. Oct;39(10):1495-504.

PETERSEN, T. N.; BRUNAK, S.; VON HEIJNE, G.; NIELSEN, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods*, 8:785-786.

PINEDA, S. S.; CHAUMEIL, P. A.; KUNERT, A.; KAAS, Q.; KING, G. F. ArchnoServer 3.0: an online resource for automated discovery, analysis and annotation of spider toxins. *Bioinformatics*, Volume 34, Issue 6, 15 March 2018, Pages 1074-1076, **2018**.

PIMENTA, A. M.; STÖCKLIN, R.; FAVREAU, P.; BOUGIS, P. E. and MARTIN-EAUCCLAIRE, M. F. (2001) Moving pieces in a proteomic puzzle: mass fingerprinting of toxic fractions from the venom of *Tityus serrulatus* (Scorpiones, Buthidae). *Rapid communications in mass spectrometry : RCM* 15(17): 1562–72.

PIMENTA, A. M. C.; DE MARCO ALMEIDA, F.; DE LIMA, M. E.; MARTIN-EAUCCLAIRE, M. F. and BOUGIS, P. E. (2003) Individual variability in *Tityus serrulatus* (Scorpiones, Buthidae) venom elicited by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid communications in mass spectrometry : RCM* 17: 413–418.

POLIS, G. A. Ecology. In: POLIS, G. A. (Org.). *The Biology of Scorpions*. Stanford: Stanford University Press, **1990**. 249p, 262p.

POLIS, G. A. & SISSOM W. D. Life History. In: POLIS, G. A. (Org.). *The Biology of Scorpions*. Stanford: Stanford University Press, **1990**. 170p.

PRENDINI L. W.; WHEELER, C. Scorpion higher phylogeny and classification, taxonomic anarchy, and standards for peer review in online publishing. *Cladistics*, v. 30, p. 446-494, **2005**.

PROSDOCIMI, F.; CERQUEIRA, G. C.; BINNECK, E.; SILVA, A. F.; REIS, A. N.; JUNQUEIRA, A. C. M.; SANTOS, A. C. F.; NHANI-JÚNIOR, A.; WUST, C. I.; CAMARGO-FILHO, F.; KESSEDJIAN, J. L.; PETRETSKI, J. H.; CAMARGO, L. P.; FERREIRA, R. G. M.; LIMA, R. P.; PEREIRA, R. M.; JARDIM, S.; SAMPAIO, V. S. and FOLGUERAS-FLATSCHART, A. V. *Bioinformática: manual do usuário. Biotec. Ci. Des.* 29: 18-31, **2002**.

PROSDOCIMI, F. Curso On-line Introdução à Bioinformática. **2007**
http://www2.bioqmed.ufrj.br/prosdocimi/FProsdocimi07_CursoBioinfo.pdf

PUCCA, M. B.; CERNI, F. A.; PINHEIRO JUNIOR, E. L.; BORDON, K. DE C. F.; AMORIM, F. G.; CORDEIRO, F. A.; LONGHIM, H. T.; CREMONEZ, C. M.; OLIVEIRA, G. H.; ARANTES, E. C. *Tityus serrulatus* venom: A lethal cocktail. *Toxicon*, v.108, p. 272-284, **2015**

RATES, B.; FERRAZ, K. K.; BORGES, M. H.; RICHARDSON, M.; DE LIMA, M. E.; PIMENTA, A. M. *Tityus serrulatus* venom peptidomics: assessing venom peptide diversity. *Toxicon*. **2008** Oct;52(5):611-8.

RENDÓN-ANAYA, M.; DELAYE, L.; POSSANI, L. D.; HERRERA-ESTRELLA, A. Global transcriptome analysis of the scorpion *Centruroides noxius*: new toxin families and evolutionary insights from an ancestral scorpion species. *PLoS One*. **2012**. 7(8):e43331.

ROOT, T. M. Neurobiology. In: POLIS, G. A. (Org.). *The biology of scorpions*. Stanford: Stanford

University Press, **1990**. p. 341-413.

ROYCE, T. E.; ROZOWSKY, J. S. & GERSTEIN, M. B. Toward a universal microarray: prediction of gene expression through nearest-neighbor probe sequence identification. *Nucleic Acids Res.* **35**, e99 (**2007**).

RUIMING, Z.; YIBAO, M.; YAWEN, H.; ZHIYONG, D.; YINGLIANG, W.; ZHIJIAN, C.; WENXIN, L. (**2010**). Comparative venom gland transcriptome analysis of the scorpion *Lychas mucronatus* reveals intraspecific toxic gene diversity and new venomous components. *BMC Genomics*, **11**: 452.

RUPPERT, E. E.; BARNES, R. D. Zoologia do Invertebrados. 7.ed. São Paulo: Rocca, **2005**. 657p.

SALEM, M. L.; SHOUKRY, N. M.; TELEB, W. K.; ABDEL-DAIM, M. M. and ABDEL-RAHMAN, M. A. (**2016**) In vitro and in vivo antitumor effects of the Egyptian scorpion *Androctonus amoreuxi* venom in an Ehrlich ascites tumor model. *SpringerPlus*. Springer International Publishing 5(570): 1–12.

SANTOS, M. D. S.; PORTO, T. J.; LIRA-DA-SILVA, R. M.; BRAZIL, T. K. Description of the male of *Tityus kuryi* Lourenço, 1997 and notes about males of *Tityus stigmurus* (Thorell, 1877) and *Tityus serrulatus* Lutz & Mello, 1922 (Scorpiones, Buthidae) *ZooKeys* 435: 49–61 (**2014**) doi: 10.3897/zookeys.435.6694

SIMÃO, F. A.; WATERHOUSE, R. M.; IOANNIDIS, P.; KRIVENTSEVA, E. V.; ZDOBNOV, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, Volume 31, Issue 19, 1 October **2015**, Pages 3210–3212, 2015.

SCHLIESKY, S.; GOWIK, U.; WEBER A. P. M. and BRÄUTIGAM, A. RNA-seq assembly – are we there yet? *Frontiers in Plant Science*. 25 September **2012** doi: 10.3389/fpls.2012.00220

SISSOM W. D. Systematics, biogeography and paleontology. In: POLIS, G. A. (Org.). *The biology of scorpions*. Stanford: Stanford University Press, **1990**. p. 137; 64; 88.

SOUZA, C.A.R.; CANDIDO, D.M.; LUCAS, S.M.; BRESCOVIT, A.D. (2009) On the *Tityus stigmurus* complex (Scorpiones, Buthidae). *Zootaxa* **1987**: 1–38.

STEIN, L., 2001. Genome annotation: from sequence to biology. *Nature Reviews* 2: 493-505

STEWART, L. M.; HIRST, M.; FERBER, M. L.; MERRYWEATHER, A. T.; CAYLEY, P. J.; POSEE, R. D. Construction of an improved baculovirus insecticide containing an insect-specific toxin gene. *Nature*, v. 352, p. 85-88, **1991**.

SCHWARTZ, E.; DIEGO-GARCIA, E.; DE LA VEJA, R. C. R. and POSSANI, L. D. Transcriptome analysis of the venom gland of the Mexican scorpion *Hadrurus gertschi* (Arachnida: Scorpiones). *BMC Genomics* **2007**, 8:119 doi:10.1186/1471-2164-8-119

VENÂNCIO, E. J.; PORTARO, F. C.; KUNIYOSHI, A. K.; CARVALHO, D. C.; PIDDE-QUEIROZ, G.; TAMBOURGI, D. V. Enzymatic properties of venoms from Brazilian scorpions of *Tityus* genus and the neutralisation potential of therapeutical antivenoms. *Toxicon*. **2013**. Mar 15. pii: S0041-0101.

VERANO-BRAGA, T.; FIGUEIREDO-REZENDE, F.; MELO, M. N.; LAUTNER, R. Q.; GOMES, E. R.; *et al.* (2010) Structure-function studies of *Tityus serrulatus* Hypotensin-I (TsHpt-I): A new agonist of B(2) kinin receptor. *Toxicon* 56(7): 1162–71.

VERANO-BRAGA, T.; DUTRA, A. A.; LEÓN, I. R.; MELO-BRAGA, M. N.; ROEPSTORFF, P.; PIMENTA, A. M.; KJELDSEN, F. Moving pieces in a venomous puzzle: unveiling post-translationally modified toxins from *Tityus serrulatus*. *J Proteome Res.* 2013 Jul 5;12(7):3460-70.

VILELA, E. Soroterapia anti-escorpiônica. *Brazil-médico.* 1917. 31,393.

WANG, E. T.; SANDBERG, R.; LUO, S.; KHREBTUKOVA, I.; ZHANG, L.; MAYR, C.; KINGSMORE, S. F.; SCHROTH, G. P. & BURGE, C. B. Alternative isoform regulation in human tissue transcriptomes. *Nature* Vol 456 | 27 November 2008 | doi:10.1038/nature07509

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for 1385 transcriptomics. *Nature reviews. Genetics*, v. 10, n. 1, p. 57–63, 2009.

WANKHEDE, R. A. (2004) Extraction, Isolation, Identification and Distribution of Soluble Fluorescent Compounds from the Cuticle of Scorpion (*Hadrurus arizonensis*) - Marshall University – Thesis
(<https://pdfs.semanticscholar.org/32af/89471c4a5c9f3c1676d32d8ee312c008c421.pdf>)

XU, X. B.; DUAN, Z. G.; DI, Z. Y.; HE, Y. W.; LI, J. L. *et al.* (2014) Proteomic analysis of the venom from the scorpion *Mesobuthus martensii*. *Journal of Proteomics* 106: 162±180.
<https://doi.org/10.1016/j.jprot.2014.04.032> PMID: 24780724

YANG, I. S.; KIM, S. Analysis of Whole Transcriptome Sequencing Data: Workflow and Software. *Genomics Inform.* Genomics & Informatics eISSN 2234-0742 *Genomics Inform* 2015;13(4):119-125

YE, J.; FANG, L. *et al.* *Nucleic. Acids Res.*, 2006, 34 (Web service issue), 293-297.

ANEXO A – Lista dos organismos utilizados para montagem do banco de dados KAAS-KEGG

| Organism | Database number of sequences |
|---|------------------------------|
| hsa <i>Homo sapiens</i> (human) | 20.228 |
| oaa <i>Ornithorhynchus anatinus</i> (platypus) | 19.839 |
| dme <i>Drosophila melanogaster</i> (fruit fly) | 13.929 |
| dpo <i>Drosophila pseudoobscura pseudoobscura</i> | 14.574 |
| dan <i>Drosophila ananassae</i> | 14.365 |
| dsi <i>Drosophila simulans</i> | 14.192 |
| mde <i>Musca domestica</i> | 14.896 |
| aag <i>Aedes aegypti</i> (yellow fever mosquito) | 14.626 |
| cqu <i>Culex quinquefasciatus</i> (southern house mosquito) | 18.896 |
| ame <i>Apis mellifera</i> (honey bee) | 10.734 |
| bim <i>Bombus impatiens</i> (common eastern bumble bee) | 10.384 |
| aec <i>Acromyrmex echinator</i> (Panamanian leafcutter ant) | 11.219 |
| cfo <i>Camponotus floridanus</i> (Florida carpenter ant) | 12.512 |
| nvi <i>Nasonia vitripennis</i> (jewel wasp) | 13.358 |
| tca <i>Tribolium castaneum</i> (red flour beetle) | 12.881 |
| dpa <i>Dendroctonus ponderosae</i> (mountain pine beetle) | 13.021 |
| nvl <i>Nicrophorus vespilloides</i> | 12.642 |
| dpl <i>Danaus plexippus</i> (monarch butterfly) | 15.245 |
| pxy <i>Plutella xylostella</i> (diamondback moth) | 18.119 |
| api <i>Acyrtosiphon pisum</i> (pea aphid) | 18.619 |
| zne <i>Zootermopsis nevadensis</i> | 12.381 |
| fcd <i>Folsomia candida</i> | 20.251 |
| isc <i>Ixodes scapularis</i> (black-legged tick) | 20.467 |
| nve <i>Nematostella vectensis</i> (sea anemone) | 24.773 |
| epa <i>Exaiptasia pallida</i> | 22.199 |
| hmg <i>Hydra vulgaris</i> | 20.057 |
| Total | 414.407 |

APÊNDICE A – Lista da anotação realizada por KEGG

| KEGG-BRITE Categories | Nº of contigs |
|--|---------------|
| 09120 Genetic Information Processing | 4.935 |
| 09121 Transcription | 815 |
| 03000 Transcription factors [BR:ko03000] | 267 |
| 03041 Spliceosome [BR:ko03041] | 232 |
| 03021 Transcription machinery [BR:ko03021] | 178 |
| 03040 Spliceosome [PATH:ko03040] | 93 |
| 03022 Basal transcription factors [PATH:ko03022] | 28 |
| 03020 RNA polymerase [PATH:ko03020] | 17 |
| 09122 Translation | 1340 |
| 03019 Messenger RNA Biogenesis [BR:ko03019] | 281 |
| 03029 Mitochondrial biogenesis [BR:ko03029] | 215 |
| 03009 Ribosome biogenesis [BR:ko03009] | 191 |
| 03016 Transfer RNA biogenesis [BR:ko03016] | 128 |
| 03013 RNA transport [PATH:ko03013] | 115 |
| 03011 Ribosome [BR:ko03011] | 96 |
| 03010 Ribosome [PATH:ko03010] | 74 |
| 03012 Translation factors [BR:ko03012] | 70 |
| 03008 Ribosome biogenesis in eukaryotes [PATH:ko03008] | 66 |
| 03015 mRNA surveillance pathway [PATH:ko03015] | 62 |
| 00970 Aminoacyl-tRNA biosynthesis [PATH:ko00970] | 42 |
| 09123 Folding, sorting and degradation | 1652 |
| 04131 Membrane trafficking [BR:ko04131] | 751 |
| 04121 Ubiquitin system [BR:ko04121] | 388 |
| 03110 Chaperones and folding catalysts [BR:ko03110] | 124 |
| 04141 Protein processing in endoplasmic reticulum [PATH:ko04141] | 114 |
| 04120 Ubiquitin mediated proteolysis [PATH:ko04120] | 84 |
| 03018 RNA degradation [PATH:ko03018] | 59 |
| 03051 Proteasome [BR:ko03051] | 51 |
| 03050 Proteasome [PATH:ko03050] | 35 |
| 04130 SNARE interactions in vesicular transport [PATH:ko04130] | 21 |
| 03060 Protein export [PATH:ko03060] | 19 |
| 04122 Sulfur relay system [PATH:ko04122] | 6 |
| 09124 Replication and repair | 1128 |
| 03036 Chromosome and associated proteins [BR:ko03036] | 670 |
| 03400 DNA repair and recombination proteins [BR:ko03400] | 192 |
| 03032 DNA replication proteins [BR:ko03032] | 102 |
| 03030 DNA replication [PATH:ko03030] | 30 |
| 03460 Fanconi anemia pathway [PATH:ko03460] | 30 |
| 03420 Nucleotide excision repair [PATH:ko03420] | 28 |
| 03440 Homologous recombination [PATH:ko03440] | 26 |
| 03410 Base excision repair [PATH:ko03410] | 22 |
| 03430 Mismatch repair [PATH:ko03430] | 18 |
| 03450 Non-homologous end-joining [PATH:ko03450] | 10 |
| 09125 RNA family | 0 |
| 03100 Non-coding RNAs [BR:ko03100] | 0 |
| 09160 Human Diseases | 3.551 |
| 09161 Cancers | 1313 |
| 05200 Pathways in cancer [PATH:ko05200] | 196 |
| 05205 Proteoglycans in cancer [PATH:ko05205] | 103 |
| 05203 Viral carcinogenesis [PATH:ko05203] | 101 |

| | |
|---|------------|
| 05225 Hepatocellular carcinoma [PATH:ko05225] | 87 |
| 05206 MicroRNAs in cancer [PATH:ko05206] | 86 |
| 05202 Transcriptional misregulation in cancers [PATH:ko05202] | 69 |
| 05226 Gastric cancer [PATH:ko05226] | 63 |
| 05224 Breast cancer [PATH:ko05224] | 62 |
| 05231 Choline metabolism in cancer [PATH:ko05231] | 57 |
| 05210 Colorectal cancer [PATH:ko05210] | 47 |
| 05211 Renal cell carcinoma [PATH:ko05211] | 43 |
| 05215 Prostate cancer [PATH:ko05215] | 39 |
| 05220 Chronic myeloid leukemia [PATH:ko05220] | 38 |
| 05222 Small cell lung cancer [PATH:ko05222] | 37 |
| 05230 Central carbon metabolism in cancer [PATH:ko05230] | 34 |
| 05214 Glioma [PATH:ko05214] | 33 |
| 05212 Pancreatic cancer [PATH:ko05212] | 32 |
| 05213 Endometrial cancer [PATH:ko05213] | 31 |
| 05223 Non-small cell lung cancer [PATH:ko05223] | 30 |
| 05217 Basal cell carcinoma [PATH:ko05217] | 29 |
| 05221 Acute myeloid leukemia [PATH:ko05221] | 24 |
| 05218 Melanoma [PATH:ko05218] | 24 |
| 05219 Bladder cancer [PATH:ko05219] | 19 |
| 05216 Thyroid cancer [PATH:ko05216] | 15 |
| 05204 Chemical carcinogenesis [PATH:ko05204] | 14 |
| 09162 Immune diseases | 50 |
| 05322 Systemic lupus erythematosus [PATH:ko05322] | 23 |
| 05323 Rheumatoid arthritis [PATH:ko05323] | 17 |
| 05340 Primary immunodeficiency [PATH:ko05340] | 6 |
| 05321 Inflammatory bowel disease (IBD) [PATH:ko05321] | 4 |
| 05310 Asthma [PATH:ko05310] | 0 |
| 05320 Autoimmune thyroid disease [PATH:ko05320] | 0 |
| 05330 Allograft rejection [PATH:ko05330] | 0 |
| 05332 Graft-versus-host disease [PATH:ko05332] | 0 |
| 09163 Neurodegenerative diseases | 286 |
| 05016 Huntington's disease [PATH:ko05016] | 98 |
| 05010 Alzheimer's disease [PATH:ko05010] | 86 |
| 05012 Parkinson's disease [PATH:ko05012] | 66 |
| 05014 Amyotrophic lateral sclerosis (ALS) [PATH:ko05014] | 22 |
| 05020 Prion diseases [PATH:ko05020] | 14 |
| 09164 Substance dependence | 142 |
| 05034 Alcoholism [PATH:ko05034] | 58 |
| 05032 Morphine addiction [PATH:ko05032] | 33 |
| 05031 Amphetamine addiction [PATH:ko05031] | 29 |
| 05030 Cocaine addiction [PATH:ko05030] | 19 |
| 05033 Nicotine addiction [PATH:ko05033] | 3 |
| 09165 Cardiovascular diseases | 157 |
| 05418 Fluid shear stress and atherosclerosis [PATH:ko05418] | 68 |
| 05414 Dilated cardiomyopathy (DCM) [PATH:ko05414] | 29 |
| 05410 Hypertrophic cardiomyopathy (HCM) [PATH:ko05410] | 25 |
| 05412 Arrhythmogenic right ventricular cardiomyopathy (ARVC) [PATH:ko05412] | 19 |
| 05416 Viral myocarditis [PATH:ko05416] | 16 |
| 09166 Endocrine and metabolic diseases | 281 |
| 04934 Cushing's syndrome [PATH:ko04934] | 78 |
| 04932 Non-alcoholic fatty liver disease (NAFLD) [PATH:ko04932] | 65 |
| 04931 Insulin resistance [PATH:ko04931] | 59 |

| | |
|--|--------------|
| 04933 AGE-RAGE signaling pathway in diabetic complications [PATH:ko04933] | 43 |
| 04930 Type II diabetes mellitus [PATH:ko04930] | 24 |
| 04940 Type I diabetes mellitus [PATH:ko04940] | 6 |
| 04950 Maturity onset diabetes of the young [PATH:ko04950] | 6 |
| 09167 Infectious diseases | 1178 |
| 05165 Human papillomavirus infection [PATH:ko05165] | 133 |
| 05166 HTLV-I infection [PATH:ko05166] | 124 |
| 05169 Epstein-Barr virus infection [PATH:ko05169] | 98 |
| 05168 Herpes simplex infection [PATH:ko05168] | 72 |
| 05167 Kaposi's sarcoma-associated herpesvirus infection [PATH:ko05167] | 63 |
| 05164 Influenza A [PATH:ko05164] | 60 |
| 05152 Tuberculosis [PATH:ko05152] | 57 |
| 05161 Hepatitis B [PATH:ko05161] | 53 |
| 05100 Bacterial invasion of epithelial cells [PATH:ko05100] | 48 |
| 05142 Chagas disease (American trypanosomiasis) [PATH:ko05142] | 44 |
| 05162 Measles [PATH:ko05162] | 41 |
| 05160 Hepatitis C [PATH:ko05160] | 41 |
| 05120 Epithelial cell signaling in Helicobacter pylori infection [PATH:ko05120] | 40 |
| 05132 Salmonella infection [PATH:ko05132] | 40 |
| 05145 Toxoplasmosis [PATH:ko05145] | 38 |
| 05110 Vibrio cholerae infection [PATH:ko05110] | 35 |
| 05131 Shigellosis [PATH:ko05131] | 35 |
| 05146 Amoebiasis [PATH:ko05146] | 35 |
| 05134 Legionellosis [PATH:ko05134] | 28 |
| 05133 Pertussis [PATH:ko05133] | 27 |
| 05130 Pathogenic Escherichia coli infection [PATH:ko05130] | 26 |
| 05140 Leishmaniasis [PATH:ko05140] | 18 |
| 05143 African trypanosomiasis [PATH:ko05143] | 10 |
| 05144 Malaria [PATH:ko05144] | 9 |
| 05150 Staphylococcus aureus infection [PATH:ko05150] | 3 |
| 03200 Viral proteins [BR:ko03200] | 0 |
| 09168 Drug resistance | 144 |
| 01522 Endocrine resistance [PATH:ko01522] | 51 |
| 01521 EGFR tyrosine kinase inhibitor resistance [PATH:ko01521] | 38 |
| 01524 Platinum drug resistance [PATH:ko01524] | 38 |
| 01523 Antifolate resistance [PATH:ko01523] | 17 |
| 01501 beta-Lactam resistance [PATH:ko01501] | 0 |
| 01502 Vancomycin resistance [PATH:ko01502] | 0 |
| 01503 Cationic antimicrobial peptide (CAMP) resistance [PATH:ko01503] | 0 |
| 01504 Antimicrobial resistance genes [BR:ko01504] | 0 |
| 09150 Organismal Systems | 3.085 |
| 09151 Immune system | 642 |
| 04062 Chemokine signaling pathway [PATH:ko04062] | 66 |
| 04611 Platelet activation [PATH:ko04611] | 58 |
| 04621 NOD-like receptor signaling pathway [PATH:ko04621] | 58 |
| 04625 C-type lectin receptor signaling pathway [PATH:ko04625] | 43 |
| 04660 T cell receptor signaling pathway [PATH:ko04660] | 41 |
| 04666 Fc gamma R-mediated phagocytosis [PATH:ko04666] | 41 |
| 04670 Leukocyte transendothelial migration [PATH:ko04670] | 39 |

| | |
|--|------------|
| 04624 Toll and Imd signaling pathway [PATH:ko04624] | 36 |
| 04620 Toll-like receptor signaling pathway [PATH:ko04620] | 32 |
| 04662 B cell receptor signaling pathway [PATH:ko04662] | 31 |
| 04650 Natural killer cell mediated cytotoxicity [PATH:ko04650] | 29 |
| 04664 Fc epsilon RI signaling pathway [PATH:ko04664] | 29 |
| 04658 Th1 and Th2 cell differentiation [PATH:ko04658] | 25 |
| 04659 Th17 cell differentiation [PATH:ko04659] | 24 |
| 04623 Cytosolic DNA-sensing pathway [PATH:ko04623] | 21 |
| 04622 RIG-I-like receptor signaling pathway [PATH:ko04622] | 20 |
| 04657 IL-17 signaling pathway [PATH:ko04657] | 20 |
| 04612 Antigen processing and presentation [PATH:ko04612] | 17 |
| 04610 Complement and coagulation cascades [PATH:ko04610] | 6 |
| 04640 Hematopoietic cell lineage [PATH:ko04640] | 5 |
| | |
| 04672 Intestinal immune network for IgA production [PATH:ko04672] | 1 |
| 09152 Endocrine system | 918 |
| 04910 Insulin signaling pathway [PATH:ko04910] | 78 |
| 04919 Thyroid hormone signaling pathway [PATH:ko04919] | 72 |
| 04921 Oxytocin signaling pathway [PATH:ko04921] | 69 |
| 04915 Estrogen signaling pathway [PATH:ko04915] | 59 |
| 04926 Relaxin signaling pathway [PATH:ko04926] | 57 |
| 04916 Melanogenesis [PATH:ko04916] | 56 |
| 04912 GnRH signaling pathway [PATH:ko04912] | 52 |
| 04914 Progesterone-mediated oocyte maturation [PATH:ko04914] | 52 |
| 04922 Glucagon signaling pathway [PATH:ko04922] | 50 |
| 04925 Aldosterone synthesis and secretion [PATH:ko04925] | 50 |
| 04928 Parathyroid hormone synthesis, secretion and action [PATH:ko04928] | 48 |
| 04911 Insulin secretion [PATH:ko04911] | 39 |
| 04917 Prolactin signaling pathway [PATH:ko04917] | 34 |
| 04918 Thyroid hormone synthesis [PATH:ko04918] | 34 |
| 04924 Renin secretion [PATH:ko04924] | 34 |
| 04920 Adipocytokine signaling pathway [PATH:ko04920] | 28 |
| 04927 Cortisol synthesis and secretion [PATH:ko04927] | 28 |
| 04923 Regulation of lipolysis in adipocyte [PATH:ko04923] | 26 |
| 03320 PPAR signaling pathway [PATH:ko03320] | 23 |
| 04913 Ovarian Steroidogenesis [PATH:ko04913] | 18 |
| 04614 Renin-angiotensin system [PATH:ko04614] | 11 |
| 09153 Circulatory system | 143 |
| 04270 Vascular smooth muscle contraction [PATH:ko04270] | 65 |
| 04261 Adrenergic signaling in cardiomyocytes [PATH:ko04261] | 57 |
| 04260 Cardiac muscle contraction [PATH:ko04260] | 21 |
| 09154 Digestive system | 264 |
| 04972 Pancreatic secretion [PATH:ko04972] | 49 |
| 04970 Salivary secretion [PATH:ko04970] | 43 |
| 04971 Gastric acid secretion [PATH:ko04971] | 33 |
| 04976 Bile secretion [PATH:ko04976] | 30 |
| 04974 Protein digestion and absorption [PATH:ko04974] | 29 |
| 04973 Carbohydrate digestion and absorption [PATH:ko04973] | 19 |
| 04978 Mineral absorption [PATH:ko04978] | 19 |
| 04975 Fat digestion and absorption [PATH:ko04975] | 16 |
| 04979 Cholesterol metabolism [PATH:ko04979] | 15 |
| 04977 Vitamin digestion and absorption [PATH:ko04977] | 11 |
| 09155 Excretory system | 88 |

| | |
|--|--------------|
| 04961 Endocrine and other factor-regulated calcium reabsorption [PATH:ko04961] | 27 |
| 04962 Vasopressin-regulated water reabsorption [PATH:ko04962] | 23 |
| 04960 Aldosterone-regulated sodium reabsorption [PATH:ko04960] | 19 |
| 04964 Proximal tubule bicarbonate reclamation [PATH:ko04964] | 12 |
| 04966 Collecting duct acid secretion [PATH:ko04966] | 7 |
| 09156 Nervous system | 427 |
| 04722 Neurotrophin signaling pathway [PATH:ko04722] | 66 |
| 04728 Dopaminergic synapse [PATH:ko04728] | 55 |
| 04723 Retrograde endocannabinoid signaling [PATH:ko04723] | 53 |
| 04725 Cholinergic synapse [PATH:ko04725] | 45 |
| 04724 Glutamatergic synapse [PATH:ko04724] | 42 |
| 04726 Serotonergic synapse [PATH:ko04726] | 39 |
| 04727 GABAergic synapse [PATH:ko04727] | 35 |
| 04720 Long-term potentiation [PATH:ko04720] | 31 |
| 04721 Synaptic vesicle cycle [PATH:ko04721] | 31 |
| 04730 Long-term depression [PATH:ko04730] | 30 |
| 09157 Sensory system | 103 |
| 04750 Inflammatory mediator regulation of TRP channels [PATH:ko04750] | 45 |
| 04745 Phototransduction - fly [PATH:ko04745] | 23 |
| 04740 Olfactory transduction [PATH:ko04740] | 19 |
| 04742 Taste transduction [PATH:ko04742] | 11 |
| 04744 Phototransduction [PATH:ko04744] | 5 |
| 09158 Development | 156 |
| 04360 Axon guidance [PATH:ko04360] | 86 |
| 04380 Osteoclast differentiation [PATH:ko04380] | 39 |
| 04320 Dorso-ventral axis formation [PATH:ko04320] | 31 |
| 09149 Aging | 149 |
| 04211 Longevity regulating pathway - mammal [PATH:ko04211] | 53 |
| 04212 Longevity regulating pathway - worm [PATH:ko04212] | 51 |
| 04213 Longevity regulating pathway - multiple species [PATH:ko04213] | 45 |
| 09159 Environmental adaptation | 195 |
| 04714 Thermogenesis [PATH:ko04714] | 121 |
| 04713 Circadian entrainment [PATH:ko04713] | 42 |
| 04710 Circadian rhythm [PATH:ko04710] | 13 |
| 04626 Plant-pathogen interaction [PATH:ko04626] | 9 |
| 04711 Circadian rhythm - fly [PATH:ko04711] | 6 |
| 04712 Circadian rhythm - plant [PATH:ko04712] | 4 |
| 09100 Metabolism | 2.870 |
| 09101 Carbohydrate metabolism | 361 |
| 00520 Amino sugar and nucleotide sugar metabolism [PATH:ko00520] | 49 |
| 00562 Inositol phosphate metabolism [PATH:ko00562] | 42 |
| 00010 Glycolysis / Gluconeogenesis [PATH:ko00010] | 34 |
| 00630 Glyoxylate and dicarboxylate metabolism [PATH:ko00630] | 30 |
| 00020 Citrate cycle (TCA cycle) [PATH:ko00020] | 28 |
| 00620 Pyruvate metabolism [PATH:ko00620] | 28 |
| 00051 Fructose and mannose metabolism [PATH:ko00051] | 24 |
| 00640 Propanoate metabolism [PATH:ko00640] | 24 |
| 00052 Galactose metabolism [PATH:ko00052] | 23 |
| 00500 Starch and sucrose metabolism [PATH:ko00500] | 23 |

| | |
|--|------------|
| 00030 Pentose phosphate pathway [PATH:ko00030] | 20 |
| 00650 Butanoate metabolism [PATH:ko00650] | 15 |
| 00040 Pentose and glucuronate interconversions [PATH:ko00040] | 13 |
| 00053 Ascorbate and aldarate metabolism [PATH:ko00053] | 8 |
| 00660 C5-Branched dibasic acid metabolism [PATH:ko00660] | 0 |
| 09102 Energy metabolism | 120 |
| 00190 Oxidative phosphorylation [PATH:ko00190] | 59 |
| 00680 Methane metabolism [PATH:ko00680] | 18 |
| 00710 Carbon fixation in photosynthetic organisms [PATH:ko00710] | 15 |
| 00720 Carbon fixation pathways in prokaryotes [PATH:ko00720] | 14 |
| 00910 Nitrogen metabolism [PATH:ko00910] | 9 |
| 00920 Sulfur metabolism [PATH:ko00920] | 5 |
| 00195 Photosynthesis [PATH:ko00195] | 0 |
| 00196 Photosynthesis - antenna proteins [PATH:ko00196] | 0 |
| 00194 Photosynthesis proteins [BR:ko00194] | 0 |
| 09103 Lipid metabolism | 303 |
| 00564 Glycerophospholipid metabolism [PATH:ko00564] | 57 |
| 01004 Lipid biosynthesis proteins [BR:ko01004] | 43 |
| 00561 Glycerolipid metabolism [PATH:ko00561] | 37 |
| 00071 Fatty acid degradation [PATH:ko00071] | 27 |
| 00600 Sphingolipid metabolism [PATH:ko00600] | 24 |
| 00565 Ether lipid metabolism [PATH:ko00565] | 21 |
| 00590 Arachidonic acid metabolism [PATH:ko00590] | 19 |
| 00062 Fatty acid elongation [PATH:ko00062] | 17 |
| 01040 Biosynthesis of unsaturated fatty acids [PATH:ko01040] | 15 |
| 00061 Fatty acid biosynthesis [PATH:ko00061] | 8 |
| 00592 alpha-Linolenic acid metabolism [PATH:ko00592] | 8 |
| 00072 Synthesis and degradation of ketone bodies [PATH:ko00072] | 7 |
| 00140 Steroid hormone biosynthesis [PATH:ko00140] | 6 |
| 00591 Linoleic acid metabolism [PATH:ko00591] | 6 |
| 00120 Primary bile acid biosynthesis [PATH:ko00120] | 4 |
| 00100 Steroid biosynthesis [PATH:ko00100] | 3 |
| 00073 Cutin, suberine and wax biosynthesis [PATH:ko00073] | 1 |
| 00121 Secondary bile acid biosynthesis [PATH:ko00121] | 0 |
| 09104 Nucleotide metabolism | 178 |
| 00230 Purine metabolism [PATH:ko00230] | 111 |
| 00240 Pyrimidine metabolism [PATH:ko00240] | 67 |
| 09105 Amino acid metabolism | 337 |
| 01007 Amino acid related enzymes [BR:ko01007] | 53 |
| 00310 Lysine degradation [PATH:ko00310] | 50 |
| 00280 Valine, leucine and isoleucine degradation [PATH:ko00280] | 40 |
| 00260 Glycine, serine and threonine metabolism [PATH:ko00260] | 33 |
| 00270 Cysteine and methionine metabolism [PATH:ko00270] | 31 |
| 00330 Arginine and proline metabolism [PATH:ko00330] | 30 |
| 00250 Alanine, aspartate and glutamate metabolism [PATH:ko00250] | 24 |
| 00380 Tryptophan metabolism [PATH:ko00380] | 22 |
| 00350 Tyrosine metabolism [PATH:ko00350] | 16 |
| 00340 Histidine metabolism [PATH:ko00340] | 13 |
| 00220 Arginine biosynthesis [PATH:ko00220] | 10 |
| 00360 Phenylalanine metabolism [PATH:ko00360] | 9 |

| | |
|--|------------|
| 00400 Phenylalanine, tyrosine and tryptophan biosynthesis [PATH:ko00400] | 4 |
| 00290 Valine, leucine and isoleucine biosynthesis [PATH:ko00290] | 2 |
| 00300 Lysine biosynthesis [PATH:ko00300] | 0 |
| 09106 Metabolism of other amino acids | 85 |
| 00480 Glutathione metabolism [PATH:ko00480] | 34 |
| 00410 beta-Alanine metabolism [PATH:ko00410] | 20 |
| 00450 Selenocompound metabolism [PATH:ko00450] | 12 |
| 00430 Taurine and hypotaurine metabolism [PATH:ko00430] | 6 |
| 00440 Phosphonate and phosphinate metabolism [PATH:ko00440] | 5 |
| 00460 Cyanoamino acid metabolism [PATH:ko00460] | 5 |
| 00471 D-Glutamine and D-glutamate metabolism [PATH:ko00471] | 2 |
| 00472 D-Arginine and D-ornithine metabolism [PATH:ko00472] | 1 |
| 00473 D-Alanine metabolism [PATH:ko00473] | 0 |
| 09107 Glycan biosynthesis and metabolism | 408 |
| 01003 Glycosyltransferases [BR:ko01003] | 104 |
| 00536 Glycosaminoglycan binding proteins [BR:ko00536] | 39 |
| 00513 Various types of N-glycan biosynthesis [PATH:ko00513] | 38 |
| 00510 N-Glycan biosynthesis [PATH:ko00510] | 32 |
| 00537 Glycosylphosphatidylinositol (GPI)-anchored proteins [BR:ko00537] | 22 |
| 00535 Proteoglycans [BR:ko00535] | 20 |
| 00563 Glycosylphosphatidylinositol(GPI)-anchor biosynthesis [PATH:ko00563] | 20 |
| 00511 Other glycan degradation [PATH:ko00511] | 19 |
| 00515 Mannose type O-glycan biosynthesis [PATH:ko00515] | 16 |
| 00531 Glycosaminoglycan degradation [PATH:ko00531] | 16 |
| 00534 Glycosaminoglycan biosynthesis - heparan sulfate / heparin [PATH:ko00534] | 15 |
| 00514 Other types of O-glycan biosynthesis [PATH:ko00514] | 14 |
| 00601 Glycosphingolipid biosynthesis - lacto and neolacto series [PATH:ko00601] | 13 |
| 00603 Glycosphingolipid biosynthesis - globo and isoglobo series [PATH:ko00603] | 12 |
| 00532 Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate [PATH:ko00532] | 10 |
| 00512 Mucin type O-glycan biosynthesis [PATH:ko00512] | 9 |
| 00604 Glycosphingolipid biosynthesis - ganglio series [PATH:ko00604] | 6 |
| 00533 Glycosaminoglycan biosynthesis - keratan sulfate [PATH:ko00533] | 3 |
| 00540 Lipopolysaccharide biosynthesis [PATH:ko00540] | 0 |
| 01005 Lipopolysaccharide biosynthesis proteins [BR:ko01005] | 0 |
| 00550 Peptidoglycan biosynthesis [PATH:ko00550] | 0 |
| 01011 Peptidoglycan biosynthesis and degradation proteins [BR:ko01011] | 0 |
| 09108 Metabolism of cofactors and vitamins | 121 |
| 00760 Nicotinate and nicotinamide metabolism [PATH:ko00760] | 19 |
| 00790 Folate biosynthesis [PATH:ko00790] | 18 |
| 00670 One carbon pool by folate [PATH:ko00670] | 17 |
| 00770 Pantothenate and CoA biosynthesis [PATH:ko00770] | 14 |
| 00860 Porphyrin and chlorophyll metabolism [PATH:ko00860] | 14 |

| | |
|--|-----------|
| 00730 Thiamine metabolism [PATH:ko00730] | 10 |
| 00130 Ubiquinone and other terpenoid-quinone biosynthesis [PATH:ko00130] | 9 |
| 00830 Retinol metabolism [PATH:ko00830] | 8 |
| 00740 Riboflavin metabolism [PATH:ko00740] | 6 |
| 00750 Vitamin B6 metabolism [PATH:ko00750] | 3 |
| 00780 Biotin metabolism [PATH:ko00780] | 2 |
| 00785 Lipoic acid metabolism [PATH:ko00785] | 1 |
| 09109 Metabolism of terpenoids and polyketides | 66 |
| 00981 Insect hormone biosynthesis [PATH:ko00981] | 23 |
| 00900 Terpenoid backbone biosynthesis [PATH:ko00900] | 20 |
| 01006 Prenyltransferases [BR:ko01006] | 14 |
| 00903 Limonene and pinene degradation [PATH:ko00903] | 5 |
| 00281 Geraniol degradation [PATH:ko00281] | 2 |
| 00908 Zeatin biosynthesis [PATH:ko00908] | 1 |
| 01051 Biosynthesis of ansamycins [PATH:ko01051] | 1 |
| 00902 Monoterpenoid biosynthesis [PATH:ko00902] | 0 |
| 00909 Sesquiterpenoid and triterpenoid biosynthesis [PATH:ko00909] | 0 |
| 00904 Diterpenoid biosynthesis [PATH:ko00904] | 0 |
| 00906 Carotenoid biosynthesis [PATH:ko00906] | 0 |
| 00905 Brassinosteroid biosynthesis [PATH:ko00905] | 0 |
| 01008 Polyketide biosynthesis proteins [BR:ko01008] | 0 |
| 01052 Type I polyketide structures [PATH:ko01052] | 0 |
| 00522 Biosynthesis of 12-, 14- and 16-membered macrolides [PATH:ko00522] | 0 |
| 01059 Biosynthesis of enediynes antibiotics [PATH:ko01059] | 0 |
| 01056 Biosynthesis of type II polyketide backbone [PATH:ko01056] | 0 |
| 01057 Biosynthesis of type II polyketide products [PATH:ko01057] | 0 |
| 00253 Tetracycline biosynthesis [PATH:ko00253] | 0 |
| 00523 Polyketide sugar unit biosynthesis [PATH:ko00523] | 0 |
| 01054 Nonribosomal peptide structures [PATH:ko01054] | 0 |
| 01053 Biosynthesis of siderophore group nonribosomal peptides [PATH:ko01053] | 0 |
| 01055 Biosynthesis of vancomycin group antibiotics [PATH:ko01055] | 0 |
| 09110 Biosynthesis of other secondary metabolites | 36 |
| 00950 Isoquinoline alkaloid biosynthesis [PATH:ko00950] | 10 |
| 00965 Betalain biosynthesis [PATH:ko00965] | 5 |
| 00521 Streptomycin biosynthesis [PATH:ko00521] | 5 |
| 00960 Tropane, piperidine and pyridine alkaloid biosynthesis [PATH:ko00960] | 3 |
| 00524 Neomycin, kanamycin and gentamicin biosynthesis [PATH:ko00524] | 3 |
| 00401 Novobiocin biosynthesis [PATH:ko00401] | 2 |
| 00333 Prodigiosin biosyntheses [PATH:ko00333] | 2 |
| 00940 Phenylpropanoid biosynthesis [PATH:ko00940] | 1 |
| 00944 Flavone and flavonol biosynthesis [PATH:ko00944] | 1 |
| 00901 Indole alkaloid biosynthesis [PATH:ko00901] | 1 |
| 00311 Penicillin and cephalosporin biosynthesis [PATH:ko00311] | 1 |
| 00261 Monobactam biosynthesis [PATH:ko00261] | 1 |
| 00254 Aflatoxin biosynthesis [PATH:ko00254] | 1 |

| | |
|--|--------------|
| 00945 Stilbenoid, diarylheptanoid and gingerol biosynthesis [PATH:ko00945] | 0 |
| 00941 Flavonoid biosynthesis [PATH:ko00941] | 0 |
| 00942 Anthocyanin biosynthesis [PATH:ko00942] | 0 |
| 00943 Isoflavonoid biosynthesis [PATH:ko00943] | 0 |
| 00403 Indole diterpene alkaloid biosynthesis [PATH:ko00403] | 0 |
| 01058 Acridone alkaloid biosynthesis [PATH:ko01058] | 0 |
| 00232 Caffeine metabolism [PATH:ko00232] | 0 |
| 00966 Glucosinolate biosynthesis [PATH:ko00966] | 0 |
| 00402 Benzoxazinoid biosynthesis [PATH:ko00402] | 0 |
| 00332 Carbapenem biosynthesis [PATH:ko00332] | 0 |
| 00331 Clavulanic acid biosynthesis [PATH:ko00331] | 0 |
| 00525 Acarbose and validamycin biosynthesis [PATH:ko00525] | 0 |
| 00231 Puromycin biosynthesis [PATH:ko00231] | 0 |
| 00404 Staurosporine biosynthesis [PATH:ko00404] | 0 |
| 00405 Phenazine biosynthesis [PATH:ko00405] | 0 |
| 09111 Xenobiotics biodegradation and metabolism | 78 |
| 00983 Drug metabolism - other enzymes [PATH:ko00983] | 31 |
| 00982 Drug metabolism - cytochrome P450 [PATH:ko00982] | 15 |
| 00980 Metabolism of xenobiotics by cytochrome P450 [PATH:ko00980] | 13 |
| 00625 Chloroalkane and chloroalkene degradation [PATH:ko00625] | 6 |
| 00362 Benzoate degradation [PATH:ko00362] | 3 |
| 00627 Aminobenzoate degradation [PATH:ko00627] | 3 |
| 00930 Caprolactam degradation [PATH:ko00930] | 3 |
| 00643 Styrene degradation [PATH:ko00643] | 2 |
| 00791 Atrazine degradation [PATH:ko00791] | 1 |
| 00626 Naphthalene degradation [PATH:ko00626] | 1 |
| 00364 Fluorobenzoate degradation [PATH:ko00364] | 0 |
| 00361 Chlorocyclohexane and chlorobenzene degradation [PATH:ko00361] | 0 |
| 00623 Toluene degradation [PATH:ko00623] | 0 |
| 00622 Xylene degradation [PATH:ko00622] | 0 |
| 00633 Nitrotoluene degradation [PATH:ko00633] | 0 |
| 00642 Ethylbenzene degradation [PATH:ko00642] | 0 |
| 00351 1,1,1-Trichloro-2,2-bis(4-chlorophenyl)ethane (DDT) degradation [PATH:ko00351] | 0 |
| 00363 Bisphenol degradation [PATH:ko00363] | 0 |
| 00621 Dioxin degradation [PATH:ko00621] | 0 |
| 00624 Polycyclic aromatic hydrocarbon degradation [PATH:ko00624] | 0 |
| 00365 Furfural degradation [PATH:ko00365] | 0 |
| 00984 Steroid degradation [PATH:ko00984] | 0 |
| 09112 Enzyme families | 777 |
| 01001 Protein kinases [BR:ko01001] | 273 |
| 01009 Protein phosphatase and associated proteins [BR:ko01009] | 256 |
| 01002 Peptidases [BR:ko01002] | 224 |
| 00199 Cytochrome P450 [BR:ko00199] | 24 |
| 01000 Enzymes [BR:ko01000] | 0 |
| 00004 KEGG modules in global maps only | 0 |
| 09130 Environmental Information Processing | 2.755 |
| 09131 Membrane transport | 362 |

| | |
|---|-------------|
| 02000 Transporters [BR:ko02000] | 323 |
| 02010 ABC transporters [PATH:ko02010] | 23 |
| 02044 Secretion system [BR:ko02044] | 14 |
| 03070 Bacterial secretion system [PATH:ko03070] | 2 |
| 02060 Phosphotransferase system (PTS) [PATH:ko02060] | 0 |
| 09132 Signal transduction | 1824 |
| 04010 MAPK signaling pathway [PATH:ko04010] | 113 |
| 04151 PI3K-Akt signaling pathway [PATH:ko04151] | 109 |
| 04013 MAPK signaling pathway - fly [PATH:ko04013] | 103 |
| 04015 Rap1 signaling pathway [PATH:ko04015] | 98 |
| 04150 mTOR signaling pathway [PATH:ko04150] | 96 |
| 04014 Ras signaling pathway [PATH:ko04014] | 90 |
| 04024 cAMP signaling pathway [PATH:ko04024] | 85 |
| 04310 Wnt signaling pathway [PATH:ko04310] | 80 |
| 04022 cGMP - PKG signaling pathway [PATH:ko04022] | 75 |
| 04390 Hippo signaling pathway [PATH:ko04390] | 73 |
| 04068 FoxO signaling pathway [PATH:ko04068] | 71 |
| 04072 Phospholipase D signaling pathway [PATH:ko04072] | 71 |
| 04371 Apelin signaling pathway [PATH:ko04371] | 66 |
| 04020 Calcium signaling pathway [PATH:ko04020] | 66 |
| 04152 AMPK signaling pathway [PATH:ko04152] | 64 |
| 04391 Hippo signaling pathway - fly [PATH:ko04391] | 60 |
| 04070 Phosphatidylinositol signaling system [PATH:ko04070] | 60 |
| 04071 Sphingolipid signaling pathway [PATH:ko04071] | 60 |
| 04012 ErbB signaling pathway [PATH:ko04012] | 46 |
| 04066 HIF-1 signaling pathway [PATH:ko04066] | 43 |
| 04350 TGF-beta signaling pathway [PATH:ko04350] | 42 |
| 04668 TNF signaling pathway [PATH:ko04668] | 35 |
| 04630 Jak-STAT signaling pathway [PATH:ko04630] | 29 |
| 04370 VEGF signaling pathway [PATH:ko04370] | 28 |
| 04330 Notch signaling pathway [PATH:ko04330] | 26 |
| 04064 NF-kappa B signaling pathway [PATH:ko04064] | 26 |
| 04392 Hippo signaling pathway - multiple species [PATH:ko04392] | 23 |
| 04340 Hedgehog signaling pathway [PATH:ko04340] | 22 |
| 04011 MAPK signaling pathway - yeast [PATH:ko04011] | 21 |
| 04341 Hedgehog signaling pathway - fly [PATH:ko04341] | 20 |
| 02020 Two-component system [PATH:ko02020] | 12 |
| 04016 MAPK signaling pathway - plant [PATH:ko04016] | 11 |
| 02022 Two-component system [BR:ko02022] | 0 |
| 04075 Plant hormone signal transduction [PATH:ko04075] | 0 |
| 09133 Signaling molecules and interaction | 569 |
| 04516 Cell adhesion molecules and their ligands [BR:ko04516] | 104 |
| 04040 Ion channels [BR:ko04040] | 90 |
| 04031 GTP-binding proteins [BR:ko04031] | 86 |
| 04030 G protein-coupled receptors [BR:ko04030] | 68 |
| 04090 CD Molecules [BR:ko04090] | 65 |
| 04080 Neuroactive ligand-receptor interaction [PATH:ko04080] | 44 |
| 04512 ECM-receptor interaction [PATH:ko04512] | 26 |
| 04091 Lectins [BR:ko04091] | 22 |
| 03310 Nuclear receptors [BR:ko03310] | 20 |
| 04514 Cell adhesion molecules (CAMs) [PATH:ko04514] | 16 |
| 04060 Cytokine-cytokine receptor interaction [PATH:ko04060] | 15 |
| 04052 Cytokines and growth factors [BR:ko04052] | 12 |

| | |
|---|--------------|
| 02042 Bacterial toxins [BR:ko02042] | 1 |
| 04050 Cytokine receptors [BR:ko04050] | 0 |
| 09140 Cellular Processes | 2.206 |
| 09141 Transport and catabolism | 968 |
| 04147 Exosome [BR:ko04147] | 397 |
| 04144 Endocytosis [PATH:ko04144] | 136 |
| 04140 Autophagy - animal [PATH:ko04140] | 88 |
| 04142 Lysosome [PATH:ko04142] | 86 |
| 04146 Peroxisome [PATH:ko04146] | 58 |
| 04138 Autophagy - yeast [PATH:ko04138] | 58 |
| 04145 Phagosome [PATH:ko04145] | 54 |
| 04137 Mitophagy - animal [PATH:ko04137] | 41 |
| 04136 Autophagy - other eukaryotes [PATH:ko04136] | 27 |
| 04139 Mitophagy - yeast [PATH:ko04139] | 19 |
| 02048 Prokaryotic Defense System [BR:ko02048] | 4 |
| 09143 Cell growth and death | 582 |
| 04110 Cell cycle [PATH:ko04110] | 85 |
| 04218 Cellular senescence [PATH:ko04218] | 80 |
| 04114 Oocyte meiosis [PATH:ko04114] | 66 |
| 04111 Cell cycle - yeast [PATH:ko04111] | 61 |
| 04210 Apoptosis [PATH:ko04210] | 59 |
| 04214 Apoptosis - fly [PATH:ko04214] | 56 |
| 04217 Necroptosis [PATH:ko04217] | 54 |
| 04113 Meiosis - yeast [PATH:ko04113] | 48 |
| 04115 p53 signaling pathway [PATH:ko04115] | 30 |
| 04216 Ferroptosis [PATH:ko04216] | 22 |
| 04215 Apoptosis - multiple species [PATH:ko04215] | 18 |
| 04112 Cell cycle - Caulobacter [PATH:ko04112] | 3 |
| 09144 Cellular community - eukaryotes | 336 |
| 04510 Focal adhesion [PATH:ko04510] | 97 |
| 04530 Tight junction [PATH:ko04530] | 79 |
| 04550 Signaling pathways regulating pluripotency of stem cells [PATH:ko04550] | 62 |
| 04520 Adherens junction [PATH:ko04520] | 53 |
| 04540 Gap junction [PATH:ko04540] | 45 |
| 09145 Cellular community - prokaryotes | 8 |
| 02024 Quorum sensing [PATH:ko02024] | 6 |
| 02025 Biofilm formation - Pseudomonas aeruginosa [PATH:ko02025] | 1 |
| 02026 Biofilm formation - Escherichia coli [PATH:ko02026] | 1 |
| 05111 Biofilm formation - Vibrio cholerae [PATH:ko05111] | 0 |
| 09142 Cell motility | 312 |
| 04812 Cytoskeleton proteins [BR:ko04812] | 224 |
| 04810 Regulation of actin cytoskeleton [PATH:ko04810] | 88 |
| 02030 Bacterial chemotaxis [PATH:ko02030] | 0 |
| 02035 Bacterial motility proteins [BR:ko02035] | 0 |
| 02040 Flagellar assembly [PATH:ko02040] | 0 |
| 09190 Unclassified | 223 |
| 09191 Metabolism | 78 |
| 99999 Others | 34 |
| 99982 Energy metabolism | 15 |
| 99983 Lipid metabolism | 9 |
| 99981 Carbohydrate metabolism | 7 |
| 99985 Amino acid metabolism | 6 |

| | |
|--|-----------|
| 99987 Metabolism of cofactors and vitamins | 6 |
| 99986 Glycan biosynthesis and metabolism | 1 |
| 99984 Nucleotide metabolism | 0 |
| 99988 Biosynthesis and biodegradation of secondary metabolites | 0 |
| 09192 Genetic information processing | 44 |
| 99973 Transcription related proteins | 12 |
| 99975 Protein folding and associated processing | 12 |
| 99976 Replication, recombination and repair proteins | 11 |
| 99974 Translation proteins | 8 |
| 99998 Others | 1 |
| 09193 Cellular processes and signaling | 81 |
| 99995 Signal transduction mechanisms | 43 |
| 99992 Membrane and intracellular structural molecules | 29 |
| 99993 Cell motility and secretion | 4 |
| 99978 Cell growth | 3 |
| 99994 Inorganic ion transport and metabolism | 2 |
| 99977 Transport | 0 |
| 09196 Viral protein family | 0 |
| 99979 Unclassified viral proteins | 0 |
| 09194 Poorly characterized | 20 |
| 99997 Function unknown | 12 |
| 99996 General function prediction only | 8 |

APÊNDICE B – As 25 vias metabólicas (KEGG) mais abundantes na montagem K24_B145

Tabela do Apêndice 1 - Números de *contigs* que abrigam as 25 vias metabólicas KEGG mais abundantes da montagem K24_B145. *Vias também presentes, entre as 25 mais anotadas, no estudo transcriptômico do télson de *Tityus serrulatus* realizado por Alvarenga e colaboradores (2012).

| Via metabólica | Nº <i>contigs</i> |
|---|-------------------|
| Chaperones and folding catalysts [BR:ko03110] | 124 |
| Chromosome and associated proteins [BR:ko03036] | 670 |
| Cytoskeleton proteins [BR:ko04812] | 224 |
| DNA repair and recombination proteins [BR:ko03400] | 192 |
| Endocytosis [PATH:ko04144] | 136 |
| Exosome [BR:ko04147] | 397 |
| HTLV-I infection [PATH:ko05166] | 124 |
| Human papillomavirus infection [PATH:ko05165] | 133 |
| Membrane trafficking [BR:ko04131] | 751 |
| Messenger RNA Biogenesis [BR:ko03019] | 281 |
| Mitochondrial biogenesis [BR:ko03029] | 215 |
| Pathways in cancer [PATH:ko05200] | 196 |
| Peptidases [BR:ko01002] | 224 |
| Protein kinases [BR:ko01001] | 273 |
| Protein phosphatase and associated proteins [BR:ko01009] | 256 |
| Protein processing in endoplasmic reticulum [PATH:ko04141]* | 114 |
| Ribosome biogenesis [BR:ko03009] | 191 |
| RNA transport [PATH:ko03013]* | 115 |
| Spliceosome [BR:ko03041]* | 232 |
| Thermogenesis [PATH:ko04714] | 121 |
| Transcription factors [BR:ko03000] | 267 |
| Transcription machinery [BR:ko03021] | 178 |
| Transfer RNA biogenesis [BR:ko03016] | 128 |
| Transporters [BR:ko02000] | 323 |
| Ubiquitin system [BR:ko04121] | 388 |

APÊNDICE C – Componentes de peçonha anotados pelo BLAST2GO

| Componentes de peçonha que deram <i>hit</i> com <i>contigs</i> anotados pelo BLAST2GO | Classe |
|---|----------------|
| angiotensin-converting enzyme-like | ACE |
| angiotensin-converting enzyme-like [Limulus polyphemus] | ACE |
| CRISP Allergen PR-1-like | Alérgeno |
| mite allergen Lep d 7-like | Alérgeno |
| MYSP_BLOTA ame: Full=Paramyosin ame: Allergen=Blo t 11 | Alérgeno |
| profilin allergen | Alérgeno |
| SCX4_TITSE ame: Full=Toxin Ts4 ame: Full=Non-toxic xP Short=N ame: Full=P-Allerg-alpha* ame: Full=PT-Immun-alpha* ame: Full=Tityustoxin VI Short=Ts VI Short= -VI Short= Short= ame: Full=Tityustoxin-6 Short=Ts-6 Flags: Precursor | Alérgeno |
| Venom allergen partial | Alérgeno |
| type-1 angiotensin II receptor-associated -like [Limulus polyphemus] | Angiotensina |
| anticoagulant peptide | Anticoagulante |
| antimicrobial peptide precursor [Tityus costatus] | Antimicrobiana |
| venom antimicrobial peptide-7 | Antimicrobiana |
| endothelin-converting [Pediculus humanus corporis] | ECE |
| endothelin-converting enzyme 1 | ECE |
| endothelin-converting enzyme 1 isoform X2 | ECE |
| endothelin-converting enzyme 1 isoform X2 [Camponotus floridanus] | ECE |
| endothelin-converting enzyme 1 isoform X3 [Harpegnathos saltator] | ECE |
| endothelin-converting enzyme 1-like | ECE |
| endothelin-converting enzyme 1-like isoform X1 | ECE |
| endothelin-converting enzyme 1-like isoform X1 [Agrilus planipennis] | ECE |
| endothelin-converting enzyme 1-like isoform X1 [Apis florea] | ECE |
| endothelin-converting enzyme 1-like isoform X1 [Copidosoma floridanum] | ECE |
| endothelin-converting enzyme 1-like isoform X3 | ECE |
| endothelin-converting enzyme 2-like isoform X1 | ECE |
| endothelin-converting partial | ECE |
| HYAL1_MESMA ame: Full=Hyaluronidase-1 Short= 1 Short=HYA1 ame: Full=Hyaluronoglucosaminidase-1 ame: Full=venom spreading factor Flags: Precursor | Hialuronidase |
| HYAL1_TITSE ame: Full=Hyaluronidase 1 Short= yal-1 ame: Full=Hyaluronoglucosaminidase ame: Full=venom spreading factor Flags: Precursor | Hialuronidase |
| HYAL2_TITSE ame: Full=Hyaluronidase 2 Short= yal-2 ame: Full=Hyaluronoglucosaminidase ame: Full=venom spreading factor Flags: Precursor | Hialuronidase |
| Bradykinin-potentiating peptide-like peptide Bpp-1 [Androctonus bicolor] | Hipotensor |
| insulin growth factor binding | IGF |
| insulin-like growth factor 2 mRNA-binding 1 isoform X1 | IGF |
| insulin-like growth factor-binding 7 | IGF |
| Insulin-like growth factor-binding complex acid labile partial | IGF |
| insulin-like growth factor-binding complex acid labile subunit | IGF |
| insulin-like growth factor-binding -related 1 | IGF |
| insulin-like growth factor-binding -related 1 [Priapulus caudatus] | IGF |
| hormone-sensitive lipase | Lipase |
| hormone-sensitive lipase-like | Lipase |
| monoacylglycerol lipase ABHD12-like | Lipase |
| monoacylglycerol lipase ABHD12-like [Parasteatoda tepidariorum] | Lipase |

| | |
|---|---------------|
| monoglyceride lipase isoform X2 | Lipase |
| Pancreatic lipase-related 1 | Lipase |
| pancreatic lipase-related 2-like | Lipase |
| Pancreatic lipase-related partial | Lipase |
| pancreatic triacylglycerol lipase-like | Lipase |
| sn1-specific diacylglycerol lipase alpha isoform X4 | Lipase |
| sn1-specific diacylglycerol lipase alpha-like isoform X3 | Lipase |
| sn1-specific diacylglycerol lipase beta | Lipase |
| KAX45_TITCO ame: Full=Potassium channel toxin alpha-KTx Flags: Precursor | Neurotoxina |
| KBX1_TITSE ame: Full=Potassium channel toxin -beta ame: Full=Potassium channel toxin beta-KTx 1 ame: Full=Tityustoxin K-beta Short= K beta Short= -K beta Short= beta ame: Full=Ts8 ame: Full= 2 Flags: Precursor | Neurotoxina |
| KEX12_TITSE ame: Full=Potassium channel toxin epsilon-KTx ame: Full=Ts12 ame: Full=ep2 Flags: Precursor | Neurotoxina |
| potassium channel blocker p x21-1 | Neurotoxina |
| potassium channel blocker partial | Neurotoxina |
| potassium channel blocker x-5 [Androctonus bicolor] | Neurotoxina |
| SCX3_TITSE ame: Full=Toxin-5 ame: Full=Tityustoxin V Short=Ts V Short= ame: Full=Toxin V Contains: ame: Full=Alpha-mammal toxin Ts3 ame: Full=PT-Mice-alpha ame: Full=Tityustoxin IV Short=Toxin IV Short=Ts IV Short= ame: Full=Tityustoxin IV-5 Short=Toxin IV-5 Short= -5 ame: Full=Toxin-4 Contains: ame: Full=Toxin-3 ame: Full=Tityustoxin III Short=Toxin III Short=Ts III Short= Flags: Precursor | Neurotoxina |
| sodium channel toxin partial | Neurotoxina |
| anti-proliferation factor btg1 tob | Outros |
| Cysteine-rich venom 6 | Outros |
| cysteine-rich venom LEI1-like isoform X1 [Parasteatoda tepidariorum] | Outros |
| venom -5 | Outros |
| venom -7 | Outros |
| venom -9 | Outros |
| venom p- partial | Outros |
| venom p-1 | Outros |
| venom partial | Outros |
| venom peptide 16x | Outros |
| venom peptide 6Tx1 | Outros |
| venom peptide 6Tx3 | Outros |
| venom peptide Htglin alpha | Outros |
| venom peptide meuPep26 | Outros |
| venom peptide meuPep34 | Outros |
| venom peptide partial | Outros |
| venom phosphodiesterase 2-like | Outros |
| venom phosphodiesterase 2-like [Branchiostoma belcheri] | Outros |
| venom VP3 | Outros |
| venom VP6 | Outros |
| VP30_LYCMC name: Full=venom Flags: Precursor | Outros |
| 85 88 kDa calcium-independent phospholipase A2 | Phospholipase |
| basic phospholipase A2 Ts-G6D49-like isoform X1 [Octopus bimaculoides] | Phospholipase |
| calcium-independent phospholipase A2-gamma isoform X3 [Camponotus floridanus] | Phospholipase |
| cytosolic phospholipase A2-like isoform X5 [Crassostrea gigas] | Phospholipase |

| | |
|---|---------------|
| group XIIA secretory phospholipase A2 | Phospholipase |
| group XV phospholipase A2-like | Phospholipase |
| inactive phospholipase C 2 | Phospholipase |
| patatin-like phospholipase domain-containing 2 | Phospholipase |
| Phospholipase A2 | Phospholipase |
| phospholipase A2 [Ixodes scapularis] | Phospholipase |
| phospholipase A-2-activating | Phospholipase |
| phospholipase A2-like | Phospholipase |
| phospholipase ABHD3 | Phospholipase |
| phospholipase B-like 2 | Phospholipase |
| phospholipase D1-like isoform X1 [Parasteatoda tepidariorum] | Phospholipase |
| Phospholipase D3 | Phospholipase |
| phospholipase D3-like | Phospholipase |
| phospholipase DDHD2-like | Phospholipase |
| phospholipase partial | Phospholipase |
| 26S protease regulatory subunit 10B | Protease |
| 26S protease regulatory subunit 4 | Protease |
| 26S protease regulatory subunit 6A-B | Protease |
| 26S protease regulatory subunit 6B | Protease |
| 26S protease regulatory subunit 7 | Protease |
| 26S protease regulatory subunit 8 | Protease |
| ADAM 17-like protease | Protease |
| ADAM 9 [Pediculus humanus corporis] | Protease |
| ADAMTS 1 isoform X3 | Protease |
| ADAMTS 3 | Protease |
| ADAMTS 5 | Protease |
| ADAMTS partial | Protease |
| antarease-like Zn- | Protease |
| antarease-like Zn- partial [Tityus serrulatus] | Protease |
| astacin-like metalloprotease toxin 1 | Protease |
| astacin-like metalloprotease toxin 5 | Protease |
| Astacin-like metalloprotease partial | Protease |
| ATP-dependent Clp protease ATP-binding subunit clpX- mitochondrial isoform X1 | Protease |
| ATP-dependent Clp protease proteolytic mitochondrial-like | Protease |
| ATP-dependent Zinc metalloprotease YME1 | Protease |
| brain-specific serine protease 4-like isoform X2 | Protease |
| CAAX prenyl protease 1 homolog | Protease |
| CAAX prenyl protease 2-like | Protease |
| cathepsin B [Ixodes scapularis] | Protease |
| cathepsin D | Protease |
| cathepsin L | Protease |
| cathepsin L partial | Protease |
| cathepsin L-associated | Protease |
| cathepsin L-like | Protease |
| cathepsin O-like | Protease |
| Cathepsin partial | Protease |

| | |
|--|-----------|
| Cathepsin Z | Protease |
| chymotrypsin-like protease - partial | Protease |
| chymotrypsinogen A-like | Protease |
| chymotrypsinogen B-like | Protease |
| cysteine protease ATG4B-like | Protease |
| cysteine protease ATG4D-like | Protease |
| furin-like protease 1 partial | Protease |
| furin-like protease 2 | Protease |
| furin-like protease isoforms 1 1-X 2 | Protease |
| immunoglobulin a1 protease | Protease |
| lon protease homolog peroxisomal-like | Protease |
| lon protease mitochondrial | Protease |
| mannan-binding lectin serine protease 2- partial | Protease |
| membrane-bound transcription factor site-1 protease-like | Protease |
| membrane-bound transcription factor site-2 protease-like | Protease |
| metalloprotease TIK11 | Protease |
| metalloprotease TIK11 isoform X3 [Felis catus] | Protease |
| metalloserulase 2 | Protease |
| metalloserulase 5 | Protease |
| metalloserulase 6 | Protease |
| metalloserulase 8 | Protease |
| protease-associated domain-containing of 21 [Ixodes scapularis] | Protease |
| sentrin-specific protease 6-like isoform X3 | Protease |
| sentrin-specific protease 8-like | Protease |
| serine protease P54 | Protease |
| transglutaminase protease partial | Protease |
| Ubiquitin-like-specific protease 1 | Protease |
| ufm1-specific protease 2-like | Protease |
| VAMP-associated involved in inositol [Ixodes scapularis] | Protease |
| venom serine protease | Protease |
| acidic mammalian chitinase | Quitinase |
| chitinase domain-containing 1 | Quitinase |
| Chitinase domain-containing partial | Quitinase |
| chitinase-3 1 | Quitinase |
| endochitinase | Quitinase |
| probable chitinase 3 | Quitinase |
| KS4_MESMA ame: Full=Toxin xKS4 ame: Full=KS3 toxin Flags: Precursor [Mesobuthus martensii] | Toxina |
| meutoxin-1 precursor | Toxina |
| P9797_TITST ame: Full=Peptide partial [Tityus stigmurus] | Toxina |
| plancitoxin-1 [Tribolium castaneum] | Toxina |
| ras-related C3 botulinum toxin substrate 1-like [Parasteatoda tepidariorum] | Toxina |
| ras-related C3 botulinum toxin substrate 1-like isoform X2 [Priapulus caudatus] | Toxina |
| scorpion toxin | Toxina |
| SCX12_TITOB ame: Full=Toxin To12 ame: Full=T-beta* Flags: Precursor | Toxina |
| SCX5_TITBA ame: Full=Toxin x5 ame: Full=T-alpha* Flags: Precursor | Toxina |

| | |
|--|--------|
| SCX8_TITPA ame: Full=Toxin Tpa8 ame: Full=T Flags: Precursor | Toxina |
| Stonustoxin subunit partial | Toxina |
| toxin 14 | Toxina |
| U24-ctenitoxin-Pn1a-like [Parasteatoda tepidariorum] | Toxina |
| U9-ctenitoxin-Pr1a-like isoform X1 | Toxina |
| venom toxin | Toxina |
| venom toxin meuTx18 | Toxina |
| venom toxin-like peptide | Toxina |

APÊNDICE D – Lista das proteínas 100% similares às ORFs dos *contigs* da montagem K24_B145

| Banco de Dados | Identidade | Nome |
|----------------|----------------|--|
| NR | gb AMO02503.1 | endothelin-converting protein 3 |
| NR | gb AMO02504.1 | endothelin-converting protein 4 |
| NR | gb AMO02505.1 | endothelin-converting protein 5 |
| NR | gb AMO02507.1 | endothelin-converting protein 7 |
| NR | gb AMO02518.1 | xaa-pro aminopeptidase 3 |
| NR | gb AMO02519.1 | methionine aminopeptidase 2 |
| NR | gb AMO02522.1 | aminopeptidase N1 |
| NR | gb AMO02525.1 | glutamyl aminopeptidase |
| NR | gb AMO02526.1 | aminopeptidase A |
| NR | gb AMO02533.1 | xaa-pro aminopeptidase 1d |
| NR | gb AMO02535.1 | methionine aminopeptidase 1 |
| NR | gb AMO02536.1 | aminopeptidase N7 |
| NR | gb AMO02540.1 | astacin-like protein metalloproteinase 18 |
| NR | gb AMO02542.1 | astacin-like protein metalloproteinase 20 |
| NR | gb AMO02545.1 | carboxypeptidase D |
| NR | gb AMO02546.1 | carboxypeptidase E |
| NR | gb AMO02547.1 | carboxypeptidase M |
| NR | gb AMO02551.1 | carboxypeptidase B |
| NR | gb AMO02552.1 | carboxypeptidase Q |
| NR | gb AMO02555.1 | metalloprotease 13 |
| NR | gb AMO02557.1 | metalloprotease 15 |
| NR | gb AMO02559.1 | metalloprotease 17 |
| NR | gb AMO02561.1 | metalloprotease 19 |
| NR | gb AMO02563.1 | serine protease 1 |
| NR | gb AMO02566.1 | serine protease 4 |
| NR | gb AMO02569.1 | serine protease 7 |
| NR | XP_015905891.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit |
| NR | XP_015905291.1 | ubiquitin-conjugating enzyme E2 H |
| NR | XP_015907099.1 | protein lin-7 homolog C |
| NR | XP_015907787.1 | protein transport protein Sec61 subunit alpha |
| NR | XP_015912859.1 | actin, clone 403-like |
| NR | XP_015919889.1 | AP-2 complex subunit mu |
| NR | XP_015920067.1 | serine/threonine-protein phosphatase 4 catalytic subunit |
| NR | XP_015923060.1 | F-box/SPRY domain-containing protein 1-like isoform X2 |
| NR | XP_015924332.1 | tubulin beta chain |
| NR | gb AIJ02111.3 | metalloprotease 3 |
| NR | XP_019865923.1 | PREDICTED: ADP-ribosylation factor 6 |
| NR | gb ABI52808.1 | 40S ribosomal protein S23 |
| NR | XP_021000233.1 | serine/threonine-protein phosphatase 2A catalytic subunit beta isoform |
| NR | WP_086378842.1 | IS66 family insertion sequence hypothetical protein |
| NR | XP_022197100.1 | AP-1 complex subunit sigma-2 isoform X2 |

NR sp|P69940.2 RecName: Full=Potassium channel toxin TsTXK-beta; AltName: Full=Potassium channel toxin beta-KTx 1; AltName: Full=Tityustoxin K-beta; Short=TsTX K beta; Short=TsTX-K beta; Short=TsTXKbeta; AltName: Full=Ts8; AltName: Full=TsK2; Flags: Precursor

NR emb|CAP02971.1 putative antitoxin RelB homolog induced by DNA damage (plasmid)

NR gb|ACD11814.1 hypothetical protein

NR gb|ACD11889.1 hypothetical protein

NR gb|ACD11890.1 hypothetical protein

NR gb|ACD11935.1 hypothetical protein

NR gb|ACD11982.1 hypothetical protein

NR XP_002410492.1 Ca2+ sensor, putative

NR XP_002414033.1 DEAD box ATP-dependent RNA helicase, putative

NR emb|CAY61893.1 60s ribosomal protein L10

NR sp|O77463.1 RecName: Full=Toxin Ts4; AltName: Full=Non-toxic protein TsNTxP; Short=NTxP; AltName: Full=P-Allerg-alpha* NaTx4.1; AltName: Full=PT-Immun-alpha* NaTx4.2; AltName: Full=Tityustoxin VI; Short=Ts VI; Short=TsTX-VI; Short=TsTXVI; Short=TsVI; AltName: Full=Tityustoxin-6; Short=Ts-6; Flags: Precursor

NR gb|ADY39555.1 putative homeobox protein

NR gb|ADY62663.1 putative LIM protein

NR gb|EGJ65646.1 transposase, IS4 family

NR WP_004693838.1 MULTISPECIES: peroxiredoxin

NR WP_010326440.1 MULTISPECIES: acyl-CoA thioesterase

NR WP_010326927.1 MULTISPECIES: IS30 family transposase

NR sp|P86822.2 RecName: Full=Ts19; AltName: Full=Potassium channel toxin beta-Ktx 2; Contains: RecName: Full=Ts19 fragment I; AltName: Full=Potassium channel toxin beta-Ktx 2-1; Contains: RecName: Full=Ts19 fragment II; Short=Ts19 Frag-II; AltName: Full=Potassium channel toxin beta-Ktx 2-2; Contains: RecName: Full=Ts19 fragment III; Short=Ts19 Frag-III; Flags: Precursor

NR emb|CDJ26726.1 astacin-like metallopeptidase 5b protein

NR emb|CDJ26732.1 cathepsin L-like cysteine peptidase 3 protein

NR emb|CDJ26737.1 cathepsin F-like cysteine peptidase protein

NR emb|CDJ26739.1 cysteine peptidase 1 family C54 protein

NR emb|CDJ26741.1 cysteine peptidase 3 family C54 protein

NR sp|S6CWV8.1 RecName: Full=Antimicrobial peptide TsAP-1; AltName: Full=Non-disulfide-bridged peptide 4.22; Short=NDBP-4.22; Flags: Precursor

NR gb|AHZ63126.1 putative profilin allergen

NR gb|KDR22828.1 Protein yippee-like 5

NR gb|KFM61734.1 Ras-related C3 botulinum toxin substrate 1

NR gb|KFM67587.1 Ubiquitin-conjugating enzyme E2 G1

NR gb|KFM68186.1 Eukaryotic peptide chain release factor subunit 1

NR gb|KFM70872.1 Ubiquitin-conjugating enzyme E2 G2

NR gb|KFM74024.1 60S ribosomal protein L26

NR gb|KFM76942.1 DNA-directed RNA polymerase II subunit RPB2

NR gb|KFM78957.1 SUMO-conjugating enzyme UBC9

NR gb|KFM79737.1 ADP-ribosylation factor 1

NR gb|KFM80559.1 Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1

NR gb|KFM82434.1 DNA-directed RNA polymerase II subunit RPB7

NR gb|AIJ02109.2 metalloserrulase 1

| | | |
|-----------|----------------|---|
| NR | gb AIJ02110.2 | metalloserulase 2 |
| NR | gb AIJ02113.2 | metalloserulase 5 |
| NR | gb AIJ02114.2 | metalloserulase 6 |
| NR | gb AIJ02118.2 | endothelin-converting enzyme 1 |
| NR | sp Q5G8B5.1 | RecName: Full=Putative antimicrobial peptide clone 4; Flags: Precursor |
| NR | gb AJF94912.1 | cathepsin L3 |
| NR | XP_012889614.1 | PREDICTED: 40S ribosomal protein S13 |
| NR | XP_013789536.1 | histone H2B-like |
| NR | XP_013793949.1 | AP-2 complex subunit sigma-like |
| NR | XP_013775675.1 | small nuclear ribonucleoprotein Sm D2 |
| NR | XP_013779577.1 | serine/threonine-protein phosphatase 6 catalytic subunit isoform X1 |
| NR | XP_013782567.1 | 60S ribosomal protein L37a isoform X2 |
| NR | XP_013790347.1 | actin-related protein 2/3 complex subunit 4 isoform X1 |
| NR | NP_001301107.1 | ras-related protein Rab-11B |
| NR | gb ALX72359.1 | cell protein CPR4 |
| SwissProt | sp Q9Y333.1 | RecName: Full=U6 snRNA-associated Sm-like protein LSm2; AltName: Full=Protein G7b; AltName: Full=Small nuclear ribonuclear protein D homolog; AltName: Full=snRNP core Sm-like protein Sm-x5 |
| SwissProt | sp P18603.1 | RecName: Full=Actin, clone 403; Flags: Precursor |
| SwissProt | sp P26990.3 | RecName: Full=ADP-ribosylation factor 6 |
| SwissProt | sp P42325.2 | RecName: Full=Neurocalcin homolog; Short=DrosNcA |
| SwissProt | sp P46114.1 | RecName: Full=Potassium channel toxin alpha-KTx 4.1; AltName: Full=TSK4; AltName: Full=Tityustoxin K-alpha; Short=TsTX-K-alpha; Short=TyKalpha; AltName: Full=Toxin II-9; AltName: Full=Ts7 |
| SwissProt | sp Q0GY45.1 | RecName: Full=Potassium channel toxin TtrKIK; Flags: Precursor |
| SwissProt | sp P21251.2 | RecName: Full=Calmodulin; Short=CaM |
| SwissProt | sp P10987.4 | RecName: Full=Actin-5C; Flags: Precursor |
| SwissProt | sp P02572.3 | RecName: Full=Actin-42A; Flags: Precursor |
| SwissProt | sp P49393.2 | RecName: Full=40S ribosomal protein S13 |
| SwissProt | sp P06603.1 | RecName: Full=Tubulin alpha-1 chain |
| SwissProt | sp P05214.1 | RecName: Full=Tubulin alpha-3 chain; AltName: Full=Alpha-tubulin 3/7; AltName: Full=Alpha-tubulin isotype M-alpha-3/7; AltName: Full=Tubulin alpha-3/alpha-7 chain; Contains: RecName: Full=Detyrosinated tubulin alpha-3 chain |
| SwissProt | sp P18288.1 | RecName: Full=Tubulin alpha chain, testis-specific; Contains: RecName: Full=Detyrosinated tubulin alpha chain, testis-specific |
| SwissProt | sp P02554.1 | RecName: Full=Tubulin beta chain; AltName: Full=Beta-tubulin |
| SwissProt | sp P51643.2 | RecName: Full=ADP-ribosylation factor 1 |
| SwissProt | sp Q10453.2 | RecName: Full=Histone H3.3 type 1 |
| SwissProt | sp Q9BQE3.1 | RecName: Full=Tubulin alpha-1C chain; AltName: Full=Alpha-tubulin 6; AltName: Full=Tubulin alpha-6 chain; Contains: RecName: Full=Detyrosinated tubulin alpha-1C chain |
| SwissProt | sp P0C8X5.1 | RecName: Full=Toxin Tst3; AltName: Full=P-alpha* NaTx3.5; Flags: Precursor |
| SwissProt | sp Q24117.1 | RecName: Full=Dynein light chain 1, cytoplasmic; AltName: Full=8 kDa dynein light chain; AltName: Full=Cut up protein |
| SwissProt | sp P0CB56.1 | RecName: Full=Potassium channel toxin alpha-KTx 4.6; AltName: Full=Tst26 |
| SwissProt | sp P30883.1 | RecName: Full=Tubulin beta-4 chain |
| SwissProt | sp P60517.1 | RecName: Full=Gamma-aminobutyric acid receptor-associated protein; AltName: Full=GABA(A) receptor-associated protein; Flags: Precursor |

| | | |
|-----------|-------------|---|
| SwissProt | sp P0C8W4.2 | RecName: Full=Potassium channel toxin TstKMK; AltName: Full=Toxin 5536; AltName: Full=beta-Ktx; Flags: Precursor |
| SwissProt | sp P02299.4 | RecName: Full=Histone H3 |
| SwissProt | sp P61273.1 | RecName: Full=Dynein light chain 1, cytoplasmic; AltName: Full=Dynein light chain LC8-type 1 |
| SwissProt | sp P61209.2 | RecName: Full=ADP-ribosylation factor 1 |
| SwissProt | sp Q8WSF1.3 | RecName: Full=Histone H3.3 |
| SwissProt | sp Q8STF0.3 | CALM_STRIE RecName: Full=Calmodulin; Short=CaM |
| SwissProt | sp P62145.2 | CALM_APLCA RecName: Full=Calmodulin; Short=CaM |
| SwissProt | sp P62277.2 | RS13_HUMAN RecName: Full=40S ribosomal protein S13; AltName: Full=Small ribosomal subunit protein uS15 |
| SwissProt | sp P84079.2 | ARF1_RAT RecName: Full=ADP-ribosylation factor 1 |
| SwissProt | sp P62330.2 | ARF6_HUMAN RecName: Full=ADP-ribosylation factor 6 |
| SwissProt | sp Q86FP7.1 | RS23_DERVA RecName: Full=40S ribosomal protein S23 |
| SwissProt | sp Q9GRJ3.1 | RS23_LUMRU RecName: Full=40S ribosomal protein S23 |
| SwissProt | sp Q6P9V9.1 | TBA1B_RAT RecName: Full=Tubulin alpha-1B chain; AltName: Full=Alpha-tubulin 2; AltName: Full=Tubulin alpha-2 chain; Contains: RecName: Full=Detyrosinated tubulin alpha-1B chain |
| SwissProt | sp P84245.2 | H33_RAT RecName: Full=Histone H3.3 |
| SwissProt | sp P68370.1 | TBA1A_RAT RecName: Full=Tubulin alpha-1A chain; AltName: Full=Alpha-tubulin 1; AltName: Full=Tubulin alpha-1 chain; Contains: RecName: Full=Detyrosinated tubulin alpha-1A chain |
| SwissProt | sp P68365.1 | TBA1C_CRIGR RecName: Full=Tubulin alpha-1C chain; AltName: Full=Alpha-tubulin 3; AltName: Full=Alpha-tubulin III; AltName: Full=Tubulin alpha-3 chain; Contains: RecName: Full=Detyrosinated tubulin alpha-1C chain |
| SwissProt | sp P69071.2 | H3_DERIM RecName: Full=Histone H3, embryonic |
| SwissProt | sp S6D3A7.1 | RecName: Full=Antimicrobial peptide TsAP-2; AltName: Full=Non-disulfide-bridged peptide 4.23; Short=NDBP-4.23; Flags: Precursor |
| SwissProt | sp P86392.3 | RecName: Full=Venom metalloproteinase antarease TserMP_A; Short=VMMPA |
| SwissProt | sp V9Z9A3.1 | RecName: Full=Venom metalloproteinase antarease-like TserMP_B; Short=VMMPA |
| SwissProt | sp P41340.1 | RecName: Full=Actin-3; Flags: Precursor |
| SwissProt | sp Q5G8A6.1 | RecName: Full=Scorpine-like peptide Tco 41.46-2; Contains: RecName: Full=Scorpine-like peptide Tco 42.14; AltName: Full=Tco-beta-KTx; Short=TcobetaKTx; Flags: Precursor |
| SwissProt | sp Q64400.3 | RecName: Full=Histone H3.2 |
| SwissProt | sp P0C176.1 | RecName: Full=Peptide TsPep3; AltName: Full=Ts13 |
| Scorpion | AAB37719.2 | toxin IV-5 [Tityus serrulatus] |
| Scorpion | AAK12730.2 | RNA polymerase II largest subunit, partial [Heterometrus spinifer] |
| Scorpion | AAK12731.1 | RNA polymerase II largest subunit, partial [Heterometrus spinifer] |
| Scorpion | ACC95861.1 | shaker cognate b, partial [Mesobuthus martensii] |
| Scorpion | ACY42772.1 | G protein-coupled receptor kinase, partial [Heterometrus spinifer] |
| Scorpion | ACY42916.1 | myosin, partial [Heterometrus spinifer] |
| Scorpion | ACY42968.1 | myosin, partial [Heterometrus spinifer] |
| Scorpion | ACY43122.1 | arg methyltransferase, partial [Heterometrus spinifer] |
| Scorpion | ACY43225.1 | RNA helicase, partial [Heterometrus spinifer] |
| Scorpion | ACY43383.1 | ATPase, partial [Heterometrus spinifer] |
| Scorpion | ACY43586.1 | protein kinase, partial [Heterometrus spinifer] |
| Scorpion | ACY43641.1 | RNA helicase, partial [Heterometrus spinifer] |

| | | |
|----------|------------|--|
| Scorpion | ACY43697.1 | RNA polymerase subunit 2, partial [Heterometrus spinifer] |
| Scorpion | ACY43756.1 | RNA polymerase subunit 1, partial [Heterometrus spinifer] |
| Scorpion | ACY44496.1 | signal recognition particle, partial [Heterometrus spinifer] |
| Scorpion | ACY44554.1 | spliceosome-associated protein, partial [Heterometrus spinifer] |
| Scorpion | ACY44820.1 | clathrin heavy chain, partial [Heterometrus spinifer] |
| Scorpion | ACY45091.1 | acetylglucosaminyl-transferase, partial [Heterometrus spinifer] |
| Scorpion | ACY45410.1 | CDC5-related protein, partial [Heterometrus spinifer] |
| Scorpion | ACY45566.1 | F-box protein, partial [Heterometrus spinifer] |
| Scorpion | ACY45962.1 | methionine aminopeptidase, partial [Heterometrus spinifer] |
| Scorpion | ADY39499.1 | putative tropomyosin [Hottentotta judaicus] |
| Scorpion | ADY39528.1 | putative basic leucine zipper [Hottentotta judaicus] |
| Scorpion | ADY39535.1 | putative 60S Ribosomal protein L7-like, partial [Hottentotta judaicus] |
| Scorpion | ADY39541.1 | putative fructose 1,6-bisphosphate, partial [Hottentotta judaicus] |
| Scorpion | ADY39546.1 | putative basic leucine zipper [Hottentotta judaicus] |
| Scorpion | ADY39564.1 | putative sarco/endoplasmic reticulum Ca ²⁺ ATPase, partial [Hottentotta judaicus] |
| Scorpion | ADY39567.1 | hypothetical protein, partial [Hottentotta judaicus] |
| Scorpion | ADY39572.1 | putative N-acetylgalactosaminyl transferase [Hottentotta judaicus] |
| Scorpion | ADY39591.1 | hypothetical protein, partial [Hottentotta judaicus] |
| Scorpion | ADY39601.1 | putative DNA-binding protein, partial [Hottentotta judaicus] |
| Scorpion | ADY39603.1 | actin, partial [Hottentotta judaicus] |
| Scorpion | ADY39619.1 | putative 40S ribosomal protein 10S [Hottentotta judaicus] |
| Scorpion | ADY62662.1 | putative RNA-binding protein, partial [Hottentotta judaicus] |
| Scorpion | AEX09231.1 | translation elongation factor, partial [Pandinus cavimanus] |
| Scorpion | AGT39267.1 | 40S ribosomal protein, partial [Mesobuthus martensii] |
| Scorpion | AIX87780.1 | cellular protein AbCp-58, partial [Androctonus bicolor] |
| Scorpion | AIX87801.1 | cellular protein AbCp-85, partial [Androctonus bicolor] |
| Scorpion | AIX87802.1 | cellular protein AbCp-86, partial [Androctonus bicolor] |
| Scorpion | AMO02537.1 | astacin-like protein metallopeptidase 15, partial [Tityus serrulatus] |
| Scorpion | AMO02539.1 | astacin-like protein metallopeptidase 17, partial [Tityus serrulatus] |
| Scorpion | AMO02564.1 | serine protease 2, partial [Tityus serrulatus] |
| Scorpion | AMO02568.1 | serine protease 6 [Tityus serrulatus] |
| Scorpion | CAE54106.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAE54109.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAE54116.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAE54223.1 | chaperonin, partial [Mesobuthus eupeus] |
| Scorpion | CAE54232.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAE54244.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAE54247.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAE54250.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAE54257.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAE54263.1 | chaperonin, partial [Mesobuthus gibbosus] |
| Scorpion | CAX51428.1 | calmodulin-like protein, partial [Opisthacanthus cayaporum] |
| Scorpion | CDJ26709.1 | astacin-like metallopeptidase 13 protein, partial [Tityus serrulatus] |
| Scorpion | CDJ26713.1 | trypsin-like serine peptidase 3 protein, partial [Tityus serrulatus] |
| Scorpion | CDJ26730.1 | cathepsin L-like cysteine peptidase 1 protein, partial [Tityus serrulatus] |

| | | |
|----------|------------|--|
| Scorpion | JAG85159.1 | putative cysteine-rich venom protein [Tityus bahiensis] |
| Scorpion | JAG85162.1 | putative serpin-like proteinase inhibitor [Tityus bahiensis] |
| Scorpion | JAG85166.1 | putative cathepsin D-like protein [Tityus bahiensis] |
| Scorpion | JAG85167.1 | putative cathepsin B-like protein [Tityus bahiensis] |
| Scorpion | JAG85168.1 | putative venom serpin-like proteinase inhibitor [Tityus bahiensis] |
| Scorpion | JAG85169.1 | putative cathepsin F-like protein [Tityus bahiensis] |
| Scorpion | JAG85170.1 | putative angiotensin-converting enzyme, partial [Tityus bahiensis] |
| Scorpion | JAG85176.1 | putative cysteine proteinase inhibitor, partial [Tityus bahiensis] |
| Scorpion | JAG85178.1 | putative cysteine-rich protein, partial [Tityus bahiensis] |
| Scorpion | JAG85185.1 | putative sodium channel toxin [Tityus bahiensis] |
| Scorpion | JAG85195.1 | putative potassium channel toxin [Tityus bahiensis] |
| Scorpion | JAG85209.1 | putative chitinase, partial [Tityus bahiensis] |
| Scorpion | JAG85218.1 | putative phospholipase A2, partial [Tityus bahiensis] |
| Scorpion | JAG85223.1 | putative potassium channel toxin [Tityus bahiensis] |
| Scorpion | JAG85224.1 | putative hypotensin-like protein [Tityus bahiensis] |
| Scorpion | JAV47733.1 | ubiquitin-conjugating enzyme E2-17 kDa [Hadrurus spadix] |
| Scorpion | JAV47734.1 | ubiquitin-like protein 3 [Hadrurus spadix] |
| Scorpion | JAV47741.1 | tubulin alpha-1C chain [Hadrurus spadix] |
| Scorpion | JAV47747.1 | tropomyosin [Hadrurus spadix] |
| Scorpion | JAV47772.1 | eukaryotic initiation factor 4A-I [Hadrurus spadix] |
| Scorpion | JAV47835.1 | Protein transport protein Sec61 subunit alpha [Hadrurus spadix] |
| Scorpion | JAV47852.1 | 40S ribosomal protein S7 [Hadrurus spadix] |
| Scorpion | JAV47853.1 | 40S ribosomal protein S5 [Hadrurus spadix] |
| Scorpion | JAV47854.1 | 40S ribosomal protein S4 [Hadrurus spadix] |
| Scorpion | JAV47866.1 | 40S ribosomal protein S13 [Hadrurus spadix] |
| Scorpion | JAV47869.1 | 60S ribosomal protein L8 [Hadrurus spadix] |
| Scorpion | JAV47876.1 | 60S ribosomal protein L3 [Hadrurus spadix] |
| Scorpion | JAV47878.1 | 60S ribosomal protein L37a [Hadrurus spadix] |
| Scorpion | JAV47885.1 | 60S ribosomal protein L26 [Hadrurus spadix] |
| Scorpion | JAV47887.1 | 60S ribosomal protein L17 [Hadrurus spadix] |
| Scorpion | JAV47897.1 | RING-box protein 1A [Hadrurus spadix] |
| Scorpion | JAV47901.1 | ras-like GTP-binding protein Rho1 [Hadrurus spadix] |
| Scorpion | JAV47908.1 | ras-related protein Rap-1b [Hadrurus spadix] |
| Scorpion | JAV47911.1 | ribosome-attached membrane protein 4 [Hadrurus spadix] |
| Scorpion | JAV47916.1 | Ras-related protein Rab-2A [Hadrurus spadix] |
| Scorpion | JAV47922.1 | ras-related protein Rab-11B [Hadrurus spadix] |
| Scorpion | JAV47927.1 | pur-alpha [Hadrurus spadix] |
| Scorpion | JAV48002.1 | neurocalcin [Hadrurus spadix] |
| Scorpion | JAV48133.1 | guanine nucleotide-binding protein subunit beta-2 [Hadrurus spadix] |
| Scorpion | JAV48134.1 | guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1 [Hadrurus spadix] |
| Scorpion | JAV48146.1 | gamma-aminobutyric acid receptor-associated protein [Hadrurus spadix] |
| Scorpion | JAV48195.1 | translation elongation factor EF-1 alpha/Tu [Hadrurus spadix] |
| Scorpion | JAV48223.1 | cysteine-rich protein 1 [Hadrurus spadix] |
| Scorpion | JAV48232.1 | cofilin [Hadrurus spadix] |
| Scorpion | JAV48238.1 | clathrin-adaptor protein [Hadrurus spadix] |

| | | |
|----------|---------------|--|
| Scorpion | JAV48239.1 | clathrin heavy chain 1 [Hadrurus spadix] |
| Scorpion | JAV48370.1 | ADP ribosylation factor 102F [Hadrurus spadix] |
| Scorpion | JAV48374.1 | actin-related protein 3 [Hadrurus spadix] |
| Scorpion | JAV48375.1 | actin [Hadrurus spadix] |
| Scorpion | JAV48376.1 | actin [Hadrurus spadix] |
| Scorpion | JAW06971.1 | putative sodium channel toxin [Tityus serrulatus] |
| Scorpion | JAW06974.1 | putative potassium channel toxin [Tityus serrulatus] |
| Scorpion | JAW06976.1 | putative sodium channel toxin [Tityus serrulatus] |
| Scorpion | JAW06982.1 | putative metalloproteinase, partial [Tityus serrulatus] |
| Scorpion | JAW06995.1 | putative cysteine-rich protein [Tityus serrulatus] |
| Scorpion | JAW06999.1 | putative antimicrobial peptide [Tityus serrulatus] |
| Scorpion | JAW07004.1 | putative sodium channel toxin [Tityus serrulatus] |
| Scorpion | JAW07010.1 | putative potassium channel toxin [Tityus serrulatus] |
| Scorpion | JAW07014.1 | putative anticoagulant peptide [Tityus serrulatus] |
| Scorpion | JAW07015.1 | putative chitinase, partial [Tityus serrulatus] |
| Scorpion | JAW07016.1 | putative sodium channel toxin [Tityus serrulatus] |
| Scorpion | JAW07018.1 | putative cathepsin B like, partial [Tityus serrulatus] |
| Scorpion | JAW07019.1 | putative phospholipase A2 [Tityus serrulatus] |
| Scorpion | JAW07022.1 | putative antimicrobial peptide [Tityus serrulatus] |
| Scorpion | JAW07023.1 | putative phospholipase A2, partial [Tityus serrulatus] |
| Scorpion | JAW07028.1 | putative metalloprotease, partial [Tityus serrulatus] |
| Scorpion | JAW07031.1 | putative cysteine-rich protein [Tityus serrulatus] |
| Scorpion | JAW07032.1 | putative cysteine-rich protein [Tityus serrulatus] |
| Scorpion | JAW07033.1 | putative angiotensin converting enzyme (ACE-like) [Tityus serrulatus] |
| Scorpion | JAW07034.1 | putative cysteine-rich protein [Tityus serrulatus] |
| Scorpion | JAW07036.1 | cysteine-rich perotein [Tityus serrulatus] |
| Scorpion | JAW07037.1 | putative potassium channel toxin [Tityus serrulatus] |
| Scorpion | JAW07038.1 | putative hypotensin [Tityus serrulatus] |
| Scorpion | JAW07043.1 | putative antimicrobial peptide [Tityus serrulatus] |
| Scorpion | JAW07044.1 | putative antimicrobial peptide, partial [Tityus serrulatus] |
| Scorpion | sp POC175.1 | RecName: Full=Potassium channel toxin epsilon-KTx 1.2; AltName: Full=Ts12; AltName: Full=TsPep2; Flags: Precursor |
| Scorpion | sp P56219.2 | RecName: Full=Potassium channel toxin alpha-KTx 4.2; AltName: Full=Neurotoxin Ts-kappa; Short=TsK; Short=TsKappa; AltName: Full=Ts9; Flags: Precursor |
| Scorpion | sp P84190.1 | RecName: Full=Hypotensin-2; AltName: Full=Hypotensin II; Short=TsHpt-II; Contains: RecName: Full=Hypotensin-4; AltName: Full=Hypotensin IV; Short=TsHpt-IV |
| Scorpion | tr A0A0U1S4Q6 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1S4Q9 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1S8A6 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1SCE8 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1SCF4 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1SCG3 | Uncharacterized protein OS=Isometrus maculatus OX=497827 PE=2 SV=1 |

| | | |
|----------|---------------|--|
| Scorpion | tr A0A0U1SKC7 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1SPK7 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1TYE4 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1TYG3 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1TYI8 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1TYJ5 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1U016 | Uncharacterized protein OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A0U1U053 | Uncharacterized protein (Fragment) OS=Isometrus maculatus OX=497827 PE=2 SV=1 |
| Scorpion | tr A0A1E1WVU8 | Putative atrial natriuretic peptide-converting enzyme-like protein OS=Tityus obscurus OX=1221240 PE=4 SV=1 |
| Scorpion | tr A0A1E1WVW7 | Profilin OS=Tityus obscurus OX=1221240 PE=3 SV=1 |
| Scorpion | tr A0A1E1WWB5 | Putative serine proteinase (Fragment) OS=Tityus obscurus OX=1221240 PE=4 SV=1 |
| Scorpion | tr A0A1U9WSY8 | Histone 3a (Fragment) OS=Tityus bahiensis OX=50343 GN=H3a PE=3 SV=1 |
| Scorpion | tr A0A2I9LNI8 | Actin OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNP4 | Alpha-enolase OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNP6 | ADP Ribosylation Factor OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNS0 | Voltage-dependent anion-selective channel protein 2 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNS3 | ATP Synthase OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNT3 | ATP/ADP Translocase OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNT6 | Calglandulin OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNV2 | Clatherine heavy chain OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNV6 | Calcium transporting ATPase OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNV8 | Coatomer subunit beta OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNW1 | Collagen type IV alpha binding protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNY1 | DnaJ-like protein subfamily B member 4 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNY8 | Chloride intracellular channel exc-4 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNZ0 | Elongation factor 2 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNZ3 | DnaJ-like protein subfamily OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LNZ7 | Filamin-A OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP07 | Enhancer binding protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP13 | Glycosyltransferase OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP24 | Hexokinase-1 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP25 | Heat shock 70 kDa protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP31 | Junctophilin-1 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP32 | Headcase protein-like protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP47 | Interferon regulatory factor 2-binding protein-like A OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP60 | Basic leucine zipper and W2 domain-containing protein 2 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |

| | | |
|----------|---------------|--|
| Scorpion | tr A0A2I9LP67 | MAP kinase-interacting serine/threonine-protein kinase 1 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP84 | Myocyte-specific enhancer factor 2 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP92 | Myosin heavy chain OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LP96 | Sodium/potassium-transporting ATPase subunit alpha-B OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPD2 | Mesencephalic astrocyte-derived neurotrophic factor-like protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPE8 | Protein disulfide-isomerase A3 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPF2 | Plastin-2 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPF5 | Peroxiredoxin 1 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPG4 | Blo t profilin allergen OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPH4 | 26S protease regulatory subunit 10B OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPH7 | Profilin allergen OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPI0 | 26S protease regulatory subunit 8 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPI1 | Rab GDP dissociation inhibitor beta OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPI6 | Ras related protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPJ7 | 60S ribosomal protein L10a OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPJ8 | 60S Ribosomal Protein L13 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPK0 | 60S ribosomal protein L11 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPK2 | 60S ribosomal protein L17 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPK4 | 60S ribosomal protein L21 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPK6 | 60S ribosomal protein L13a OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPL7 | 60S Ribosomal Protein L3 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPM1 | 40S ribosomal protein S16 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPM3 | 60S ribosomal protein L7 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPN3 | 40S Ribosomal Protein S19 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPP0 | 40S ribosomal protein S10 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPP5 | 40S ribosomal protein S4 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPQ0 | Protein slowmo OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPQ5 | Protein transport protein Sec23A OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPQ9 | Surfeit OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPR1 | S-adenosylmethionine synthase OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPR4 | Signal sequence receptor OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPR9 | Transcriptional activator protein Pur-beta OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPS2 | Uncharacterized protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPS7 | T complex protein subunit 1 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPS8 | Protein Tob1 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPT3 | Transitional endoplasmic reticulum ATPase OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPT5 | Translationally controlled tumor protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPT6 | Translation initiation factor 5A OS=Centruroides hentzi OX=88313 PE=4 SV=1 |

| | | |
|----------|----------------|---|
| Scorpion | tr A0A2I9LPT9 | Eukaryotic translation initiation factor 5 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPU1 | Transmembrane 9 superfamily member 3 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPU8 | Tumor suppressor candidate 3 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPV3 | Translocon-associated protein subunit gamma OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPV9 | Vigilin OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPW1 | Tubulin beta chain OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPW7 | Tubulin, alpha 1c OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPW8 | Ubiquitin-conjugating enzyme E2-17 kDa OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPX4 | Vesicle trafficking protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr A0A2I9LPZ8 | V-type proton ATPase catalytic subunit A OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Scorpion | tr C9X4H4 | Uncharacterized protein (Fragment) OS=Tityus discrepans OX=57059 PE=2 SV=1 |
| Scorpion | tr D0UJ54 | G protein-coupled receptor kinase (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0UJJ9 | Myosin (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0UK54 | Arg methyltransferase (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0UKF8 | RNA helicase (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0UKW5 | ATPase (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0ULB4 | Histone deacetylase (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0ULM4 | RNA helicase (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0ULS9 | RNA polymerase subunit 2 (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0UPD4 | Syntaxin (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0UQ03 | Clathrin heavy chain (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr D0UQL8 | Clathrin coat assembly protein (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | tr Q5SER5 | Histone H3 (Fragment) OS=Taurulus bubalis OX=61643 PE=3 SV=1 |
| Scorpion | tr Q9BNE0 | RNA polymerase II largest subunit (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=2 |
| Scorpion | tr Q9BNV0 | Elongation factor 1-alpha (Fragment) OS=Hadrurus arizonensis OX=88316 PE=2 SV=1 |
| Scorpion | XP_023209351.1 | serine/threonine-protein phosphatase 4 catalytic subunit B [Centruroides sculpturatus] |
| Scorpion | XP_023209365.1 | leucine-rich repeat-containing protein 49-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023209388.1 | transmembrane protein 151B-like [Centruroides sculpturatus] |
| Scorpion | XP_023209409.1 | plasma membrane calcium-transporting ATPase 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023209437.1 | LIM domain transcription factor LMO4-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023209464.1 | tyrosine-protein phosphatase Lar-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023209472.1 | tyrosine-protein phosphatase Lar-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023209476.1 | tyrosine-protein phosphatase Lar-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023209484.1 | tyrosine-protein phosphatase Lar-like isoform X4 [Centruroides sculpturatus] |

Scorpion XP_023209492.1 tyrosine-protein phosphatase Lar-like isoform X5 [Centruroides sculpturatus]

Scorpion XP_023209529.1 E3 ubiquitin-protein ligase RNF185-like [Centruroides sculpturatus]

Scorpion XP_023209574.1 signal recognition particle 9 kDa protein-like [Centruroides sculpturatus]

Scorpion XP_023209576.1 LHFPL tetraspan subfamily member 6 protein-like [Centruroides sculpturatus]

Scorpion XP_023209602.1 bone morphogenetic protein 4-like [Centruroides sculpturatus]

Scorpion XP_023209607.1 protein king tubby 1-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023209617.1 A-kinase anchor protein 17A-like [Centruroides sculpturatus]

Scorpion XP_023209638.1 cleavage and polyadenylation specificity factor subunit 3-like [Centruroides sculpturatus]

Scorpion XP_023209643.1 ADP-ribosylation factor 2 [Centruroides sculpturatus]

Scorpion XP_023209655.1 uncharacterized protein LOC111612647 [Centruroides sculpturatus]

Scorpion XP_023209656.1 ubiquitin-conjugating enzyme E2 A-like [Centruroides sculpturatus]

Scorpion XP_023209716.1 COMM domain-containing protein 2-like [Centruroides sculpturatus]

Scorpion XP_023209719.1 gonadotropin-releasing hormone receptor-like, partial [Centruroides sculpturatus]

Scorpion XP_023209730.1 AP-2 complex subunit mu [Centruroides sculpturatus]

Scorpion XP_023209769.1 sodium-dependent noradrenaline transporter-like [Centruroides sculpturatus]

Scorpion XP_023209804.1 actin-related protein 2/3 complex subunit 5-like [Centruroides sculpturatus]

Scorpion XP_023209811.1 histone-binding protein RBBP7, partial [Centruroides sculpturatus]

Scorpion XP_023209824.1 transcription factor Sox-8-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023209844.1 ATP-dependent RNA helicase DDX3Y-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023209846.1 ATP-dependent RNA helicase DDX3Y-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023209853.1 inactive tyrosine-protein kinase transmembrane receptor ROR1-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023209870.1 OTU domain-containing protein 5-B-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023209872.1 dnaJ homolog subfamily A member 1-like [Centruroides sculpturatus]

Scorpion XP_023209891.1 ras-specific guanine nucleotide-releasing factor 1-like [Centruroides sculpturatus]

Scorpion XP_023209921.1 FACT complex subunit SPT16-like [Centruroides sculpturatus]

Scorpion XP_023209939.1 AP-1 complex subunit beta-1-like [Centruroides sculpturatus]

Scorpion XP_023209963.1 glycogen synthase kinase-3 beta-like, partial [Centruroides sculpturatus]

Scorpion XP_023209964.1 LIX1-like protein [Centruroides sculpturatus]

Scorpion XP_023209968.1 ankyrin repeat domain-containing protein 54-like [Centruroides sculpturatus]

Scorpion XP_023209985.1 clarin-2-like [Centruroides sculpturatus]

Scorpion XP_023210052.1 uncharacterized protein ZK1073.1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023210059.1 zinc finger SWIM domain-containing protein 8-like [Centruroides sculpturatus]

Scorpion XP_023210141.1 GTPase KRas-like [Centruroides sculpturatus]

Scorpion XP_023210178.1 zinc finger MIZ domain-containing protein 1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023210180.1 zinc finger MIZ domain-containing protein 1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023210195.1 polycomb protein SCM1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023210196.1 polycomb protein SCM1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023210248.1 phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase TPTE2-like [Centruroides sculpturatus]

Scorpion XP_023210271.1 myosin heavy chain, muscle-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023210272.1 myosin heavy chain, muscle-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023210273.1 myosin heavy chain, muscle-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023210274.1 myosin heavy chain, muscle-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023210276.1 myosin heavy chain, muscle-like isoform X6 [Centruroides sculpturatus]

Scorpion XP_023210277.1 myosin heavy chain, muscle-like isoform X7 [Centruroides sculpturatus]

Scorpion XP_023210281.1 myosin heavy chain, muscle-like isoform X10 [Centruroides sculpturatus]

Scorpion XP_023210282.1 myosin heavy chain, muscle-like isoform X11 [Centruroides sculpturatus]

Scorpion XP_023210284.1 myosin heavy chain, muscle-like isoform X13 [Centruroides sculpturatus]

Scorpion XP_023210286.1 myosin heavy chain, muscle-like isoform X15 [Centruroides sculpturatus]

Scorpion XP_023210290.1 myosin heavy chain, muscle-like isoform X18 [Centruroides sculpturatus]

Scorpion XP_023210345.1 transmembrane protein 115-like [Centruroides sculpturatus]

Scorpion XP_023210382.1 Usher syndrome type-1G protein homolog [Centruroides sculpturatus]

Scorpion XP_023210398.1 uncharacterized protein LOC111613282 [Centruroides sculpturatus]

Scorpion XP_023210399.1 UPF0472 protein C16orf72 homolog [Centruroides sculpturatus]

Scorpion XP_023210439.1 membrane magnesium transporter 1-like [Centruroides sculpturatus]

Scorpion XP_023210445.1 vang-like protein 2 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023210447.1 vang-like protein 2-B isoform X2 [Centruroides sculpturatus]

Scorpion XP_023210448.1 vang-like protein 2-B [Centruroides sculpturatus]

Scorpion XP_023210450.1 solute carrier family 41 member 1-like [Centruroides sculpturatus]

Scorpion XP_023210469.1 BMP and activin membrane-bound inhibitor homolog [Centruroides sculpturatus]

Scorpion XP_023210477.1 aryl hydrocarbon receptor-like, partial [Centruroides sculpturatus]

Scorpion XP_023210661.1 disks large homolog 1-like, partial [Centruroides sculpturatus]

Scorpion XP_023210759.1 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase-like [Centruroides sculpturatus]

Scorpion XP_023210764.1 protein RCC2 homolog [Centruroides sculpturatus]

Scorpion XP_023210803.1 serine/threonine-protein kinase MRCK alpha-like [Centruroides sculpturatus]

Scorpion XP_023210839.1 uncharacterized protein LOC111613730 [Centruroides sculpturatus]

Scorpion XP_023210844.1 GTP-binding protein Di-Ras2-like [Centruroides sculpturatus]

Scorpion XP_023210850.1 uncharacterized protein LOC111613739 isoform X2 [Centruroides sculpturatus]

Scorpion XP_023210884.1 polyglutamine-binding protein 1-like [Centruroides sculpturatus]

Scorpion XP_023210939.1 transcription factor 4-like [Centruroides sculpturatus]

Scorpion XP_023210950.1 ATP-binding cassette sub-family F member 1-like [Centruroides sculpturatus]

Scorpion XP_023210989.1 dipeptidyl peptidase 9-like [Centruroides sculpturatus]

Scorpion XP_023211057.1 flotillin-2-like [Centruroides sculpturatus]

Scorpion XP_023211064.1 flotillin-2-like [Centruroides sculpturatus]

Scorpion XP_023211099.1 vinculin-like [Centruroides sculpturatus]

Scorpion XP_023211161.1 flotillin-2-like [Centruroides sculpturatus]

Scorpion XP_023211184.1 uncharacterized protein LOC111614034 [Centruroides sculpturatus]

Scorpion XP_023211194.1 SPRY domain-containing SOCS box protein 3-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023211222.1 WD repeat and FYVE domain-containing protein 3-like, partial [Centruroides sculpturatus]

Scorpion XP_023211223.1 guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1 [Centruroides sculpturatus]

| | | |
|----------|----------------|---|
| Scorpion | XP_023211262.1 | calcium/calmodulin-dependent protein kinase type 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023211275.1 | myosin heavy chain, non-muscle-like [Centruroides sculpturatus] |
| Scorpion | XP_023211337.1 | polypeptide N-acetylgalactosaminyltransferase 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023211338.1 | histone H3-like [Centruroides sculpturatus] |
| Scorpion | XP_023211431.1 | zinc finger protein 585B-like [Centruroides sculpturatus] |
| Scorpion | XP_023211454.1 | casein kinase I-like [Centruroides sculpturatus] |
| Scorpion | XP_023211493.1 | cytoplasmic polyadenylation element-binding protein 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023211496.1 | 60S ribosomal protein L24-like [Centruroides sculpturatus] |
| Scorpion | XP_023211528.1 | cullin-associated NEDD8-dissociated protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023211595.1 | monoglyceride lipase-like [Centruroides sculpturatus] |
| Scorpion | XP_023211596.1 | forkhead box protein C2-like [Centruroides sculpturatus] |
| Scorpion | XP_023211610.1 | pleiotropic regulator 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023211640.1 | 40S ribosomal protein S19-like [Centruroides sculpturatus] |
| Scorpion | XP_023211642.1 | ER degradation-enhancing alpha-mannosidase-like protein 3 [Centruroides sculpturatus] |
| Scorpion | XP_023211721.1 | UPF0769 protein C21orf59 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023211818.1 | nuclear receptor subfamily 2 group C member 1-A-like [Centruroides sculpturatus] |
| Scorpion | XP_023211888.1 | phosphatidylinositol 5-phosphate 4-kinase type-2 alpha-like [Centruroides sculpturatus] |
| Scorpion | XP_023211900.1 | uncharacterized protein LOC111614766, partial [Centruroides sculpturatus] |
| Scorpion | XP_023211952.1 | dnaJ homolog shv-like [Centruroides sculpturatus] |
| Scorpion | XP_023211963.1 | sorting nexin-17-like [Centruroides sculpturatus] |
| Scorpion | XP_023211987.1 | forkhead box protein F1-like [Centruroides sculpturatus] |
| Scorpion | XP_023212064.1 | NEDD8-conjugating enzyme Ubc12-like [Centruroides sculpturatus] |
| Scorpion | XP_023212070.1 | ubiquitin thioesterase OTUB1-like [Centruroides sculpturatus] |
| Scorpion | XP_023212108.1 | enolase-like [Centruroides sculpturatus] |
| Scorpion | XP_023212139.1 | cyclin-dependent kinase-like 4 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023212142.1 | cyclin-dependent kinase-like 4 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023212143.1 | U4/U6 small nuclear ribonucleoprotein Prp4-like [Centruroides sculpturatus] |
| Scorpion | XP_023212169.1 | protein limb expression 1 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023212267.1 | failed axon connections-like [Centruroides sculpturatus] |
| Scorpion | XP_023212273.1 | etoposide-induced protein 2.4 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023212275.1 | 60S ribosomal protein L28-like [Centruroides sculpturatus] |
| Scorpion | XP_023212317.1 | TOG array regulator of axonemal microtubules protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023212390.1 | tubulin alpha-1A chain isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023212555.1 | TBC1 domain family member 1-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023212585.1 | U6 snRNA-associated Sm-like protein LSm4 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023212589.1 | 60S ribosomal protein L7-like [Centruroides sculpturatus] |
| Scorpion | XP_023212609.1 | zinc finger homeobox protein 4-like [Centruroides sculpturatus] |
| Scorpion | XP_023212663.1 | coatomer subunit zeta-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023212716.1 | GTP-binding protein SAR1b-like [Centruroides sculpturatus] |
| Scorpion | XP_023212722.1 | octopamine receptor beta-2R-like [Centruroides sculpturatus] |

Scorpion XP_023212739.1 protein transport protein Sec24C-like [Centruroides sculpturatus]

Scorpion XP_023212758.1 small ubiquitin-related modifier 3-like [Centruroides sculpturatus]

Scorpion XP_023212818.1 DNA polymerase zeta catalytic subunit-like [Centruroides sculpturatus]

Scorpion XP_023212825.1 ras-related protein Rap-1A [Centruroides sculpturatus]

Scorpion XP_023212836.1 ras-associated and pleckstrin homology domains-containing protein 1-like [Centruroides sculpturatus]

Scorpion XP_023213013.1 lysosomal thioesterase PPT2-A-like, partial [Centruroides sculpturatus]

Scorpion XP_023213014.1 U2 snRNP-associated SURP motif-containing protein-like [Centruroides sculpturatus]

Scorpion XP_023213017.1 steroid hormone receptor ERR1-like [Centruroides sculpturatus]

Scorpion XP_023213034.1 uncharacterized protein LOC111615835 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023213108.1 transforming growth factor beta receptor type 3-like [Centruroides sculpturatus]

Scorpion XP_023213206.1 gastrula zinc finger protein XICGF57.1-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023213209.1 protein slowmo-like [Centruroides sculpturatus]

Scorpion XP_023213232.1 protein turtle-like [Centruroides sculpturatus]

Scorpion XP_023213245.1 transient-receptor-potential-like protein [Centruroides sculpturatus]

Scorpion XP_023213395.1 26S proteasome non-ATPase regulatory subunit 6-like [Centruroides sculpturatus]

Scorpion XP_023213423.1 DNA-directed RNA polymerase III subunit RPC10-like [Centruroides sculpturatus]

Scorpion XP_023213451.1 zinc finger FYVE domain-containing protein 1-like [Centruroides sculpturatus]

Scorpion XP_023213454.1 innexin inx2-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023213532.1 myb/SANT-like DNA-binding domain-containing protein 4 [Centruroides sculpturatus]

Scorpion XP_023213534.1 uncharacterized protein LOC111616346 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023213535.1 uncharacterized protein LOC111616346 isoform X2 [Centruroides sculpturatus]

Scorpion XP_023213569.1 ankyrin repeat and SOCS box protein 3-like [Centruroides sculpturatus]

Scorpion XP_023213570.1 chromodomain-helicase-DNA-binding protein 8-like, partial [Centruroides sculpturatus]

Scorpion XP_023213576.1 early growth response protein 1-B-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023213582.1 early growth response protein 1-B-like isoform X5 [Centruroides sculpturatus]

Scorpion XP_023213596.1 phosphatidylinositol 5-phosphate 4-kinase type-2 alpha-like [Centruroides sculpturatus]

Scorpion XP_023213617.1 DNA-directed RNA polymerases I, II, and III subunit RPABC1-like, partial [Centruroides sculpturatus]

Scorpion XP_023213621.1 ryanodine receptor-like [Centruroides sculpturatus]

Scorpion XP_023213673.1 citrate synthase, mitochondrial-like [Centruroides sculpturatus]

Scorpion XP_023213706.1 neurexin-1-like [Centruroides sculpturatus]

Scorpion XP_023213735.1 neurobeachin-like [Centruroides sculpturatus]

Scorpion XP_023213835.1 NAD(P) transhydrogenase, mitochondrial-like [Centruroides sculpturatus]

Scorpion XP_023213841.1 octopamine receptor-like [Centruroides sculpturatus]

Scorpion XP_023213862.1 activin receptor type-1C-like [Centruroides sculpturatus]

Scorpion XP_023213865.1 DNA-directed RNA polymerase II subunit RPB4 [Centruroides sculpturatus]

Scorpion XP_023213882.1 ankyrin repeat domain-containing protein 50-like [Centruroides sculpturatus]

Scorpion XP_023213884.1 DNA-binding protein inhibitor ID-2-like [Centruroides sculpturatus]

Scorpion XP_023213893.1 tetratricopeptide repeat protein 28-like, partial [Centruroides sculpturatus]

Scorpion XP_023213916.1 LOW QUALITY PROTEIN: histone acetyltransferase KAT5-like, partial [Centruroides sculpturatus]

Scorpion XP_023213951.1 dystonin-like [Centruroides sculpturatus]

Scorpion XP_023213983.1 serine/threonine-protein phosphatase 2B catalytic subunit-like, partial [Centruroides sculpturatus]

Scorpion XP_023214001.1 Ca(2+)/calmodulin-responsive adenylate cyclase-like [Centruroides sculpturatus]

Scorpion XP_023214002.1 Ca(2+)/calmodulin-responsive adenylate cyclase-like [Centruroides sculpturatus]

Scorpion XP_023214020.1 ADP-ribosylation factor-like protein 8B-A [Centruroides sculpturatus]

Scorpion XP_023214037.1 uncharacterized protein LOC111616910, partial [Centruroides sculpturatus]

Scorpion XP_023214068.1 T-cell leukemia homeobox protein 1-like [Centruroides sculpturatus]

Scorpion XP_023214077.1 formin-like protein CG32138 [Centruroides sculpturatus]

Scorpion XP_023214095.1 protein tracheless-like, partial [Centruroides sculpturatus]

Scorpion XP_023214098.1 BTB/POZ domain-containing protein KCTD3-like, partial [Centruroides sculpturatus]

Scorpion XP_023214123.1 uncharacterized protein LOC111617010, partial [Centruroides sculpturatus]

Scorpion XP_023214139.1 26S proteasome regulatory subunit 4 [Centruroides sculpturatus]

Scorpion XP_023214165.1 10 kDa heat shock protein, mitochondrial-like [Centruroides sculpturatus]

Scorpion XP_023214225.1 frizzled-5-like [Centruroides sculpturatus]

Scorpion XP_023214236.1 LOW QUALITY PROTEIN: casein kinase I-like [Centruroides sculpturatus]

Scorpion XP_023214248.1 (E3-independent) E2 ubiquitin-conjugating enzyme-like, partial [Centruroides sculpturatus]

Scorpion XP_023214252.1 transcription factor Sox-9-B-like [Centruroides sculpturatus]

Scorpion XP_023214263.1 cleavage and polyadenylation specificity factor subunit 4-like [Centruroides sculpturatus]

Scorpion XP_023214272.1 60S ribosomal protein L9-like [Centruroides sculpturatus]

Scorpion XP_023214428.1 E3 ubiquitin-protein ligase MIB2-like [Centruroides sculpturatus]

Scorpion XP_023214452.1 erythroid differentiation-related factor 1-like [Centruroides sculpturatus]

Scorpion XP_023214481.1 DNA helicase INO80-like [Centruroides sculpturatus]

Scorpion XP_023214488.1 sarcolemmal membrane-associated protein-like, partial [Centruroides sculpturatus]

Scorpion XP_023214596.1 chaoptin-like, partial [Centruroides sculpturatus]

Scorpion XP_023214607.1 vacuolar protein sorting-associated protein 29 isoform X2 [Centruroides sculpturatus]

Scorpion XP_023214620.1 coatomer subunit beta'-like, partial [Centruroides sculpturatus]

Scorpion XP_023214627.1 uncharacterized protein LOC111617577, partial [Centruroides sculpturatus]

Scorpion XP_023214634.1 signal peptidase complex catalytic subunit SEC11A-like [Centruroides sculpturatus]

Scorpion XP_023214667.1 ATP-binding cassette sub-family A member 2-like [Centruroides sculpturatus]

Scorpion XP_023214821.1 CCAAT/enhancer-binding protein epsilon-like [Centruroides sculpturatus]

Scorpion XP_023214901.1 calcium-transporting ATPase type 2C member 1-like [Centruroides sculpturatus]

Scorpion XP_023215002.1 retinoblastoma-like protein 2, partial [Centruroides sculpturatus]

Scorpion XP_023215004.1 ras-related protein Rab-2-like [Centruroides sculpturatus]

Scorpion XP_023215009.1 uncharacterized protein LOC111617893 [Centruroides sculpturatus]

Scorpion XP_023215038.1 T-box transcription factor TBX20-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023215039.1 T-box transcription factor TBX20-like isoform X2 [Centruroides sculpturatus]

| | | |
|----------|----------------|--|
| Scorpion | XP_023215103.1 | toll-like receptor 6 [Centruroides sculpturatus] |
| Scorpion | XP_023215106.1 | trafficking protein particle complex subunit 3-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023215176.1 | proteasome subunit alpha type-5-like [Centruroides sculpturatus] |
| Scorpion | XP_023215182.1 | F-box/LRR-repeat protein 14-like [Centruroides sculpturatus] |
| Scorpion | XP_023215186.1 | protein Wnt-5b-like [Centruroides sculpturatus] |
| Scorpion | XP_023215188.1 | integrator complex subunit 4-like [Centruroides sculpturatus] |
| Scorpion | XP_023215236.1 | protein disulfide-isomerase A6 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023215241.1 | E3 ubiquitin-protein ligase ZNRF1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023215253.1 | uncharacterized protein LOC111618058 isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023215255.1 | uncharacterized protein LOC111618058 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023215285.1 | RING finger and SPRY domain-containing protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023215302.1 | uncharacterized protein LOC111618089 [Centruroides sculpturatus] |
| Scorpion | XP_023215308.1 | homeobox protein cut-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023215333.1 | solute carrier family 35 member E1 homolog isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023215338.1 | chondroitin sulfate N-acetylgalactosaminyltransferase 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023215380.1 | myocyte-specific enhancer factor 2C-like [Centruroides sculpturatus] |
| Scorpion | XP_023215447.1 | N-alpha-acetyltransferase 15, NatA auxiliary subunit-like [Centruroides sculpturatus] |
| Scorpion | XP_023215466.1 | tubulin alpha-1A chain-like [Centruroides sculpturatus] |
| Scorpion | XP_023215471.1 | septin-4-like [Centruroides sculpturatus] |
| Scorpion | XP_023215472.1 | septin-1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023215509.1 | LOW QUALITY PROTEIN: transportin-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023215567.1 | 40S ribosomal protein S6-like [Centruroides sculpturatus] |
| Scorpion | XP_023215594.1 | neurofibromin-like [Centruroides sculpturatus] |
| Scorpion | XP_023215609.1 | myosin-11-like [Centruroides sculpturatus] |
| Scorpion | XP_023215636.1 | adenylate kinase isoenzyme 1-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023215667.1 | leucine-rich repeat-containing protein 24-like [Centruroides sculpturatus] |
| Scorpion | XP_023215693.1 | sterile alpha and TIR motif-containing protein 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023215694.1 | sterile alpha and TIR motif-containing protein 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023215715.1 | protein FAM117B-like [Centruroides sculpturatus] |
| Scorpion | XP_023215720.1 | RISC-loading complex subunit tarbp2-like [Centruroides sculpturatus] |
| Scorpion | XP_023215722.1 | estradiol 17-beta-dehydrogenase 8-like [Centruroides sculpturatus] |
| Scorpion | XP_023215742.1 | metabotropic glutamate receptor 3-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023215757.1 | Kv channel-interacting protein 4-like [Centruroides sculpturatus] |
| Scorpion | XP_023215771.1 | homeobox protein Hox-B4a-like [Centruroides sculpturatus] |
| Scorpion | XP_023215773.1 | F-box/WD repeat-containing protein 1A-like [Centruroides sculpturatus] |
| Scorpion | XP_023215776.1 | uncharacterized protein LOC111618466 [Centruroides sculpturatus] |
| Scorpion | XP_023215826.1 | UPF0428 protein CXorf56 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023215829.1 | protein Dr1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023215858.1 | tubulin alpha-1C chain-like [Centruroides sculpturatus] |

| | | |
|----------|----------------|--|
| Scorpion | XP_023215859.1 | surfeit locus protein 4 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023215915.1 | UV excision repair protein RAD23 homolog B-like [Centruroides sculpturatus] |
| Scorpion | XP_023215922.1 | calcium-independent protein kinase C-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023215924.1 | calcium-independent protein kinase C-like isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023215944.1 | transmembrane protein 39A-like [Centruroides sculpturatus] |
| Scorpion | XP_023215954.1 | microtubule-associated proteins 1A/1B light chain 3B-like [Centruroides sculpturatus] |
| Scorpion | XP_023215970.1 | casein kinase II subunit beta [Centruroides sculpturatus] |
| Scorpion | XP_023215989.1 | vacuolar protein sorting-associated protein 16 homolog isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023216015.1 | uncharacterized protein LOC111618684 isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023216080.1 | protein sprint-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023216094.1 | protein FAM49B-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023216102.1 | protein FAM49B-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023216148.1 | BTB/POZ domain-containing protein 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023216153.1 | engulfment and cell motility protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023216231.1 | recombining binding protein suppressor of hairless-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023216238.1 | rho GTPase-activating protein 100F-like [Centruroides sculpturatus] |
| Scorpion | XP_023216301.1 | GTPase-activating Rap/Ran-GAP domain-like protein 3 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023216308.1 | uncharacterized protein LOC111618912 [Centruroides sculpturatus] |
| Scorpion | XP_023216316.1 | homeobox protein six1-like [Centruroides sculpturatus] |
| Scorpion | XP_023216343.1 | GTP-binding protein ypt2-like [Centruroides sculpturatus] |
| Scorpion | XP_023216365.1 | protein trachealess-like [Centruroides sculpturatus] |
| Scorpion | XP_023216401.1 | 60S ribosomal protein L8 [Centruroides sculpturatus] |
| Scorpion | XP_023216461.1 | HEAT repeat-containing protein 5B-like [Centruroides sculpturatus] |
| Scorpion | XP_023216464.1 | LHFPL tetraspan subfamily member 2a protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023216473.1 | RING finger protein 11-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023216492.1 | catenin delta-2-like [Centruroides sculpturatus] |
| Scorpion | XP_023216559.1 | vesicle-trafficking protein SEC22b-like [Centruroides sculpturatus] |
| Scorpion | XP_023216569.1 | potassium channel subfamily T member 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023216585.1 | histone acetyltransferase KAT2B-like [Centruroides sculpturatus] |
| Scorpion | XP_023216596.1 | prohibitin-2-like [Centruroides sculpturatus] |
| Scorpion | XP_023216611.1 | ETS homologous factor-like isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023216651.1 | integrator complex subunit 2-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023216742.1 | lysoplasmalogenase-like protein TMEM86A isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023216749.1 | sec1 family domain-containing protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023216761.1 | protein odd-skipped-related 2-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023216825.1 | protein tyrosine phosphatase type IVA 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023216879.1 | T-cell leukemia homeobox protein 3-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023216880.1 | T-cell leukemia homeobox protein 3-like isoform X2 [Centruroides sculpturatus] |

| | | |
|----------|----------------|--|
| Scorpion | XP_023216899.1 | DNA methyltransferase 1-associated protein 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023216900.1 | DNA methyltransferase 1-associated protein 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023216972.1 | forkhead box protein O-like [Centruroides sculpturatus] |
| Scorpion | XP_023217023.1 | ubiquitin carboxyl-terminal hydrolase 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023217031.1 | rap1 GTPase-activating protein 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023217041.1 | rap1 GTPase-activating protein 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023217047.1 | F-box only protein 33-like [Centruroides sculpturatus] |
| Scorpion | XP_023217085.1 | GTP-binding protein SAR1-like [Centruroides sculpturatus] |
| Scorpion | XP_023217102.1 | zinc finger MIZ domain-containing protein 1-like isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023217103.1 | zinc finger MIZ domain-containing protein 1-like isoform X5 [Centruroides sculpturatus] |
| Scorpion | XP_023217167.1 | uncharacterized protein LOC111619628 [Centruroides sculpturatus] |
| Scorpion | XP_023217184.1 | homeodomain-only protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023217215.1 | MAP kinase-activating death domain protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023217221.1 | developmentally-regulated GTP-binding protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023217243.1 | pleiotropic regulator 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023217334.1 | ras-related protein Rab-14 [Centruroides sculpturatus] |
| Scorpion | XP_023217361.1 | probable actin-related protein 2/3 complex subunit 2 [Centruroides sculpturatus] |
| Scorpion | XP_023217447.1 | probable phospholipid-transporting ATPase IM [Centruroides sculpturatus] |
| Scorpion | XP_023217546.1 | proteasome subunit beta type-3-like [Centruroides sculpturatus] |
| Scorpion | XP_023217552.1 | uncharacterized protein LOC111619957 [Centruroides sculpturatus] |
| Scorpion | XP_023217558.1 | UDP-glucose 4-epimerase-like [Centruroides sculpturatus] |
| Scorpion | XP_023217569.1 | GTP-binding protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023217593.1 | N-chimaerin-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023217637.1 | COP9 signalosome complex subunit 7b-like [Centruroides sculpturatus] |
| Scorpion | XP_023217640.1 | kelch-like protein diablo [Centruroides sculpturatus] |
| Scorpion | XP_023217663.1 | putative fatty acyl-CoA reductase CG5065, partial [Centruroides sculpturatus] |
| Scorpion | XP_023217705.1 | eukaryotic translation initiation factor 3 subunit D-like [Centruroides sculpturatus] |
| Scorpion | XP_023217706.1 | proliferation-associated protein 2G4-like [Centruroides sculpturatus] |
| Scorpion | XP_023217707.1 | proliferation-associated protein 2G4-like [Centruroides sculpturatus] |
| Scorpion | XP_023217708.1 | exocyst complex component 5-like [Centruroides sculpturatus] |
| Scorpion | XP_023217713.1 | WD repeat-containing protein 48-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023217723.1 | WD repeat-containing protein 48-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023217724.1 | serine/threonine-protein phosphatase 5-like [Centruroides sculpturatus] |
| Scorpion | XP_023217749.1 | protein kish-B-like [Centruroides sculpturatus] |
| Scorpion | XP_023217777.1 | uncharacterized protein LOC111620132 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023217833.1 | sarcoplasmic calcium-binding proteins I, III, and IV-like [Centruroides sculpturatus] |
| Scorpion | XP_023217901.1 | ADP-ribosylation factor-like protein 5B [Centruroides sculpturatus] |
| Scorpion | XP_023217947.1 | ras-related protein Rab-43-like [Centruroides sculpturatus] |
| Scorpion | XP_023217971.1 | inhibitor of growth protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023218004.1 | kinesin-like protein KIF3A [Centruroides sculpturatus] |

| | | |
|----------|----------------|--|
| Scorpion | XP_023218056.1 | palmitoyltransferase ZDHHC7-like [Centruroides sculpturatus] |
| Scorpion | XP_023218094.1 | 3-methyl-2-oxobutanoate dehydrogenase [lipoamide] kinase, mitochondrial-like [Centruroides sculpturatus] |
| Scorpion | XP_023218113.1 | density-regulated protein homolog isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023218114.1 | density-regulated protein homolog isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023218118.1 | E3 ubiquitin-protein ligase RGLG4-like [Centruroides sculpturatus] |
| Scorpion | XP_023218132.1 | NAD-dependent protein deacylase sirtuin-5, mitochondrial-like [Centruroides sculpturatus] |
| Scorpion | XP_023218172.1 | ATPase ASNA1 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023218201.1 | dehydrogenase/reductase SDR family protein 7-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023218210.1 | metalloprotease TIKI1-like [Centruroides sculpturatus] |
| Scorpion | XP_023218211.1 | adenosine monophosphate-protein transferase FICD homolog [Centruroides sculpturatus] |
| Scorpion | XP_023218212.1 | histone chaperone asf1-like [Centruroides sculpturatus] |
| Scorpion | XP_023218361.1 | BTB/POZ domain-containing protein 3-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023218381.1 | U5 small nuclear ribonucleoprotein 40 kDa protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023218407.1 | guanine nucleotide-binding protein G(s) subunit alpha-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023218408.1 | CCAAT/enhancer-binding protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023218412.1 | guanine nucleotide-binding protein G(o) subunit alpha isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023218444.1 | zinc transporter 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023218466.1 | LOW QUALITY PROTEIN: ras-related protein Rab-11A-like [Centruroides sculpturatus] |
| Scorpion | XP_023218497.1 | SOSS complex subunit C-like [Centruroides sculpturatus] |
| Scorpion | XP_023218522.1 | small conductance calcium-activated potassium channel protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023218550.1 | UDP-N-acetylglucosamine--dolichyl-phosphate N-acetylglucosaminophosphotransferase-like [Centruroides sculpturatus] |
| Scorpion | XP_023218586.1 | uncharacterized protein LOC111620826 [Centruroides sculpturatus] |
| Scorpion | XP_023218642.1 | mRNA decay activator protein ZFP36L1-like [Centruroides sculpturatus] |
| Scorpion | XP_023218654.1 | discoidin domain-containing receptor 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023218690.1 | NEDD4 family-interacting protein 2-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023218694.1 | histone deacetylase complex subunit SAP30 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023218700.1 | vacuolar protein sorting-associated protein 26B-like [Centruroides sculpturatus] |
| Scorpion | XP_023218728.1 | spectrin alpha chain-like [Centruroides sculpturatus] |
| Scorpion | XP_023218752.1 | ubiquitin-conjugating enzyme E2Q-like protein 1 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023218758.1 | transmembrane protein 170A-like [Centruroides sculpturatus] |
| Scorpion | XP_023218761.1 | ras-related protein Rab-9B-like [Centruroides sculpturatus] |
| Scorpion | XP_023218779.1 | ankyrin repeat domain-containing protein 29-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023218828.1 | nuclear transcription factor Y subunit alpha-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023218829.1 | nuclear transcription factor Y subunit alpha-like isoform X2 [Centruroides sculpturatus] |

| | | |
|----------|----------------|--|
| Scorpion | XP_023218948.1 | platelet-activating factor acetylhydrolase IB subunit gamma-like [Centruroides sculpturatus] |
| Scorpion | XP_023218949.1 | uncharacterized protein LOC111621114 [Centruroides sculpturatus] |
| Scorpion | XP_023218952.1 | trafficking protein particle complex subunit 4-like [Centruroides sculpturatus] |
| Scorpion | XP_023219021.1 | 60S ribosomal protein L13a-like [Centruroides sculpturatus] |
| Scorpion | XP_023219062.1 | actin-related protein 2/3 complex subunit 1A-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023219063.1 | actin-related protein 2/3 complex subunit 1A-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219198.1 | ETS domain-containing protein Elk-3-like [Centruroides sculpturatus] |
| Scorpion | XP_023219257.1 | serine/threonine-protein kinase 24-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219259.1 | serine/threonine-protein kinase 24-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023219265.1 | NF-kappa-B inhibitor-interacting Ras-like protein 1 [Centruroides sculpturatus] |
| Scorpion | XP_023219289.1 | glutamate-gated chloride channel-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023219290.1 | glutamate-gated chloride channel-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219318.1 | transcription factor Sp9-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219331.1 | nuclear hormone receptor E75-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219332.1 | nuclear hormone receptor E75-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023219333.1 | nuclear hormone receptor E75-like isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023219335.1 | nuclear hormone receptor E75-like isoform X6 [Centruroides sculpturatus] |
| Scorpion | XP_023219358.1 | sodium/potassium-transporting ATPase subunit alpha-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219373.1 | ecdysone receptor-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023219379.1 | ecdysone receptor-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219456.1 | EF-hand calcium-binding domain-containing protein 14-like [Centruroides sculpturatus] |
| Scorpion | XP_023219495.1 | heparan sulfate glucosamine 3-O-sulfotransferase 3A1-like [Centruroides sculpturatus] |
| Scorpion | XP_023219505.1 | zinc finger protein rotund-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023219506.1 | zinc finger protein rotund-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219653.1 | KRR1 small subunit processome component homolog [Centruroides sculpturatus] |
| Scorpion | XP_023219699.1 | chloride intracellular channel exc-4-like [Centruroides sculpturatus] |
| Scorpion | XP_023219706.1 | transmembrane protein 214-A-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023219747.1 | mitochondrial Rho GTPase 1-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023219758.1 | protein fem-1 homolog A-like [Centruroides sculpturatus] |
| Scorpion | XP_023219766.1 | structural maintenance of chromosomes protein 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023219842.1 | U2 small nuclear ribonucleoprotein A'-like [Centruroides sculpturatus] |
| Scorpion | XP_023219861.1 | mediator of RNA polymerase II transcription subunit 30-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023219878.1 | neural-cadherin-like [Centruroides sculpturatus] |
| Scorpion | XP_023219883.1 | uncharacterized protein LOC111621857 [Centruroides sculpturatus] |
| Scorpion | XP_023219900.1 | zinc finger CCHC domain-containing protein 24-like [Centruroides sculpturatus] |
| Scorpion | XP_023219983.1 | rap guanine nucleotide exchange factor 4-like, partial [Centruroides sculpturatus] |

Scorpion XP_023219984.1 rap guanine nucleotide exchange factor 4-like [Centruroides sculpturatus]

Scorpion XP_023220040.1 DET1- and DDB1-associated protein 1-like [Centruroides sculpturatus]

Scorpion XP_023220095.1 SIN3-HDAC complex-associated factor-like [Centruroides sculpturatus]

Scorpion XP_023220112.1 semaphorin-2A-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220113.1 semaphorin-2A-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023220159.1 transcription factor AP-2-epsilon-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023220160.1 transcription factor AP-2-beta-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023220162.1 PITH domain-containing protein CG6153-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220227.1 ras-like GTP-binding protein Rho1 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220229.1 isoamyl acetate-hydrolyzing esterase 1 homolog [Centruroides sculpturatus]

Scorpion XP_023220267.1 dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B-like [Centruroides sculpturatus]

Scorpion XP_023220299.1 vigilin-like [Centruroides sculpturatus]

Scorpion XP_023220312.1 CTD nuclear envelope phosphatase 1A-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220342.1 kinesin heavy chain-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220344.1 probable small nuclear ribonucleoprotein G [Centruroides sculpturatus]

Scorpion XP_023220373.1 Golgi to ER traffic protein 4 homolog isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220393.1 uncharacterized protein LOC111622269, partial [Centruroides sculpturatus]

Scorpion XP_023220436.1 phosphorylase b kinase gamma catalytic chain, skeletal muscle/heart isoform-like [Centruroides sculpturatus]

Scorpion XP_023220440.1 actin-related protein 2/3 complex subunit 3-like [Centruroides sculpturatus]

Scorpion XP_023220461.1 serine/threonine-protein phosphatase PP1-gamma catalytic subunit [Centruroides sculpturatus]

Scorpion XP_023220478.1 FGFR1 oncogene partner 2 homolog isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220498.1 protein EFR3 homolog B-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220499.1 protein EFR3 homolog B-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023220537.1 glycopeptide N-tetradecanoyltransferase 2-like [Centruroides sculpturatus]

Scorpion XP_023220605.1 caltractin-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220646.1 irregular chiasm C-roughest protein-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023220660.1 eukaryotic peptide chain release factor subunit 1 [Centruroides sculpturatus]

Scorpion XP_023220685.1 gamma-aminobutyric acid receptor subunit beta-like [Centruroides sculpturatus]

Scorpion XP_023220691.1 suppressor of fused homolog [Centruroides sculpturatus]

Scorpion XP_023220692.1 alpha-centractin [Centruroides sculpturatus]

Scorpion XP_023220725.1 leucine-rich repeat-containing protein 15-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220760.1 pre-mRNA-splicing factor 38B-like [Centruroides sculpturatus]

Scorpion XP_023220830.1 basic leucine zipper and W2 domain-containing protein 2-like [Centruroides sculpturatus]

Scorpion XP_023220837.1 fas apoptotic inhibitory molecule 1-like [Centruroides sculpturatus]

Scorpion XP_023220908.1 eukaryotic translation initiation factor 5A-1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220911.1 26S proteasome regulatory subunit 7 [Centruroides sculpturatus]

Scorpion XP_023220935.1 follistatin-related protein 5-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023220936.1 follistatin-related protein 5-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023220937.1 follistatin-related protein 5-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023220983.1 T-complex protein 1 subunit alpha-like [Centruroides sculpturatus]

Scorpion XP_023221000.1 protein argonaute-2-like [Centruroides sculpturatus]

Scorpion XP_023221036.1 serine/threonine-protein phosphatase 2A catalytic subunit beta isoform-like [Centruroides sculpturatus]

Scorpion XP_023221075.1 uncharacterized protein LOC111622872 [Centruroides sculpturatus]

Scorpion XP_023221089.1 complexin-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023221092.1 complexin-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023221094.1 complexin-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023221121.1 E3 ubiquitin-protein ligase MARCH6-like [Centruroides sculpturatus]

Scorpion XP_023221122.1 alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase B-like [Centruroides sculpturatus]

Scorpion XP_023221148.1 Golgi pH regulator B-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023221149.1 Golgi pH regulator B-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023221193.1 elongation of very long chain fatty acids protein 6-like [Centruroides sculpturatus]

Scorpion XP_023221199.1 transmembrane protein 203-like [Centruroides sculpturatus]

Scorpion XP_023221205.1 serine/threonine-protein phosphatase 6 catalytic subunit [Centruroides sculpturatus]

Scorpion XP_023221206.1 xenotropic and polytropic retrovirus receptor 1-like [Centruroides sculpturatus]

Scorpion XP_023221258.1 brain tumor protein-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023221259.1 brain tumor protein-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023221275.1 uncharacterized protein C16orf52 homolog [Centruroides sculpturatus]

Scorpion XP_023221348.1 voltage-dependent T-type calcium channel subunit alpha-1I-like [Centruroides sculpturatus]

Scorpion XP_023221356.1 equilibrative nucleoside transporter 1-like [Centruroides sculpturatus]

Scorpion XP_023221393.1 myophilin-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023221396.1 sorting nexin-3-like [Centruroides sculpturatus]

Scorpion XP_023221418.1 solute carrier family 35 member G1-like [Centruroides sculpturatus]

Scorpion XP_023221442.1 josephin-1-like [Centruroides sculpturatus]

Scorpion XP_023221444.1 60S ribosomal protein L10-like [Centruroides sculpturatus]

Scorpion XP_023221502.1 protein lin-9 homolog [Centruroides sculpturatus]

Scorpion XP_023221590.1 DNA-directed RNA polymerase III subunit RPC2-like [Centruroides sculpturatus]

Scorpion XP_023221654.1 rab11 family-interacting protein 4A-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023221657.1 rab11 family-interacting protein 4-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023221711.1 cofilin/actin-depolymerizing factor homolog isoform X1 [Centruroides sculpturatus]

Scorpion XP_023221712.1 cofilin/actin-depolymerizing factor homolog isoform X2 [Centruroides sculpturatus]

Scorpion XP_023221733.1 E3 ubiquitin-protein ligase Su(dx)-like [Centruroides sculpturatus]

Scorpion XP_023221773.1 dynactin subunit 6-like [Centruroides sculpturatus]

Scorpion XP_023221782.1 uncharacterized protein LOC111623440 [Centruroides sculpturatus]

Scorpion XP_023221828.1 probable ubiquitin carboxyl-terminal hydrolase FAF-X isoform X1 [Centruroides sculpturatus]

Scorpion XP_023221864.1 coronin-1C-like [Centruroides sculpturatus]

Scorpion XP_023221866.1 coronin-1C-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023221867.1 small nuclear ribonucleoprotein Sm D3-like [Centruroides sculpturatus]

| | | |
|----------|----------------|---|
| Scorpion | XP_023221875.1 | epidermal growth factor receptor-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023221876.1 | epidermal growth factor receptor-like [Centruroides sculpturatus] |
| Scorpion | XP_023221882.1 | uncharacterized protein LOC111623514 [Centruroides sculpturatus] |
| Scorpion | XP_023221908.1 | probable serine incorporator isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023221922.1 | dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A-like [Centruroides sculpturatus] |
| Scorpion | XP_023221974.1 | coactosin-like protein [Centruroides sculpturatus] |
| Scorpion | XP_023222056.1 | polymerase delta-interacting protein 2-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023222101.1 | protein pellino-like [Centruroides sculpturatus] |
| Scorpion | XP_023222113.1 | protein pellino-like [Centruroides sculpturatus] |
| Scorpion | XP_023222119.1 | cleavage stimulation factor subunit 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023222204.1 | ubiquitin carboxyl-terminal hydrolase 46-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023222205.1 | ubiquitin carboxyl-terminal hydrolase 46-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023222232.1 | peptidyl-prolyl cis-trans isomerase H-like [Centruroides sculpturatus] |
| Scorpion | XP_023222238.1 | 26S proteasome non-ATPase regulatory subunit 6-like [Centruroides sculpturatus] |
| Scorpion | XP_023222249.1 | ribosome biogenesis protein NSA2 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023222294.1 | neuroglobin-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023222299.1 | signal peptidase complex catalytic subunit SEC11A isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023222301.1 | signal peptidase complex catalytic subunit SEC11A isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023222307.1 | N6-adenosine-methyltransferase catalytic subunit-like [Centruroides sculpturatus] |
| Scorpion | XP_023222359.1 | uncharacterized protein LOC111623868 [Centruroides sculpturatus] |
| Scorpion | XP_023222398.1 | 60S ribosomal protein L10a-2 [Centruroides sculpturatus] |
| Scorpion | XP_023222475.1 | synaptic functional regulator FMR1-like [Centruroides sculpturatus] |
| Scorpion | XP_023222506.1 | signal recognition particle 54 kDa protein [Centruroides sculpturatus] |
| Scorpion | XP_023222508.1 | enhancer of rudimentary homolog [Centruroides sculpturatus] |
| Scorpion | XP_023222546.1 | 60S ribosomal protein L21-like [Centruroides sculpturatus] |
| Scorpion | XP_023222600.1 | E3 ubiquitin-protein ligase UBR2-like [Centruroides sculpturatus] |
| Scorpion | XP_023222601.1 | ras-related and estrogen-regulated growth inhibitor-like protein [Centruroides sculpturatus] |
| Scorpion | XP_023222604.1 | leptin receptor overlapping transcript-like 1 [Centruroides sculpturatus] |
| Scorpion | XP_023222643.1 | rho-related GTP-binding protein RhoE-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023222650.1 | transforming growth factor beta receptor type 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023222674.1 | heparan sulfate glucosamine 3-O-sulfotransferase 5-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023222730.1 | tubulin alpha-1B chain-like [Centruroides sculpturatus] |
| Scorpion | XP_023222750.1 | uncharacterized protein LOC111624174 [Centruroides sculpturatus] |
| Scorpion | XP_023222759.1 | glutamate-gated chloride channel-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023222878.1 | importin subunit beta-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023222881.1 | protein transport protein Sec61 subunit beta-like [Centruroides sculpturatus] |
| Scorpion | XP_023222889.1 | ubiquitin-conjugating enzyme E2-24 kDa-like [Centruroides sculpturatus] |

| | | |
|----------|----------------|--|
| Scorpion | XP_023222892.1 | mannose-1-phosphate guanyltransferase alpha-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023222896.1 | cytohesin-interacting protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023222923.1 | RING finger protein 44-like [Centruroides sculpturatus] |
| Scorpion | XP_023222932.1 | tankyrase-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023222941.1 | sterol O-acyltransferase 1-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023222943.1 | uncharacterized protein LOC111624351 [Centruroides sculpturatus] |
| Scorpion | XP_023222966.1 | cytohesin-1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023222974.1 | cytohesin-1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023222983.1 | cytohesin-1-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023222990.1 | fibroblast growth factor 17-like [Centruroides sculpturatus] |
| Scorpion | XP_023222995.1 | phosphatidylinositol 4-kinase type 2-beta-like [Centruroides sculpturatus] |
| Scorpion | XP_023223008.1 | protein lin-7 homolog C isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023223009.1 | protein lin-7 homolog C isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023223014.1 | WD repeat-containing protein 13-like [Centruroides sculpturatus] |
| Scorpion | XP_023223028.1 | F-box/LRR-repeat protein 15-like [Centruroides sculpturatus] |
| Scorpion | XP_023223078.1 | zinc finger with UFM1-specific peptidase domain protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023223113.1 | cyclin-dependent kinase 9-like [Centruroides sculpturatus] |
| Scorpion | XP_023223142.1 | V-type proton ATPase 21 kDa proteolipid subunit-like [Centruroides sculpturatus] |
| Scorpion | XP_023223169.1 | E3 ubiquitin-protein ligase Nedd-4-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023223170.1 | UV excision repair protein RAD23 homolog B-like [Centruroides sculpturatus] |
| Scorpion | XP_023223176.1 | E3 ubiquitin-protein ligase Nedd-4-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023223213.1 | formin-like protein CG32138 [Centruroides sculpturatus] |
| Scorpion | XP_023223233.1 | adenylate cyclase type 6-like [Centruroides sculpturatus] |
| Scorpion | XP_023223241.1 | growth/differentiation factor 8-like [Centruroides sculpturatus] |
| Scorpion | XP_023223242.1 | inhibin beta B chain-like [Centruroides sculpturatus] |
| Scorpion | XP_023223359.1 | soma ferritin-like [Centruroides sculpturatus] |
| Scorpion | XP_023223366.1 | repressor of RNA polymerase III transcription MAF1 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023223388.1 | F-box-like/WD repeat-containing protein TBL1XR1 [Centruroides sculpturatus] |
| Scorpion | XP_023223419.1 | eukaryotic translation initiation factor 3 subunit E-B-like [Centruroides sculpturatus] |
| Scorpion | XP_023223447.1 | leucine-rich repeat neuronal protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023223591.1 | tensin-2-like [Centruroides sculpturatus] |
| Scorpion | XP_023223610.1 | N-acetylneuraminase 9-O-acetyltransferase-like [Centruroides sculpturatus] |
| Scorpion | XP_023223621.1 | death-associated protein 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023223622.1 | death-associated protein 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023223639.1 | nuclear factor 1 X-type-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023223661.1 | T-complex protein 1 subunit theta-like [Centruroides sculpturatus] |
| Scorpion | XP_023223693.1 | trimeric intracellular cation channel type 1B.1-like [Centruroides sculpturatus] |
| Scorpion | XP_023223794.1 | transmembrane protein 145-like [Centruroides sculpturatus] |
| Scorpion | XP_023223804.1 | uncharacterized protein LOC111625027 [Centruroides sculpturatus] |
| Scorpion | XP_023223823.1 | ankyrin-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023223894.1 | cohesin subunit SA-1-like [Centruroides sculpturatus] |

| | | |
|----------|----------------|---|
| Scorpion | XP_023223935.1 | serine/threonine-protein kinase tousled-like 2 isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023223936.1 | serine/threonine-protein kinase tousled-like 2 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023223987.1 | E3 ubiquitin-protein ligase CBL-like isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023224007.1 | uncharacterized protein LOC111625177 [Centruroides sculpturatus] |
| Scorpion | XP_023224026.1 | probable ATP-dependent RNA helicase DDX46 [Centruroides sculpturatus] |
| Scorpion | XP_023224058.1 | zinc finger protein Noc-like [Centruroides sculpturatus] |
| Scorpion | XP_023224139.1 | malignant T-cell-amplified sequence 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023224206.1 | kelch-like protein 2 isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023224324.1 | class A basic helix-loop-helix protein 15-like [Centruroides sculpturatus] |
| Scorpion | XP_023224425.1 | solute carrier family 25 member 51-like [Centruroides sculpturatus] |
| Scorpion | XP_023224432.1 | solute carrier family 25 member 51-like [Centruroides sculpturatus] |
| Scorpion | XP_023224455.1 | NAD(P) transhydrogenase, mitochondrial-like [Centruroides sculpturatus] |
| Scorpion | XP_023224489.1 | BRISC and BRCA1-A complex member 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023224524.1 | inhibitor of growth protein 5-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023224525.1 | glyceraldehyde-3-phosphate dehydrogenase-like [Centruroides sculpturatus] |
| Scorpion | XP_023224547.1 | junctophilin-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023224568.1 | 60S ribosomal protein L13-like [Centruroides sculpturatus] |
| Scorpion | XP_023224587.1 | cilia- and flagella-associated protein 20 [Centruroides sculpturatus] |
| Scorpion | XP_023224601.1 | protein purity of essence-like [Centruroides sculpturatus] |
| Scorpion | XP_023224609.1 | MAGUK p55 subfamily member 6-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023224615.1 | MAGUK p55 subfamily member 6-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023224629.1 | mitochondrial import inner membrane translocase subunit Tim10-B-like [Centruroides sculpturatus] |
| Scorpion | XP_023224637.1 | elongation of very long chain fatty acids protein AAEL008004-like [Centruroides sculpturatus] |
| Scorpion | XP_023224647.1 | protein Hook homolog 3-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023224648.1 | protein Hook homolog 3-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023224651.1 | protein Hook homolog 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023224658.1 | coatomer subunit alpha-like [Centruroides sculpturatus] |
| Scorpion | XP_023224721.1 | NHP2-like protein 1 [Centruroides sculpturatus] |
| Scorpion | XP_023224752.1 | transmembrane protein 35A-like [Centruroides sculpturatus] |
| Scorpion | XP_023224791.1 | NADH dehydrogenase [ubiquinone] iron-sulfur protein 2, mitochondrial-like [Centruroides sculpturatus] |
| Scorpion | XP_023224899.1 | programmed cell death protein 4-like [Centruroides sculpturatus] |
| Scorpion | XP_023224901.1 | probable transaldolase isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023224903.1 | probable transaldolase isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023224907.1 | uncharacterized protein LOC111625886 isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023224921.1 | signal transducer and activator of transcription 5A-like [Centruroides sculpturatus] |
| Scorpion | XP_023224957.1 | sprouty-related, EVH1 domain-containing protein 2-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023225052.1 | partitioning defective 6 homolog beta-like [Centruroides sculpturatus] |
| Scorpion | XP_023225091.1 | putative RNA-binding protein 15 [Centruroides sculpturatus] |
| Scorpion | XP_023225110.1 | delta-1-pyrroline-5-carboxylate dehydrogenase, mitochondrial-like [Centruroides sculpturatus] |

Scorpion XP_023225139.1 tail-anchored protein insertion receptor WRB-like [Centruroides sculpturatus]

Scorpion XP_023225210.1 homeobox protein Meis2-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023225218.1 homeobox protein Meis2-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023225223.1 homeobox protein Meis2-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023225232.1 homeobox protein Meis2-like isoform X5 [Centruroides sculpturatus]

Scorpion XP_023225253.1 thyroid hormone receptor beta-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023225287.1 histone-arginine methyltransferase CARMER-like [Centruroides sculpturatus]

Scorpion XP_023225311.1 neuroglobin-like [Centruroides sculpturatus]

Scorpion XP_023225327.1 zinc finger FYVE domain-containing protein 9-like [Centruroides sculpturatus]

Scorpion XP_023225330.1 Golgi resident protein GCP60-like [Centruroides sculpturatus]

Scorpion XP_023225367.1 WD repeat-containing protein 61-like [Centruroides sculpturatus]

Scorpion XP_023225437.1 myosin heavy chain, non-muscle-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023225453.1 mitochondrial dicarboxylate carrier-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023225454.1 mitochondrial dicarboxylate carrier-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023225466.1 notch-regulated ankyrin repeat-containing protein-like [Centruroides sculpturatus]

Scorpion XP_023225469.1 neurogenic locus Notch protein-like [Centruroides sculpturatus]

Scorpion XP_023225472.1 60 kDa SS-A/Ro ribonucleoprotein-like [Centruroides sculpturatus]

Scorpion XP_023225494.1 U4/U6 small nuclear ribonucleoprotein Prp3-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023225495.1 U4/U6 small nuclear ribonucleoprotein Prp3-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023225542.1 fructose-bisphosphate aldolase-like [Centruroides sculpturatus]

Scorpion XP_023225546.1 phosphatidylinositol transfer protein alpha isoform-like [Centruroides sculpturatus]

Scorpion XP_023225590.1 E3 ubiquitin-protein ligase KCMF1-like [Centruroides sculpturatus]

Scorpion XP_023225633.1 CLIP-associating protein 2-like [Centruroides sculpturatus]

Scorpion XP_023225664.1 26S proteasome regulatory subunit 6B [Centruroides sculpturatus]

Scorpion XP_023225665.1 tubulin beta chain-like [Centruroides sculpturatus]

Scorpion XP_023225673.1 ubiquitin thioesterase ZRANB1-like [Centruroides sculpturatus]

Scorpion XP_023225717.1 BTB/POZ domain-containing protein KCTD12-like [Centruroides sculpturatus]

Scorpion XP_023225720.1 ankyrin repeat and SOCS box protein 3-like [Centruroides sculpturatus]

Scorpion XP_023225744.1 transcriptional repressor CTCF-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023225753.1 STE20-related kinase adapter protein alpha-like [Centruroides sculpturatus]

Scorpion XP_023225788.1 TOX high mobility group box family member 3-like [Centruroides sculpturatus]

Scorpion XP_023225823.1 diacylglycerol kinase 1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023225851.1 COP9 signalosome complex subunit 4-like [Centruroides sculpturatus]

Scorpion XP_023225892.1 protein KIAA0100-like [Centruroides sculpturatus]

Scorpion XP_023225901.1 importin subunit alpha-6-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023225902.1 uncharacterized protein LOC111626669 [Centruroides sculpturatus]

Scorpion XP_023225923.1 plexin-B-like [Centruroides sculpturatus]

Scorpion XP_023225936.1 vacuole membrane protein 1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023225940.1 neurocalcin homolog [Centruroides sculpturatus]

Scorpion XP_023226083.1 putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX16 [Centruroides sculpturatus]

Scorpion XP_023226085.1 tetraspanin-9-like [Centruroides sculpturatus]

| | | |
|----------|----------------|---|
| Scorpion | XP_023226090.1 | structural maintenance of chromosomes protein 1A-like [Centruroides sculpturatus] |
| Scorpion | XP_023226148.1 | receptor-type tyrosine-protein phosphatase kappa-like [Centruroides sculpturatus] |
| Scorpion | XP_023226150.1 | transcription factor AP-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023226223.1 | myosin-VIIa-like [Centruroides sculpturatus] |
| Scorpion | XP_023226278.1 | uncharacterized protein LOC111626966 [Centruroides sculpturatus] |
| Scorpion | XP_023226283.1 | polycomb group RING finger protein 1-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023226301.1 | eukaryotic translation initiation factor 4E-like [Centruroides sculpturatus] |
| Scorpion | XP_023226314.1 | sodium- and chloride-dependent taurine transporter-like [Centruroides sculpturatus] |
| Scorpion | XP_023226322.1 | V-type proton ATPase subunit B [Centruroides sculpturatus] |
| Scorpion | XP_023226425.1 | transmembrane protein FAM155A-like [Centruroides sculpturatus] |
| Scorpion | XP_023226555.1 | arrestin domain-containing protein 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023226616.1 | cAMP-dependent protein kinase regulatory subunit-like [Centruroides sculpturatus] |
| Scorpion | XP_023226637.1 | B-cell lymphoma/leukemia 11A-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023226644.1 | uncharacterized protein LOC111627325 [Centruroides sculpturatus] |
| Scorpion | XP_023226645.1 | homeobox protein six1-like [Centruroides sculpturatus] |
| Scorpion | XP_023226646.1 | cysteine-rich protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023226680.1 | SAGA-associated factor 29-like [Centruroides sculpturatus] |
| Scorpion | XP_023226683.1 | uncharacterized protein LOC111627357 [Centruroides sculpturatus] |
| Scorpion | XP_023226686.1 | serine/threonine-protein kinase tricornet-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023226687.1 | serine/threonine-protein kinase tricornet-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023226695.1 | polycomb protein SCM1-like [Centruroides sculpturatus] |
| Scorpion | XP_023226707.1 | protein YIPF5-like [Centruroides sculpturatus] |
| Scorpion | XP_023226789.1 | uncharacterized protein C45G9.7-like [Centruroides sculpturatus] |
| Scorpion | XP_023226813.1 | parafibromin-like [Centruroides sculpturatus] |
| Scorpion | XP_023226853.1 | fructose-1,6-bisphosphatase 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023226899.1 | zinc finger protein ZIC 4-like [Centruroides sculpturatus] |
| Scorpion | XP_023226909.1 | protein couch potato-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023226939.1 | high mobility group protein 20A-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023226943.1 | high mobility group protein 20A-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023226955.1 | aspartate--tRNA ligase, cytoplasmic-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023227000.1 | UBA-like domain-containing protein 2 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023227002.1 | gamma-tubulin complex component 4-like [Centruroides sculpturatus] |
| Scorpion | XP_023227020.1 | glucose-induced degradation protein 4 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023227029.1 | protein enhancer of sevenless 2B [Centruroides sculpturatus] |
| Scorpion | XP_023227103.1 | unconventional myosin-Id-like [Centruroides sculpturatus] |
| Scorpion | XP_023227141.1 | 60S ribosomal protein L35-like [Centruroides sculpturatus] |
| Scorpion | XP_023227170.1 | lachesin-like [Centruroides sculpturatus] |
| Scorpion | XP_023227219.1 | mediator of RNA polymerase II transcription subunit 24-like [Centruroides sculpturatus] |
| Scorpion | XP_023227232.1 | charged multivesicular body protein 1a-like [Centruroides sculpturatus] |

Scorpion XP_023227281.1 staphylococcal nuclease domain-containing protein 1-like [Centruroides sculpturatus]

Scorpion XP_023227331.1 F-box only protein 21-like [Centruroides sculpturatus]

Scorpion XP_023227348.1 coiled-coil domain-containing protein 6-like [Centruroides sculpturatus]

Scorpion XP_023227351.1 mitochondrial import inner membrane translocase subunit Tim23-like [Centruroides sculpturatus]

Scorpion XP_023227353.1 dnaJ homolog subfamily B member 14-like [Centruroides sculpturatus]

Scorpion XP_023227383.1 phosphoglycerate kinase-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023227384.1 ATP-binding cassette sub-family E member 1 [Centruroides sculpturatus]

Scorpion XP_023227434.1 uncharacterized protein LOC111627964 [Centruroides sculpturatus]

Scorpion XP_023227442.1 zinc finger protein 337-like [Centruroides sculpturatus]

Scorpion XP_023227447.1 protein furry-like [Centruroides sculpturatus]

Scorpion XP_023227450.1 polyadenylate-binding protein 4-like [Centruroides sculpturatus]

Scorpion XP_023227468.1 trafficking protein particle complex subunit 1-like [Centruroides sculpturatus]

Scorpion XP_023227484.1 TNF receptor-associated factor 4-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023227494.1 RNA-binding protein MEX3B-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023227503.1 eukaryotic translation initiation factor 1 [Centruroides sculpturatus]

Scorpion XP_023227508.1 activin receptor type-1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023227517.1 transcription factor BTF3-like [Centruroides sculpturatus]

Scorpion XP_023227542.1 sideroflexin-1-like [Centruroides sculpturatus]

Scorpion XP_023227551.1 hexokinase-2-like [Centruroides sculpturatus]

Scorpion XP_023227595.1 mitochondrial carnitine/acylcarnitine carrier protein-like [Centruroides sculpturatus]

Scorpion XP_023227601.1 calsynenin-1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023227605.1 kazrin-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023227607.1 kazrin-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023227609.1 actin-related protein 2/3 complex subunit 4-like [Centruroides sculpturatus]

Scorpion XP_023227620.1 glutamate receptor ionotropic, kainate 2-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023227629.1 G protein-coupled receptor kinase 1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023227641.1 proteasome subunit alpha type-1-like [Centruroides sculpturatus]

Scorpion XP_023227642.1 zinc finger CCCH domain-containing protein 10-like [Centruroides sculpturatus]

Scorpion XP_023227654.1 cGMP-inhibited 3',5'-cyclic phosphodiesterase A-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023227698.1 polycomb group RING finger protein 3-like [Centruroides sculpturatus]

Scorpion XP_023227702.1 putative nucleotidyltransferase FAM46C [Centruroides sculpturatus]

Scorpion XP_023227712.1 soluble guanylate cyclase 88E-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023227714.1 soluble guanylate cyclase 88E-like [Centruroides sculpturatus]

Scorpion XP_023227730.1 colorectal mutant cancer protein-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023227739.1 replication factor C subunit 3-like [Centruroides sculpturatus]

Scorpion XP_023227755.1 serine/threonine-protein kinase B-raf-like [Centruroides sculpturatus]

Scorpion XP_023227767.1 protein Smaug homolog 1-like [Centruroides sculpturatus]

Scorpion XP_023227769.1 protein Smaug homolog 1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023227771.1 alpha-actinin isoform X1 [Centruroides sculpturatus]

Scorpion XP_023227772.1 alpha-actinin isoform X2 [Centruroides sculpturatus]

Scorpion XP_023227773.1 alpha-actinin isoform X3 [Centruroides sculpturatus]

Scorpion XP_023227778.1 sorting nexin-6-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023227811.1 ubiquitin domain-containing protein 2-like [Centruroides sculpturatus]

Scorpion XP_023227813.1 GTP-binding protein 2-like [Centruroides sculpturatus]

Scorpion XP_023227824.1 ubiquitin-like domain-containing CTD phosphatase 1 [Centruroides sculpturatus]

Scorpion XP_023227827.1 BCL2/adenovirus E1B 19 kDa protein-interacting protein 3-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023227847.1 iroquois-class homeodomain protein IRX-6-like [Centruroides sculpturatus]

Scorpion XP_023227930.1 guanine nucleotide-binding protein subunit beta-5-like [Centruroides sculpturatus]

Scorpion XP_023227933.1 protein ABHD17B-like [Centruroides sculpturatus]

Scorpion XP_023227934.1 ATP synthase subunit alpha, mitochondrial-like [Centruroides sculpturatus]

Scorpion XP_023227944.1 cartilage oligomeric matrix protein-like [Centruroides sculpturatus]

Scorpion XP_023227948.1 60S ribosomal protein L17 [Centruroides sculpturatus]

Scorpion XP_023228028.1 transcription and mRNA export factor ENY2-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023228030.1 ELAV-like protein 2 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023228147.1 regulator complex protein LAMTOR4-like [Centruroides sculpturatus]

Scorpion XP_023228166.1 ELAV-like protein 1 isoform X3 [Centruroides sculpturatus]

Scorpion XP_023228190.1 beclin-1-like [Centruroides sculpturatus]

Scorpion XP_023228207.1 negative elongation factor B-like [Centruroides sculpturatus]

Scorpion XP_023228244.1 polyadenylate-binding protein 1-like [Centruroides sculpturatus]

Scorpion XP_023228266.1 E3 ubiquitin-protein ligase mib1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023228288.1 ELAV-like protein 2 isoform X5 [Centruroides sculpturatus]

Scorpion XP_023228353.1 LIM domain transcription factor LMO4-like [Centruroides sculpturatus]

Scorpion XP_023228404.1 voltage-gated potassium channel subunit beta-2-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023228405.1 voltage-gated potassium channel subunit beta-2-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023228436.1 ATP-citrate synthase-like [Centruroides sculpturatus]

Scorpion XP_023228483.1 sorting nexin-3-like [Centruroides sculpturatus]

Scorpion XP_023228486.1 superoxide dismutase [Cu-Zn]-like [Centruroides sculpturatus]

Scorpion XP_023228511.1 transcriptional regulator ovo-like [Centruroides sculpturatus]

Scorpion XP_023228530.1 LHFPL tetraspan subfamily member 3 protein-like [Centruroides sculpturatus]

Scorpion XP_023228541.1 troponin C-like [Centruroides sculpturatus]

Scorpion XP_023228554.1 MOB kinase activator-like 2 [Centruroides sculpturatus]

Scorpion XP_023228556.1 AP-2 complex subunit alpha-like [Centruroides sculpturatus]

Scorpion XP_023228596.1 estrogen-related receptor gamma-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023228599.1 synaptosomal-associated protein 25-like [Centruroides sculpturatus]

Scorpion XP_023228613.1 BTB/POZ domain-containing protein KCTD5-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023228622.1 26S proteasome non-ATPase regulatory subunit 12-like [Centruroides sculpturatus]

Scorpion XP_023228643.1 transitional endoplasmic reticulum ATPase isoform X1 [Centruroides sculpturatus]

Scorpion XP_023228644.1 transitional endoplasmic reticulum ATPase isoform X2 [Centruroides sculpturatus]

Scorpion XP_023228647.1 kelch-like protein 20 [Centruroides sculpturatus]

Scorpion XP_023228665.1 ADP-ribosylation factor-like protein 5A [Centruroides sculpturatus]

| | | |
|----------|----------------|---|
| Scorpion | XP_023228687.1 | mediator of RNA polymerase II transcription subunit 6-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023228708.1 | pleckstrin homology domain-containing family G member 5-like isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023228739.1 | neuropeptide Y receptor type 5-like [Centruroides sculpturatus] |
| Scorpion | XP_023228758.1 | phospholipid phosphatase-related protein type 5-like [Centruroides sculpturatus] |
| Scorpion | XP_023228802.1 | protein Wnt-6-like [Centruroides sculpturatus] |
| Scorpion | XP_023228812.1 | ubiquitin-conjugating enzyme E2 W-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023228868.1 | homeobox protein abdominal-B-like [Centruroides sculpturatus] |
| Scorpion | XP_023228912.1 | 60S ribosomal protein L8-like [Centruroides sculpturatus] |
| Scorpion | XP_023228942.1 | ER membrane protein complex subunit 8-like [Centruroides sculpturatus] |
| Scorpion | XP_023228949.1 | blood vessel epicardial substance-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023228958.1 | vegetative incompatibility protein HET-E-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023229015.1 | macoilin-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023229046.1 | ER membrane protein complex subunit 4-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023229052.1 | thyrotropin-releasing hormone receptor-like [Centruroides sculpturatus] |
| Scorpion | XP_023229197.1 | tensin-2-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023229211.1 | cAMP-dependent protein kinase catalytic subunit [Centruroides sculpturatus] |
| Scorpion | XP_023229217.1 | Krueppel-like factor 6 [Centruroides sculpturatus] |
| Scorpion | XP_023229218.1 | uncharacterized protein LOC111629531 [Centruroides sculpturatus] |
| Scorpion | XP_023229226.1 | frizzled-5-like [Centruroides sculpturatus] |
| Scorpion | XP_023229324.1 | ras-related protein M-Ras-like [Centruroides sculpturatus] |
| Scorpion | XP_023229327.1 | tyrosine-protein phosphatase non-receptor type 4-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023229328.1 | tyrosine-protein phosphatase non-receptor type 4-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023229369.1 | regulating synaptic membrane exocytosis protein 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023229440.1 | 40S ribosomal protein S3-A-like [Centruroides sculpturatus] |
| Scorpion | XP_023229447.1 | selenide, water dikinase [Centruroides sculpturatus] |
| Scorpion | XP_023229595.1 | vacuolar protein sorting-associated protein 26B-like [Centruroides sculpturatus] |
| Scorpion | XP_023229644.1 | uncharacterized protein LOC111629926 [Centruroides sculpturatus] |
| Scorpion | XP_023229663.1 | dnaJ homolog subfamily B member 4-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023229666.1 | uncharacterized protein LOC111629938 isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023229707.1 | CDGSH iron-sulfur domain-containing protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023229715.1 | cytoplasmic polyadenylation element-binding protein 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023229729.1 | 40S ribosomal protein S7-like [Centruroides sculpturatus] |
| Scorpion | XP_023229732.1 | cytoplasmic polyadenylation element-binding protein 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023229735.1 | brain tumor protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023229736.1 | mediator of RNA polymerase II transcription subunit 7-like [Centruroides sculpturatus] |

| | | |
|----------|----------------|---|
| Scorpion | XP_023229737.1 | anaphase-promoting complex subunit 15-like [Centruroides sculpturatus] |
| Scorpion | XP_023229790.1 | sodium leak channel non-selective protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023229817.1 | gamma-tubulin complex component 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023229851.1 | calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023229856.1 | calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023229872.1 | calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023229901.1 | phosphofurin acidic cluster sorting protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023229903.1 | protein cornichon-like [Centruroides sculpturatus] |
| Scorpion | XP_023229951.1 | MOG interacting and ectopic P-granules protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023229957.1 | septin-7-like isoform X5 [Centruroides sculpturatus] |
| Scorpion | XP_023229983.1 | small nuclear ribonucleoprotein Sm D2 [Centruroides sculpturatus] |
| Scorpion | XP_023230002.1 | syntaxin-1A-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023230023.1 | GTP-binding nuclear protein Ran-like [Centruroides sculpturatus] |
| Scorpion | XP_023230029.1 | UPF0577 protein KIAA1324-like homolog [Centruroides sculpturatus] |
| Scorpion | XP_023230032.1 | E3 ubiquitin-protein ligase UBR4-like [Centruroides sculpturatus] |
| Scorpion | XP_023230037.1 | importin-9-like [Centruroides sculpturatus] |
| Scorpion | XP_023230065.1 | serine/threonine-protein phosphatase 4 catalytic subunit B-like [Centruroides sculpturatus] |
| Scorpion | XP_023230203.1 | serine/threonine-protein kinase 3-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023230207.1 | exostosin-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023230209.1 | transmembrane protein 185A-like [Centruroides sculpturatus] |
| Scorpion | XP_023230246.1 | serine/threonine-protein kinase NLK-like [Centruroides sculpturatus] |
| Scorpion | XP_023230249.1 | beta-1,4-mannosyltransferase egh-like [Centruroides sculpturatus] |
| Scorpion | XP_023230270.1 | transmembrane protein 59-like [Centruroides sculpturatus] |
| Scorpion | XP_023230301.1 | IQ motif and SEC7 domain-containing protein 1-like isoform X6 [Centruroides sculpturatus] |
| Scorpion | XP_023230302.1 | IQ motif and SEC7 domain-containing protein 1-like isoform X7 [Centruroides sculpturatus] |
| Scorpion | XP_023230304.1 | IQ motif and SEC7 domain-containing protein 1-like isoform X8 [Centruroides sculpturatus] |
| Scorpion | XP_023230307.1 | LIM domain-containing protein jub-like [Centruroides sculpturatus] |
| Scorpion | XP_023230325.1 | sorting nexin-21-like [Centruroides sculpturatus] |
| Scorpion | XP_023230326.1 | sorting nexin-21-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023230398.1 | transmembrane protein 47-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023230435.1 | mitogen-activated protein kinase 14-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023230563.1 | nuclear receptor ROR-beta-like [Centruroides sculpturatus] |
| Scorpion | XP_023230585.1 | COP9 signalosome complex subunit 3-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023230631.1 | dedicator of cytokinesis protein 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023230646.1 | 116 kDa U5 small nuclear ribonucleoprotein component-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023230648.1 | 116 kDa U5 small nuclear ribonucleoprotein component-like [Centruroides sculpturatus] |
| Scorpion | XP_023230670.1 | mediator of RNA polymerase II transcription subunit 9-like [Centruroides sculpturatus] |

Scorpion XP_023230709.1 neurogenic protein big brain-like [Centruroides sculpturatus]

Scorpion XP_023230711.1 dihydropyrimidinase-like, partial [Centruroides sculpturatus]

Scorpion XP_023230740.1 uncharacterized protein LOC111630815, partial [Centruroides sculpturatus]

Scorpion XP_023230791.1 protein N-terminal asparagine amidohydrolase-like [Centruroides sculpturatus]

Scorpion XP_023230820.1 UPF0505 protein C16orf62-like [Centruroides sculpturatus]

Scorpion XP_023230827.1 ras-specific guanine nucleotide-releasing factor RalGPS1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023230828.1 ras-specific guanine nucleotide-releasing factor RalGPS1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023230829.1 ras-specific guanine nucleotide-releasing factor RalGPS1-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023230850.1 V-type proton ATPase subunit d 1-like [Centruroides sculpturatus]

Scorpion XP_023230895.1 sodium-dependent proline transporter-like [Centruroides sculpturatus]

Scorpion XP_023230923.1 semaphorin-5A-like [Centruroides sculpturatus]

Scorpion XP_023230929.1 mothers against decapentaplegic homolog 4-like [Centruroides sculpturatus]

Scorpion XP_023230954.1 uncharacterized protein LOC111631004 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023230955.1 uncharacterized protein LOC111631004 isoform X2 [Centruroides sculpturatus]

Scorpion XP_023230957.1 ATP-binding cassette sub-family G member 4-like [Centruroides sculpturatus]

Scorpion XP_023230958.1 ATP-binding cassette sub-family G member 4-like [Centruroides sculpturatus]

Scorpion XP_023231028.1 actin-related protein 8-like [Centruroides sculpturatus]

Scorpion XP_023231042.1 zinc transporter 2-like isoform X5 [Centruroides sculpturatus]

Scorpion XP_023231046.1 LARGE xylosyl- and glucuronyltransferase 1-like [Centruroides sculpturatus]

Scorpion XP_023231089.1 MICAL-like protein 2 [Centruroides sculpturatus]

Scorpion XP_023231107.1 elongation factor 1-alpha [Centruroides sculpturatus]

Scorpion XP_023231182.1 enhancer of rudimentary homolog [Centruroides sculpturatus]

Scorpion XP_023231188.1 AP-2 complex subunit sigma [Centruroides sculpturatus]

Scorpion XP_023231209.1 60S ribosomal protein L23 [Centruroides sculpturatus]

Scorpion XP_023231270.1 protein transport protein Sec61 subunit alpha [Centruroides sculpturatus]

Scorpion XP_023231276.1 innexin shaking-B-like [Centruroides sculpturatus]

Scorpion XP_023231296.1 innexin inx2-like [Centruroides sculpturatus]

Scorpion XP_023231343.1 U6 snRNA-associated Sm-like protein LSm2 [Centruroides sculpturatus]

Scorpion XP_023231372.1 proteasome subunit alpha type-2-like [Centruroides sculpturatus]

Scorpion XP_023231381.1 U6 snRNA-associated Sm-like protein LSm1 [Centruroides sculpturatus]

Scorpion XP_023231388.1 nuclear pore complex protein NUP1-like [Centruroides sculpturatus]

Scorpion XP_023231394.1 acetylcholine receptor subunit alpha-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023231429.1 homeobox protein OTX2-like [Centruroides sculpturatus]

Scorpion XP_023231453.1 DENN domain-containing protein 5A-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023231454.1 DENN domain-containing protein 5A-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023231486.1 voltage-dependent L-type calcium channel subunit beta-2-like [Centruroides sculpturatus]

Scorpion XP_023231526.1 serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit epsilon isoform-like [Centruroides sculpturatus]

Scorpion XP_023231547.1 ubiquitin-conjugating enzyme E2 Q2-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023231548.1 ubiquitin-conjugating enzyme E2 Q2-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023231557.1 mortality factor 4-like protein 1 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023231559.1 mortality factor 4-like protein 1 isoform X2 [Centruroides sculpturatus]

Scorpion XP_023231562.1 tropomyosin isoform X3 [Centruroides sculpturatus]

Scorpion XP_023231563.1 tropomyosin isoform X4 [Centruroides sculpturatus]

Scorpion XP_023231567.1 protein fem-1 homolog B-like [Centruroides sculpturatus]

Scorpion XP_023231573.1 Down syndrome cell adhesion molecule-like protein Dscam2 [Centruroides sculpturatus]

Scorpion XP_023231595.1 septin-2-like [Centruroides sculpturatus]

Scorpion XP_023231610.1 T-complex protein 1 subunit eta-like [Centruroides sculpturatus]

Scorpion XP_023231614.1 CCR4-NOT transcription complex subunit 7-like [Centruroides sculpturatus]

Scorpion XP_023231657.1 uncharacterized protein LOC111631600 [Centruroides sculpturatus]

Scorpion XP_023231698.1 guanine nucleotide exchange factor MSS4-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023231771.1 homeobox protein PKNOX1-like [Centruroides sculpturatus]

Scorpion XP_023231818.1 E3 ubiquitin-protein ligase RNF170-like, partial [Centruroides sculpturatus]

Scorpion XP_023231832.1 phosphatidylcholine:ceramide cholinephosphotransferase 2-like [Centruroides sculpturatus]

Scorpion XP_023231838.1 rho GTPase-activating protein 39-like [Centruroides sculpturatus]

Scorpion XP_023231846.1 ADP-ribosylation factor 6-like [Centruroides sculpturatus]

Scorpion XP_023231866.1 syntaxin-7-like [Centruroides sculpturatus]

Scorpion XP_023231958.1 BTB/POZ domain-containing protein kctd15-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023231961.1 BTB/POZ domain-containing protein kctd15-like isoform X6 [Centruroides sculpturatus]

Scorpion XP_023232080.1 protein tincar-like [Centruroides sculpturatus]

Scorpion XP_023232091.1 regulator of G-protein signaling loco-like [Centruroides sculpturatus]

Scorpion XP_023232144.1 ankyrin repeat domain-containing protein 29-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023232147.1 eukaryotic peptide chain release factor subunit 1-like [Centruroides sculpturatus]

Scorpion XP_023232192.1 neural cell adhesion molecule 1-A-like, partial [Centruroides sculpturatus]

Scorpion XP_023232211.1 DCN1-like protein 3 [Centruroides sculpturatus]

Scorpion XP_023232224.1 chloride intracellular channel exl-1-like [Centruroides sculpturatus]

Scorpion XP_023232246.1 uncharacterized protein LOC111632109 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023232271.1 ubiA prenyltransferase domain-containing protein 1 homolog [Centruroides sculpturatus]

Scorpion XP_023232286.1 DNA excision repair protein haywire-like [Centruroides sculpturatus]

Scorpion XP_023232313.1 DNA-directed RNA polymerase II subunit RPB2 [Centruroides sculpturatus]

Scorpion XP_023232378.1 RNA-binding protein Nova-2-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023232393.1 protein arginine N-methyltransferase 1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023232514.1 Golgi phosphoprotein 3 homolog sauron-like [Centruroides sculpturatus]

Scorpion XP_023232551.1 E3 ubiquitin-protein ligase znrf2-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023232552.1 E3 ubiquitin-protein ligase znrf2-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023232583.1 protein phosphatase 1 regulatory subunit 14C-like [Centruroides sculpturatus]

Scorpion XP_023232584.1 tyrosine-protein kinase Src42A-like [Centruroides sculpturatus]

Scorpion XP_023232595.1 protein yippee-like 5 [Centruroides sculpturatus]

Scorpion XP_023232643.1 frequenin-1-like [Centruroides sculpturatus]

Scorpion XP_023232646.1 atypical protein kinase C-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023232647.1 atypical protein kinase C-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023232650.1 rab GDP dissociation inhibitor beta-like, partial [Centruroides sculpturatus]

Scorpion XP_023232656.1 E3 ubiquitin-protein ligase arih1-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023232716.1 parafibromin-like [Centruroides sculpturatus]

Scorpion XP_023232730.1 Kv channel-interacting protein 1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023232750.1 serine palmitoyltransferase 2-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023232787.1 uncharacterized protein C20orf24 homolog [Centruroides sculpturatus]

Scorpion XP_023232887.1 influenza virus NS1A-binding protein homolog [Centruroides sculpturatus]

Scorpion XP_023232907.1 BTB/POZ domain-containing protein 2-like [Centruroides sculpturatus]

Scorpion XP_023232915.1 stomatin-2-like [Centruroides sculpturatus]

Scorpion XP_023232916.1 POU domain, class 3, transcription factor 4-like [Centruroides sculpturatus]

Scorpion XP_023232950.1 protein phosphatase PTC7 homolog [Centruroides sculpturatus]

Scorpion XP_023232964.1 ubiquitin-conjugating enzyme E2 H-like [Centruroides sculpturatus]

Scorpion XP_023232971.1 ras-related C3 botulinum toxin substrate 1-like [Centruroides sculpturatus]

Scorpion XP_023232990.1 citrate synthase, mitochondrial-like [Centruroides sculpturatus]

Scorpion XP_023232994.1 transcription factor Sox-2-like [Centruroides sculpturatus]

Scorpion XP_023233065.1 succinate--CoA ligase [ADP-forming] subunit beta, mitochondrial-like [Centruroides sculpturatus]

Scorpion XP_023233097.1 hormone receptor 4-like [Centruroides sculpturatus]

Scorpion XP_023233111.1 transmembrane protein 138-like [Centruroides sculpturatus]

Scorpion XP_023233112.1 unc-112-related protein-like [Centruroides sculpturatus]

Scorpion XP_023233152.1 moesin/ezrin/radixin homolog 1-like [Centruroides sculpturatus]

Scorpion XP_023233154.1 acetyl-CoA acetyltransferase, mitochondrial-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023233169.1 protein UBASH3A homolog [Centruroides sculpturatus]

Scorpion XP_023233182.1 bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase-like [Centruroides sculpturatus]

Scorpion XP_023233212.1 BLOC-1-related complex subunit 6-like [Centruroides sculpturatus]

Scorpion XP_023233214.1 translation initiation factor eIF-2B subunit beta-like [Centruroides sculpturatus]

Scorpion XP_023233318.1 serine/threonine-protein phosphatase 2A catalytic subunit beta isoform [Centruroides sculpturatus]

Scorpion XP_023233331.1 WD repeat domain phosphoinositide-interacting protein 2-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023233332.1 WD repeat domain phosphoinositide-interacting protein 2-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023233356.1 protein Red-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023233378.1 ankyrin repeat and KH domain-containing protein 1-like [Centruroides sculpturatus]

Scorpion XP_023233379.1 ankyrin repeat domain-containing protein 17-like [Centruroides sculpturatus]

Scorpion XP_023233382.1 AP2-associated protein kinase 1-like [Centruroides sculpturatus]

Scorpion XP_023233411.1 NEDD8 isoform X2 [Centruroides sculpturatus]

Scorpion XP_023233433.1 transcription factor Sox-5-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023233463.1 profilin-like [Centruroides sculpturatus]

Scorpion XP_023233464.1 peripheral plasma membrane protein CASK-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023233465.1 peripheral plasma membrane protein CASK-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023233467.1 eukaryotic initiation factor 4A-I-like [Centruroides sculpturatus]

Scorpion XP_023233506.1 protein TEX261-like [Centruroides sculpturatus]

Scorpion XP_023233557.1 ADP-ribosylation factor-like protein 3 [Centruroides sculpturatus]

Scorpion XP_023233568.1 F-box/SPRY domain-containing protein 1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023233581.1 ADP-ribosylation factor-like protein 6 [Centruroides sculpturatus]

Scorpion XP_023233600.1 kinesin heavy chain-like [Centruroides sculpturatus]

Scorpion XP_023233639.1 hillarin-like [Centruroides sculpturatus]

Scorpion XP_023233690.1 follistatin-related protein 5-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023233692.1 AP-1 complex subunit sigma-2-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023233697.1 follistatin-related protein 5-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023233716.1 calsyntenin-1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023233717.1 calsyntenin-1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023233740.1 serine/threonine-protein kinase STK11-like [Centruroides sculpturatus]

Scorpion XP_023233755.1 beta-1,4-glucuronyltransferase 1-like [Centruroides sculpturatus]

Scorpion XP_023233853.1 40S ribosomal protein S23 [Centruroides sculpturatus]

Scorpion XP_023233920.1 RING-box protein 1 [Centruroides sculpturatus]

Scorpion XP_023233925.1 zinc finger protein PLAGL1-like [Centruroides sculpturatus]

Scorpion XP_023233950.1 flotillin-1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023233958.1 flotillin-1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023233976.1 tankyrase-2-like [Centruroides sculpturatus]

Scorpion XP_023233978.1 neuropathy target esterase sws-like [Centruroides sculpturatus]

Scorpion XP_023233997.1 homeobox protein extradenticle-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023234007.1 homeobox protein extradenticle-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023234015.1 nuclear hormone receptor FTZ-F1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023234017.1 nuclear hormone receptor FTZ-F1-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023234061.1 60S ribosomal protein L37a [Centruroides sculpturatus]

Scorpion XP_023234070.1 uncharacterized protein LOC111633688 [Centruroides sculpturatus]

Scorpion XP_023234076.1 histone deacetylase 1-like [Centruroides sculpturatus]

Scorpion XP_023234078.1 ras-related protein Rab-10-like [Centruroides sculpturatus]

Scorpion XP_023234090.1 beta-1,4-glucuronyltransferase 1-like [Centruroides sculpturatus]

Scorpion XP_023234128.1 E3 ubiquitin-protein ligase UBR5-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023234129.1 E3 ubiquitin-protein ligase UBR5-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023234134.1 protein tiptop-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023234253.1 glutaminase liver isoform, mitochondrial-like [Centruroides sculpturatus]

Scorpion XP_023234261.1 ADP-ribosylation factor 6 [Centruroides sculpturatus]

Scorpion XP_023234297.1 protein BRICK1-like [Centruroides sculpturatus]

Scorpion XP_023234303.1 BTB/POZ domain-containing protein 3-like [Centruroides sculpturatus]

Scorpion XP_023234383.1 protein kinase C epsilon type-like [Centruroides sculpturatus]

Scorpion XP_023234425.1 leucine-rich repeat extensin-like protein 5 [Centruroides sculpturatus]

Scorpion XP_023234545.1 uncharacterized protein LOC111634100 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023234561.1 1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase-like [Centruroides sculpturatus]

Scorpion XP_023234592.1 PDZ domain-containing protein GIPC1-like [Centruroides sculpturatus]

Scorpion XP_023234600.1 protein phosphatase 1 regulatory subunit 16A-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023234608.1 uncharacterized protein LOC111634134 [Centruroides sculpturatus]

Scorpion XP_023234620.1 coronin-2B-like [Centruroides sculpturatus]

Scorpion XP_023234639.1 elongation factor 2 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023234657.1 charged multivesicular body protein 5-like [Centruroides sculpturatus]

Scorpion XP_023234722.1 5'-AMP-activated protein kinase catalytic subunit alpha-2-like [Centruroides sculpturatus]

Scorpion XP_023234775.1 eyes absent homolog 2-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023234777.1 eyes absent homolog 2-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023234807.1 ADP-ribosylation factor-related protein 1-like [Centruroides sculpturatus]

Scorpion XP_023234840.1 actin, cytoplasmic 1-like [Centruroides sculpturatus]

Scorpion XP_023234852.1 kinesin-like protein KIF2A [Centruroides sculpturatus]

Scorpion XP_023234918.1 glucose-induced degradation protein 8 homolog [Centruroides sculpturatus]

Scorpion XP_023234946.1 mediator of RNA polymerase II transcription subunit 10-like [Centruroides sculpturatus]

Scorpion XP_023234996.1 importin subunit alpha-3-like [Centruroides sculpturatus]

Scorpion XP_023235054.1 ER membrane protein complex subunit 2-B-like [Centruroides sculpturatus]

Scorpion XP_023235062.1 YTH domain-containing family protein 1-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023235067.1 cardioacceleratory peptide receptor-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023235074.1 beta-arrestin-1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023235075.1 beta-arrestin-1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023235080.1 LIM/homeobox protein Lhx9-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023235086.1 uncharacterized protein LOC111634522 isoform X1 [Centruroides sculpturatus]

Scorpion XP_023235093.1 40S ribosomal protein S15-like [Centruroides sculpturatus]

Scorpion XP_023235103.1 microtubule-associated proteins 1A/1B light chain 3C-like [Centruroides sculpturatus]

Scorpion XP_023235135.1 rho family-interacting cell polarization regulator 2-like [Centruroides sculpturatus]

Scorpion XP_023235166.1 protein-L-isoaspartate(D-aspartate) O-methyltransferase-like isoform X4 [Centruroides sculpturatus]

Scorpion XP_023235169.1 neurogenin-1-like [Centruroides sculpturatus]

Scorpion XP_023235173.1 proteasome subunit alpha type-3-like [Centruroides sculpturatus]

Scorpion XP_023235178.1 son of sevenless homolog 1-like [Centruroides sculpturatus]

Scorpion XP_023235179.1 tetratricopeptide repeat protein 8-like [Centruroides sculpturatus]

Scorpion XP_023235191.1 delta-sarcoglycan-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023235231.1 E3 ubiquitin-protein ligase MARCH5-like [Centruroides sculpturatus]

Scorpion XP_023235277.1 uncharacterized protein LOC111634674 isoform X3 [Centruroides sculpturatus]

Scorpion XP_023235326.1 eukaryotic translation initiation factor 5-like [Centruroides sculpturatus]

Scorpion XP_023235355.1 early growth response protein 1-like [Centruroides sculpturatus]

Scorpion XP_023235472.1 protein ALP1-like [Centruroides sculpturatus]

Scorpion XP_023235483.1 plexin-A2-like [Centruroides sculpturatus]

Scorpion XP_023235485.1 plexin-A2-like [Centruroides sculpturatus]

Scorpion XP_023235490.1 N-alpha-acetyltransferase 50-like [Centruroides sculpturatus]

| | | |
|----------|----------------|---|
| Scorpion | XP_023235583.1 | glyoxylate reductase/hydroxypyruvate reductase-like [Centruroides sculpturatus] |
| Scorpion | XP_023235611.1 | E3 ubiquitin-protein ligase RNF139-like [Centruroides sculpturatus] |
| Scorpion | XP_023235652.1 | 60S ribosomal protein L18-like [Centruroides sculpturatus] |
| Scorpion | XP_023235657.1 | protein singed-like [Centruroides sculpturatus] |
| Scorpion | XP_023235662.1 | nephrin-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023235731.1 | calmodulin-like protein 4 [Centruroides sculpturatus] |
| Scorpion | XP_023235757.1 | heat shock protein beta-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023235775.1 | chondroitin sulfate N-acetylgalactosaminyltransferase 2-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023235831.1 | gamma-aminobutyric acid receptor-associated protein [Centruroides sculpturatus] |
| Scorpion | XP_023235833.1 | serine/threonine-protein kinase Tao-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023235834.1 | serine/threonine-protein kinase Tao-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023235835.1 | serine/threonine-protein kinase Tao-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023235887.1 | serine/threonine-protein kinase NIM1-like [Centruroides sculpturatus] |
| Scorpion | XP_023235919.1 | myotubularin-related protein 13-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023235920.1 | myotubularin-related protein 13-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023235922.1 | myotubularin-related protein 13-like [Centruroides sculpturatus] |
| Scorpion | XP_023236009.1 | protein flightless-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023236011.1 | signal peptide peptidase-like 3 [Centruroides sculpturatus] |
| Scorpion | XP_023236172.1 | ubiquitin-conjugating enzyme E2 variant 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023236174.1 | derlin-1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023236176.1 | derlin-1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023236188.1 | ras-related protein Rab-21-like [Centruroides sculpturatus] |
| Scorpion | XP_023236253.1 | protein JTB-like [Centruroides sculpturatus] |
| Scorpion | XP_023236256.1 | creatine transporter-like [Centruroides sculpturatus] |
| Scorpion | XP_023236293.1 | probable ascorbate-specific transmembrane electron transporter 1 [Centruroides sculpturatus] |
| Scorpion | XP_023236305.1 | sodium-dependent phosphate transport protein 1, chloroplastic-like [Centruroides sculpturatus] |
| Scorpion | XP_023236409.1 | 6-phosphogluconate dehydrogenase, decarboxylating-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023236421.1 | actin-related protein 2 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023236423.1 | actin-related protein 2 isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023236433.1 | UPF0472 protein C16orf72 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023236572.1 | lissencephaly-1 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023236634.1 | unconventional myosin-XVIIIa-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023236664.1 | clathrin heavy chain-like [Centruroides sculpturatus] |
| Scorpion | XP_023236723.1 | potassium voltage-gated channel protein Shaker-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023236724.1 | potassium voltage-gated channel protein Shaker-like isoform X4 [Centruroides sculpturatus] |
| Scorpion | XP_023236730.1 | ecdysone-induced protein 78C-like [Centruroides sculpturatus] |
| Scorpion | XP_023236732.1 | vesicle transport protein GOT1B-like [Centruroides sculpturatus] |
| Scorpion | XP_023236752.1 | PWWP domain-containing protein 2A-like [Centruroides sculpturatus] |

| | | |
|----------|----------------|--|
| Scorpion | XP_023236783.1 | apoptosis-resistant E3 ubiquitin protein ligase 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023236784.1 | apoptosis-resistant E3 ubiquitin protein ligase 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023236802.1 | CUGBP Elav-like family member 4 [Centruroides sculpturatus] |
| Scorpion | XP_023236804.1 | CUGBP Elav-like family member 4 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023236814.1 | sodium- and chloride-dependent betaine transporter-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023236850.1 | matrix metalloproteinase-21-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023236853.1 | zinc finger protein 618-like [Centruroides sculpturatus] |
| Scorpion | XP_023236857.1 | matrix metalloproteinase-21-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023236904.1 | cullin-4A-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023236907.1 | arfaptin-2-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023236952.1 | cytoplasmic protein NCK2-like [Centruroides sculpturatus] |
| Scorpion | XP_023236960.1 | RING-box protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023237031.1 | protein Mo25-like [Centruroides sculpturatus] |
| Scorpion | XP_023237084.1 | glucose-fructose oxidoreductase domain-containing protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023237121.1 | tetraspanin-33-like [Centruroides sculpturatus] |
| Scorpion | XP_023237127.1 | paired box protein Pax-3-B-like [Centruroides sculpturatus] |
| Scorpion | XP_023237141.1 | GMP synthase [glutamine-hydrolyzing]-like [Centruroides sculpturatus] |
| Scorpion | XP_023237168.1 | alkaline ceramidase-like [Centruroides sculpturatus] |
| Scorpion | XP_023237190.1 | sodium- and chloride-dependent glycine transporter 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023237212.1 | band 7 protein AGAP004871-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023237241.1 | zinc finger protein 271-like [Centruroides sculpturatus] |
| Scorpion | XP_023237265.1 | probable global transcription activator SNF2L2 [Centruroides sculpturatus] |
| Scorpion | XP_023237271.1 | protein alan shepard-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023237272.1 | protein alan shepard-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023237275.1 | POU domain, class 6, transcription factor 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023237288.1 | OTU domain-containing protein 7B-like [Centruroides sculpturatus] |
| Scorpion | XP_023237319.1 | band 4.1-like protein 1 [Centruroides sculpturatus] |
| Scorpion | XP_023237341.1 | fizzy-related protein homolog isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023237342.1 | fizzy-related protein homolog isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023237376.1 | vacuolar protein sorting-associated protein 52 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023237393.1 | dynactin subunit 5-like [Centruroides sculpturatus] |
| Scorpion | XP_023237399.1 | unconventional myosin-Va-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023237440.1 | fumarate hydratase, mitochondrial-like [Centruroides sculpturatus] |
| Scorpion | XP_023237462.1 | hepatocyte nuclear factor 4-gamma-like [Centruroides sculpturatus] |
| Scorpion | XP_023237526.1 | myosin-2 essential light chain-like [Centruroides sculpturatus] |
| Scorpion | XP_023237581.1 | RNA-binding protein 8A-like [Centruroides sculpturatus] |
| Scorpion | XP_023237591.1 | ankyrin-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023237616.1 | dynein light chain 1, cytoplasmic [Centruroides sculpturatus] |
| Scorpion | XP_023237628.1 | ER lumen protein-retaining receptor-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023237629.1 | ER lumen protein-retaining receptor 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023237662.1 | protein SYS1 homolog [Centruroides sculpturatus] |

| | | |
|----------|----------------|---|
| Scorpion | XP_023237667.1 | serine-threonine kinase receptor-associated protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023237668.1 | coatomer subunit beta-like [Centruroides sculpturatus] |
| Scorpion | XP_023237669.1 | T-complex protein 1 subunit beta-like [Centruroides sculpturatus] |
| Scorpion | XP_023237701.1 | metabotropic glutamate receptor 8-like [Centruroides sculpturatus] |
| Scorpion | XP_023237714.1 | PI-PLC X domain-containing protein 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023237734.1 | disheveled-associated activator of morphogenesis 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023237743.1 | disheveled-associated activator of morphogenesis 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023237786.1 | single-strand selective monofunctional uracil DNA glycosylase-like [Centruroides sculpturatus] |
| Scorpion | XP_023237792.1 | rap guanine nucleotide exchange factor 2-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023237793.1 | rap guanine nucleotide exchange factor 2-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023237877.1 | plexin-A2-like [Centruroides sculpturatus] |
| Scorpion | XP_023237891.1 | MAGUK p55 subfamily member 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023237920.1 | uncharacterized protein LOC111636812 [Centruroides sculpturatus] |
| Scorpion | XP_023238007.1 | E3 ubiquitin-protein ligase MIB1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023238012.1 | nicalin-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023238052.1 | transforming growth factor-beta-induced protein ig-h3-like [Centruroides sculpturatus] |
| Scorpion | XP_023238053.1 | zinc finger protein Noc-like [Centruroides sculpturatus] |
| Scorpion | XP_023238060.1 | homeobox protein AKR-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023238090.1 | GTPase-activating protein and VPS9 domain-containing protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023238094.1 | nuclear receptor-binding protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023238147.1 | growth arrest-specific protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023238154.1 | uncharacterized protein LOC111636988 isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023238180.1 | transmembrane emp24 domain-containing protein eca-like [Centruroides sculpturatus] |
| Scorpion | XP_023238183.1 | uncharacterized protein LOC111637012 [Centruroides sculpturatus] |
| Scorpion | XP_023238348.1 | uncharacterized protein LOC111637156 [Centruroides sculpturatus] |
| Scorpion | XP_023238358.1 | 14-3-3 protein epsilon [Centruroides sculpturatus] |
| Scorpion | XP_023238367.1 | ubiquitin-conjugating enzyme E2 J2-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023238402.1 | transcription factor Sox-3-B-like [Centruroides sculpturatus] |
| Scorpion | XP_023238419.1 | zinc finger protein 502-like [Centruroides sculpturatus] |
| Scorpion | XP_023238425.1 | probable small nuclear ribonucleoprotein E [Centruroides sculpturatus] |
| Scorpion | XP_023238432.1 | cAMP-specific 3',5'-cyclic phosphodiesterase 7B-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023238434.1 | mannose-1-phosphate guanyltransferase beta-like [Centruroides sculpturatus] |
| Scorpion | XP_023238481.1 | uncharacterized protein LOC111637260 [Centruroides sculpturatus] |
| Scorpion | XP_023238516.1 | transient-receptor-potential-like protein isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023238624.1 | coatomer subunit delta-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023238625.1 | coatomer subunit delta-like isoform X2 [Centruroides sculpturatus] |

| | | |
|----------|----------------|--|
| Scorpion | XP_023238669.1 | ral GTPase-activating protein subunit beta-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023238670.1 | ral GTPase-activating protein subunit beta-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023238693.1 | E3 ubiquitin-protein ligase RING2-like [Centruroides sculpturatus] |
| Scorpion | XP_023238771.1 | LIM domain only protein 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023238781.1 | LIM domain only protein 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023238787.1 | receptor of activated protein C kinase 1 [Centruroides sculpturatus] |
| Scorpion | XP_023238823.1 | synaptotagmin-4-like [Centruroides sculpturatus] |
| Scorpion | XP_023238857.1 | beta-sarcoglycan-like [Centruroides sculpturatus] |
| Scorpion | XP_023238899.1 | amyloid-beta-like protein isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023238918.1 | ras-related protein Rab-30-like [Centruroides sculpturatus] |
| Scorpion | XP_023238920.1 | regulator complex protein LAMTOR1-like [Centruroides sculpturatus] |
| Scorpion | XP_023238921.1 | putative Ras-related protein Rab-33 [Centruroides sculpturatus] |
| Scorpion | XP_023238936.1 | regulator of nonsense transcripts 2-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023239062.1 | tyrosine-protein phosphatase 10D-like [Centruroides sculpturatus] |
| Scorpion | XP_023239091.1 | uncharacterized protein F36G3.2-like [Centruroides sculpturatus] |
| Scorpion | XP_023239183.1 | T-complex protein 1 subunit delta-like [Centruroides sculpturatus] |
| Scorpion | XP_023239224.1 | TWiK family of potassium channels protein 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023239365.1 | toll-like receptor Tollo [Centruroides sculpturatus] |
| Scorpion | XP_023239440.1 | integrator complex subunit 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023239478.1 | translationally-controlled tumor protein homolog [Centruroides sculpturatus] |
| Scorpion | XP_023239485.1 | anaphase-promoting complex subunit 7-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023239492.1 | anaphase-promoting complex subunit 7-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023239534.1 | 78 kDa glucose-regulated protein-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023239550.1 | uncharacterized protein LOC111638137 [Centruroides sculpturatus] |
| Scorpion | XP_023239560.1 | guanine nucleotide-binding protein G(o) subunit alpha-like [Centruroides sculpturatus] |
| Scorpion | XP_023239562.1 | protein phosphatase 1H-like [Centruroides sculpturatus] |
| Scorpion | XP_023239608.1 | transcriptional repressor p66-alpha-like [Centruroides sculpturatus] |
| Scorpion | XP_023239632.1 | cofilin/actin-depolymerizing factor homolog [Centruroides sculpturatus] |
| Scorpion | XP_023239633.1 | sister chromatid cohesion protein PDS5 homolog B-like [Centruroides sculpturatus] |
| Scorpion | XP_023239787.1 | F-box/WD repeat-containing protein 7-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023239803.1 | sulfotransferase family cytosolic 1B member 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023239809.1 | asparagine--tRNA ligase, cytoplasmic-like [Centruroides sculpturatus] |
| Scorpion | XP_023239819.1 | SUMO-conjugating enzyme UBC9-B [Centruroides sculpturatus] |
| Scorpion | XP_023239830.1 | COMM domain-containing protein 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023239834.1 | regulator of G-protein signaling 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023239870.1 | leucine-rich repeats and immunoglobulin-like domains protein 1 [Centruroides sculpturatus] |
| Scorpion | XP_023239878.1 | uncharacterized protein LOC111638404 [Centruroides sculpturatus] |
| Scorpion | XP_023239885.1 | storkhead-box protein 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023239903.1 | ubiquitin-conjugating enzyme E2 G2 [Centruroides sculpturatus] |

| | | |
|----------|----------------|---|
| Scorpion | XP_023239942.1 | C2 domain-containing protein 5-like [Centruroides sculpturatus] |
| Scorpion | XP_023239957.1 | exocyst complex component 6B-like isoform X5 [Centruroides sculpturatus] |
| Scorpion | XP_023240056.1 | probable ubiquitin-conjugating enzyme protein 17 [Centruroides sculpturatus] |
| Scorpion | XP_023240058.1 | 60S ribosomal protein L3-like [Centruroides sculpturatus] |
| Scorpion | XP_023240059.1 | inactive rhomboid protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023240071.1 | inactive rhomboid protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023240146.1 | negative elongation factor A-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023240153.1 | octopamine receptor Oamb-like [Centruroides sculpturatus] |
| Scorpion | XP_023240196.1 | glycerol-3-phosphate dehydrogenase [NAD(+)], cytoplasmic-like [Centruroides sculpturatus] |
| Scorpion | XP_023240217.1 | ras-related protein Rab-30-like [Centruroides sculpturatus] |
| Scorpion | XP_023240356.1 | leucine-rich repeat flightless-interacting protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023240360.1 | ras-related protein Rap-2c-like [Centruroides sculpturatus] |
| Scorpion | XP_023240361.1 | DNA-directed RNA polymerase II subunit RPB1-like [Centruroides sculpturatus] |
| Scorpion | XP_023240385.1 | glutamate-gated chloride channel-like [Centruroides sculpturatus] |
| Scorpion | XP_023240454.1 | vegetative cell wall protein gp1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023240463.1 | probable E3 ubiquitin-protein ligase makorin-1 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023240483.1 | tyrosine-protein phosphatase 99A-like isoform X5 [Centruroides sculpturatus] |
| Scorpion | XP_023240484.1 | dual specificity protein phosphatase 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023240489.1 | cullin-5-like [Centruroides sculpturatus] |
| Scorpion | XP_023240490.1 | charged multivesicular body protein 1b-like [Centruroides sculpturatus] |
| Scorpion | XP_023240558.1 | IDLSRF-like peptide [Centruroides sculpturatus] |
| Scorpion | XP_023240591.1 | tubulin beta chain [Centruroides sculpturatus] |
| Scorpion | XP_023240656.1 | WD repeat domain phosphoinositide-interacting protein 3-like [Centruroides sculpturatus] |
| Scorpion | XP_023240673.1 | chloride intracellular channel exc-4-like [Centruroides sculpturatus] |
| Scorpion | XP_023240700.1 | ADP-ribosylation factor-like protein 1 [Centruroides sculpturatus] |
| Scorpion | XP_023240727.1 | 60S ribosomal protein L26-like [Centruroides sculpturatus] |
| Scorpion | XP_023240750.1 | V-type proton ATPase subunit H-like [Centruroides sculpturatus] |
| Scorpion | XP_023240826.1 | single-stranded DNA-binding protein 3-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023240838.1 | THO complex subunit 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023240855.1 | autophagy-related protein 101-like [Centruroides sculpturatus] |
| Scorpion | XP_023240933.1 | tyrosine-protein kinase Src42A-like [Centruroides sculpturatus] |
| Scorpion | XP_023241000.1 | N6-adenosine-methyltransferase non-catalytic subunit-like [Centruroides sculpturatus] |
| Scorpion | XP_023241031.1 | nuclear pore complex protein Nup93-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023241032.1 | nuclear pore complex protein Nup93-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023241089.1 | protein FAM43A-like [Centruroides sculpturatus] |
| Scorpion | XP_023241106.1 | gamma-secretase subunit PEN-2-like [Centruroides sculpturatus] |
| Scorpion | XP_023241121.1 | extracellular serine/threonine protein kinase four-jointed-like [Centruroides sculpturatus] |
| Scorpion | XP_023241162.1 | 5'-AMP-activated protein kinase subunit gamma-1-like [Centruroides sculpturatus] |

| | | |
|----------|----------------|---|
| Scorpion | XP_023241186.1 | uncharacterized RNA-binding protein C23E6.01c-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023241210.1 | ras-related protein Rab-4B-like [Centruroides sculpturatus] |
| Scorpion | XP_023241212.1 | ubiquitin-like protein 3 [Centruroides sculpturatus] |
| Scorpion | XP_023241219.1 | fatty acid hydroxylase domain-containing protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023241220.1 | LOW QUALITY PROTEIN: vacuolar protein sorting-associated protein 45-like [Centruroides sculpturatus] |
| Scorpion | XP_023241366.1 | MLX-interacting protein-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023241369.1 | LHFPL tetraspan subfamily member 6 protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023241380.1 | N-acetylgalactosaminyltransferase 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023241406.1 | phosphatidylinositol 4-kinase beta-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023241426.1 | charged multivesicular body protein 4b-like [Centruroides sculpturatus] |
| Scorpion | XP_023241438.1 | GTP-binding protein REM 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023241439.1 | GTP-binding protein REM 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023241481.1 | excitatory amino acid transporter-like [Centruroides sculpturatus] |
| Scorpion | XP_023241519.1 | protein vav-like [Centruroides sculpturatus] |
| Scorpion | XP_023241650.1 | proto-oncogene tyrosine-protein kinase ROS-like [Centruroides sculpturatus] |
| Scorpion | XP_023241652.1 | phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform-like [Centruroides sculpturatus] |
| Scorpion | XP_023241701.1 | protein EFR3 homolog B-like [Centruroides sculpturatus] |
| Scorpion | XP_023241710.1 | exostosin-2-like [Centruroides sculpturatus] |
| Scorpion | XP_023241736.1 | MAP kinase-interacting serine/threonine-protein kinase 1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023241743.1 | neuropathy target esterase sws-like [Centruroides sculpturatus] |
| Scorpion | XP_023241744.1 | MAP kinase-interacting serine/threonine-protein kinase 1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023241750.1 | dystrobrevin beta-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023241805.1 | glucoside xylosyltransferase 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023241833.1 | cadherin-89D-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023241834.1 | cadherin-89D-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023241856.1 | desumoylating isopeptidase 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023241885.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit [Centruroides sculpturatus] |
| Scorpion | XP_023241912.1 | 3-phosphoinositide-dependent protein kinase 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023241989.1 | ras suppressor protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023241990.1 | T-complex protein 1 subunit theta-like [Centruroides sculpturatus] |
| Scorpion | XP_023242026.1 | calcium/calmodulin-dependent protein kinase type 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023242055.1 | ruvB-like 2 [Centruroides sculpturatus] |
| Scorpion | XP_023242072.1 | serine/threonine-protein kinase MRCK alpha-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023242079.1 | E3 ubiquitin-protein ligase RNF123-like [Centruroides sculpturatus] |
| Scorpion | XP_023242100.1 | AP-3 complex subunit beta-2-like [Centruroides sculpturatus] |
| Scorpion | XP_023242152.1 | interferon regulatory factor 2-binding protein-like A [Centruroides sculpturatus] |
| Scorpion | XP_023242153.1 | interferon regulatory factor 2-binding protein 2-B-like [Centruroides sculpturatus] |

| | | |
|----------|----------------|---|
| Scorpion | XP_023242182.1 | phosphatidylinositide phosphatase SAC1-A-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023242212.1 | protein singed-like [Centruroides sculpturatus] |
| Scorpion | XP_023242277.1 | dual specificity protein kinase splB-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023242278.1 | dual specificity protein kinase splB-like isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023242286.1 | muscle LIM protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023242328.1 | serine/threonine-protein kinase KIN2-like [Centruroides sculpturatus] |
| Scorpion | XP_023242332.1 | tyrosine-protein kinase Fyn-like [Centruroides sculpturatus] |
| Scorpion | XP_023242354.1 | lysine-specific demethylase 2B-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023242355.1 | lysine-specific demethylase 2B-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023242356.1 | lysine-specific demethylase 2B-like [Centruroides sculpturatus] |
| Scorpion | XP_023242369.1 | 26S proteasome non-ATPase regulatory subunit 11-like [Centruroides sculpturatus] |
| Scorpion | XP_023242371.1 | 26S proteasome non-ATPase regulatory subunit 11-like [Centruroides sculpturatus] |
| Scorpion | XP_023242419.1 | ran-binding protein 9-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023242469.1 | troponin C isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023242470.1 | troponin C isoform X3 [Centruroides sculpturatus] |
| Scorpion | XP_023242478.1 | protein borderless-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023242482.1 | kelch-like protein 26 [Centruroides sculpturatus] |
| Scorpion | XP_023242513.1 | E3 ubiquitin-protein ligase MGRN1-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023242517.1 | forkhead box protein N3-like [Centruroides sculpturatus] |
| Scorpion | XP_023242518.1 | netrin receptor UNC5C-like [Centruroides sculpturatus] |
| Scorpion | XP_023242598.1 | uncharacterized protein LOC111640770 [Centruroides sculpturatus] |
| Scorpion | XP_023242620.1 | actin-related protein 8-like [Centruroides sculpturatus] |
| Scorpion | XP_023242648.1 | calbindin-32-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023242649.1 | uncharacterized protein LOC111640818 [Centruroides sculpturatus] |
| Scorpion | XP_023242675.1 | selenoprotein T2-like [Centruroides sculpturatus] |
| Scorpion | XP_023242687.1 | zinc finger MYND domain-containing protein 19-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023242697.1 | COP9 signalosome complex subunit 6-like [Centruroides sculpturatus] |
| Scorpion | XP_023242722.1 | integrin-linked protein kinase-like [Centruroides sculpturatus] |
| Scorpion | XP_023242725.1 | profilin-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023242734.1 | ORM1-like protein 3 [Centruroides sculpturatus] |
| Scorpion | XP_023242775.1 | IDLSRF-like peptide [Centruroides sculpturatus] |
| Scorpion | XP_023242784.1 | ribosomal protein S6 kinase beta-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023242798.1 | dedicator of cytokinesis protein 7-like [Centruroides sculpturatus] |
| Scorpion | XP_023242829.1 | ephrin-B1-like [Centruroides sculpturatus] |
| Scorpion | XP_023242837.1 | kinesin light chain-like [Centruroides sculpturatus] |
| Scorpion | XP_023242845.1 | serine/threonine-protein kinase SIK3-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023242873.1 | rho GTPase-activating protein 33-like [Centruroides sculpturatus] |
| Scorpion | XP_023242878.1 | cell adhesion molecule 4-like, partial [Centruroides sculpturatus] |
| Scorpion | XP_023242893.1 | ras-related protein Rap-2c-like [Centruroides sculpturatus] |
| Scorpion | XP_023242895.1 | uncharacterized protein LOC111641020 [Centruroides sculpturatus] |
| Scorpion | XP_023242915.1 | zinc finger protein 16-like [Centruroides sculpturatus] |

Scorpion XP_023242978.1 translocon-associated protein subunit beta-like [Centruroides sculpturatus]

Scorpion XP_023243016.1 headcase protein-like [Centruroides sculpturatus]

Scorpion XP_023243095.1 ubiquitin carboxyl-terminal hydrolase 20-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023243144.1 syntaxin-binding protein 5-like isoform X3 [Centruroides sculpturatus]

Scorpion XP_023243172.1 RNA polymerase II subunit A C-terminal domain phosphatase SSU72-like [Centruroides sculpturatus]

Scorpion XP_023243203.1 40S ribosomal protein S5 [Centruroides sculpturatus]

Scorpion XP_023243234.1 ankyrin repeat domain-containing protein 50-like [Centruroides sculpturatus]

Scorpion XP_023243253.1 zinc finger C4H2 domain-containing protein-like [Centruroides sculpturatus]

Scorpion XP_023243255.1 tubulin gamma-1 chain [Centruroides sculpturatus]

Scorpion XP_023243310.1 heparan-sulfate 6-O-sulfotransferase 2-like [Centruroides sculpturatus]

Scorpion XP_023243338.1 integrator complex subunit 11-like [Centruroides sculpturatus]

Scorpion XP_023243349.1 WD40 repeat-containing protein SMU1 [Centruroides sculpturatus]

Scorpion XP_023243436.1 uncharacterized protein LOC111641494 [Centruroides sculpturatus]

Scorpion XP_023243528.1 uncharacterized protein LOC111641581 [Centruroides sculpturatus]

Scorpion XP_023243597.1 gastrula zinc finger protein XICGF8.2DB-like [Centruroides sculpturatus]

Scorpion XP_023243605.1 armadillo segment polarity protein-like [Centruroides sculpturatus]

Scorpion XP_023243607.1 transcription elongation factor SPT4-like [Centruroides sculpturatus]

Scorpion XP_023243645.1 FMRamide receptor-like [Centruroides sculpturatus]

Scorpion XP_023243657.1 protein disulfide-isomerase A3-like [Centruroides sculpturatus]

Scorpion XP_023243701.1 dual specificity tyrosine-phosphorylation-regulated kinase 1A-like [Centruroides sculpturatus]

Scorpion XP_023243702.1 dual specificity tyrosine-phosphorylation-regulated kinase 1A-like [Centruroides sculpturatus]

Scorpion XP_023243723.1 nuclear inhibitor of protein phosphatase 1-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023243724.1 nuclear inhibitor of protein phosphatase 1-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023243749.1 histone H2B-like [Centruroides sculpturatus]

Scorpion XP_023243774.1 G protein-activated inward rectifier potassium channel 3-like [Centruroides sculpturatus]

Scorpion XP_023243779.1 BTB/POZ domain-containing protein KCTD5-like [Centruroides sculpturatus]

Scorpion XP_023243827.1 tyrosine-protein phosphatase 10D-like [Centruroides sculpturatus]

Scorpion XP_023243829.1 transmembrane and TPR repeat-containing protein 2-like [Centruroides sculpturatus]

Scorpion XP_023243886.1 ceramide glucosyltransferase-like, partial [Centruroides sculpturatus]

Scorpion XP_023243980.1 uncharacterized protein LOC111641959 [Centruroides sculpturatus]

Scorpion XP_023244025.1 eukaryotic initiation factor 4A-III [Centruroides sculpturatus]

Scorpion XP_023244095.1 26S proteasome regulatory subunit 10B [Centruroides sculpturatus]

Scorpion XP_023244100.1 succinate--CoA ligase [ADP/GDP-forming] subunit alpha, mitochondrial-like [Centruroides sculpturatus]

Scorpion XP_023244110.1 histone acetyltransferase KAT6A-like isoform X1 [Centruroides sculpturatus]

Scorpion XP_023244111.1 histone acetyltransferase KAT6A-like isoform X2 [Centruroides sculpturatus]

Scorpion XP_023244163.1 transmembrane protein 184B-like [Centruroides sculpturatus]

Scorpion XP_023244225.1 eukaryotic translation initiation factor 5B-like [Centruroides sculpturatus]

Scorpion XP_023244230.1 ceramide synthase 6-like [Centruroides sculpturatus]

Scorpion XP_023244235.1 hypoxanthine-guanine phosphoribosyltransferase-like isoform X3 [Centruroides sculpturatus]

| | | |
|---------------|----------------------------------|--|
| Scorpion | XP_023244255.1 | homeobox protein prospero-like [Centruroides sculpturatus] |
| Scorpion | XP_023244287.1 | DNA topoisomerase 3-beta-1-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023244294.1 | cysteine-rich hydrophobic domain-containing protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023244379.1 | oxysterol-binding protein-related protein 2-like [Centruroides sculpturatus] |
| Scorpion | XP_023244450.1 | ras-related GTP-binding protein A [Centruroides sculpturatus] |
| Scorpion | XP_023244457.1 | lachesin-like [Centruroides sculpturatus] |
| Scorpion | XP_023244479.1 | tyrosine-protein kinase CSK-like [Centruroides sculpturatus] |
| Scorpion | XP_023244494.1 | D-glucuronyl C5-epimerase-like [Centruroides sculpturatus] |
| Scorpion | XP_023244522.1 | protein zyg-11 homolog B-like [Centruroides sculpturatus] |
| Scorpion | XP_023244523.1 | hippocampus abundant transcript 1 protein-like [Centruroides sculpturatus] |
| Scorpion | XP_023244556.1 | large proline-rich protein BAG6-like [Centruroides sculpturatus] |
| Scorpion | XP_023244574.1 | gamma-secretase subunit Aph-1-like [Centruroides sculpturatus] |
| Scorpion | XP_023244585.1 | lachesin-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023244587.1 | lachesin-like isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023244602.1 | uncharacterized protein LOC111642472 isoform X2 [Centruroides sculpturatus] |
| Scorpion | XP_023244619.1 | cell division cycle protein 27 homolog [Centruroides sculpturatus] |
| Scorpion | XP_023244624.1 | dr1-associated corepressor-like [Centruroides sculpturatus] |
| Scorpion | XP_023244645.1 | V-type proton ATPase catalytic subunit A-like [Centruroides sculpturatus] |
| Scorpion | XP_023244650.1 | G-protein coupled receptor moody-like [Centruroides sculpturatus] |
| Scorpion | XP_023244659.1 | erbin-like [Centruroides sculpturatus] |
| Scorpion | XP_023244699.1 | 40S ribosomal protein S13, partial [Centruroides sculpturatus] |
| Scorpion | XP_023244702.1 | histone H2B-like [Centruroides sculpturatus] |
| Scorpion | XP_023244735.1 | CCR4-NOT transcription complex subunit 9-like isoform X1 [Centruroides sculpturatus] |
| Scorpion | XP_023244820.1 | DNA damage-regulated autophagy modulator protein 1-like [Centruroides sculpturatus] |
| Scorpion | XP_023244835.1 | proteasome subunit alpha type-6-like [Centruroides sculpturatus] |
| Scorpion | XP_023244844.1 | cytoplasmic polyadenylation element-binding protein 2, partial [Centruroides sculpturatus] |
| T. serrulatus | JK732409 | Antimicrobial peptide (Alvarenga <i>et al.</i> , 2012) |
| T. serrulatus | Non published | c255_Tese_Ana_Paula_Putative_NaToxin4 (Martins, 2016) |
| T. serrulatus | Non published | c4862_Tese_Ana_Paula_Putative_NaToxin7 (Martins, 2016) |
| T. serrulatus | Non published | c4748_Tese_Ana_Paula_Putative_NaToxin12 (Martins, 2016) |
| T. serrulatus | Non published | c1432_Tese_Ana_Paula_Putative_NaToxin13 (Martins, 2016) |
| T. serrulatus | gi1214567028 | putative_metalloproteinase |
| T. serrulatus | A0A076L876 | Metalloserrulase1 (CARMO <i>et al.</i> , 2014) |
| T. serrulatus | A0A076L332 | Metalloserrulase4 (CARMO <i>et al.</i> , 2014) |
| T. serrulatus | (Bertazzi, 2003) | MetalloProtease_MQ7_1_2_aD_Meio (Bertazzi, 2003) |
| T. serrulatus | (Alvarenga <i>et al.</i> , 2012) | Ts17 |
| T. serrulatus | (Alvarenga <i>et al.</i> , 2012) | Ts18 |
| T. serrulatus | O77463 | Ts4 |
| Toxin | ABC94584.1 | protein phosphatase 1 catalytic subunit beta isoform [Scophthalmus maximus] |
| Toxin | AEA29698.1 | actin [Trichoplusia ni] |
| Toxin | EEN69223.1 | hypothetical protein BRAFLDRAFT_210558 [Branchiostoma floridae] |

| | | |
|-------|----------------|---|
| Toxin | ESO90163.1 | hypothetical protein LOTGIDRAFT_192316 [<i>Lottia gigantea</i>] |
| Toxin | JAA72107.1 | Putative serine/threonine specific protein phosphatase pp1 catalytic subunit [<i>Ixodes ricinus</i>] |
| Toxin | JAH06761.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAH13157.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAH31103.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAH89569.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAH94639.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAH97332.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAH98479.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAH98512.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAH98720.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | JAI08153.1 | hypothetical protein [<i>Anguilla anguilla</i>] |
| Toxin | KFO65212.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Corvus brachyrhynchos</i>] |
| Toxin | KFO78653.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Cuculus canorus</i>] |
| Toxin | KFO79592.1 | Serine/threonine-protein phosphatase PP1-alpha catalytic subunit, partial [<i>Cuculus canorus</i>] |
| Toxin | KFO92785.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Buceros rhinoceros silvestris</i>] |
| Toxin | KFP02935.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Calypte anna</i>] |
| Toxin | KFP45124.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Chlamydotis macqueenii</i>] |
| Toxin | KFQ28594.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Merops nubicus</i>] |
| Toxin | KFQ56349.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Nestor notabilis</i>] |
| Toxin | KFR06461.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Opisthocomus hoazin</i>] |
| Toxin | KFU95142.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Chaetura pelagica</i>] |
| Toxin | KFV09597.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Pterocles gutturalis</i>] |
| Toxin | KFV70279.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Picooides pubescens</i>] |
| Toxin | KFV81991.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Struthio camelus australis</i>] |
| Toxin | KFW89026.1 | Serine/threonine-protein phosphatase PP1-beta catalytic subunit, partial [<i>Manacus vitellinus</i>] |
| Toxin | NP_001004527.2 | serine/threonine-protein phosphatase PP1-beta catalytic subunit [<i>Danio rerio</i>] |
| Toxin | NP_001279437.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit [<i>Callorhinchus milii</i>] |
| Toxin | pdb 3EL2 A | Chain A, 1 Actin-5c |
| Toxin | pdb 4V0X A | Chain A, 1 Protein Phosphatase Pp1-gamma Catalytic Subunit Molecule: Serine/threonine-protein Phosphatase Pp1-gamma Catalytic Subunit |
| Toxin | tr A0A0F7YZ82 | Tubulin alpha chain OS= <i>Micrurus fulvius</i> OX=8637 PE=2 SV=1 |
| Toxin | tr A0A0F7Z029 | Tubulin beta chain OS= <i>Micrurus fulvius</i> OX=8637 PE=2 SV=1 |
| Toxin | tr A0A0F7Z1T6 | Tubulin alpha chain OS= <i>Crotalus adamanteus</i> OX=8729 PE=2 SV=1 |
| Toxin | tr A0A0F7Z477 | Tubulin beta chain OS= <i>Micrurus fulvius</i> OX=8637 PE=2 SV=1 |

| | | |
|-------|----------------|--|
| Toxin | tr A0A2I9LPI4 | Ras-like GTP binding protein OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Toxin | tr A0A2I9LPL9 | 60S ribosomal protein L39 OS=Centruroides hentzi OX=88313 PE=4 SV=1 |
| Toxin | tr A7XZV7 | Beta-actin (Fragment) OS=Siganus canaliculatus OX=75042 PE=2 SV=1 |
| Toxin | tr J3S1F3 | Tubulin alpha chain OS=Crotalus adamanteus OX=8729 PE=2 SV=1 |
| Toxin | tr J3S3X5 | Actin, cytoplasmic 1 OS=Crotalus adamanteus OX=8729 PE=2 SV=1 |
| Toxin | tr J3S8A3 | Calmodulin OS=Crotalus adamanteus OX=8729 PE=2 SV=1 |
| Toxin | tr J3SBS6 | ADP-ribosylation factor 6-like OS=Crotalus adamanteus OX=8729 PE=2 SV=1 |
| Toxin | tr U3ET38 | GABA OS=Micrurus fulvius OX=8637 PE=2 SV=1 |
| Toxin | tr U3F5T7 | Tubulin alpha chain OS=Micrurus fulvius OX=8637 PE=2 SV=1 |
| Toxin | tr U3FXC9 | Calmodulin OS=Micrurus fulvius OX=8637 PE=2 SV=1 |
| Toxin | tr U3FZL5 | Actin, cytoplasmic 2 OS=Micrurus fulvius OX=8637 PE=2 SV=1 |
| Toxin | XP_002731251.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit [Saccoglossus kowalevskii] |
| Toxin | XP_005043209.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit isoform X3 [Ficedula albicollis] |
| Toxin | XP_005103447.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit [Aplysia californica] |
| Toxin | XP_005798385.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit [Xiphophorus maculatus] |
| Toxin | XP_006115676.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit [Pelodiscus sinensis] |
| Toxin | XP_007257657.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit-like [Astyanax mexicanus] |
| Toxin | XP_008292125.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit [Stegastes partitus] |
| Toxin | XP_010794613.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit isoform X2 [Notothenia coriiceps] |
| Toxin | XP_011441164.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit [Crassostrea gigas] |
| Toxin | XP_011604760.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit-like [Takifugu rubripes] |
| Toxin | XP_012507965.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit [Propithecus coquereli] |
| Toxin | XP_013067120.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit [Biomphalaria glabrata] |
| Toxin | XP_013391353.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit [Lingula anatina] |
| Toxin | XP_013776160.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit [Limulus polyphemus] |
| Toxin | XP_013875462.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit-like [Austrofundulus limnaeus] |
| Toxin | XP_014967413.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit isoform X2 [Macaca mulatta] |
| Toxin | XP_015825270.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit [Nothobranchius furzeri] |
| Toxin | XP_019728112.1 | PREDICTED: serine/threonine-protein phosphatase PP1-beta catalytic subunit-like [Hippocampus comes] |
| Toxin | XP_020483811.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit-like [Labrus bergylta] |
| Toxin | XP_020502093.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit [Labrus bergylta] |

| | | |
|-----------------|----------------|--|
| Toxin | XP_022620270.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit-like [Seriola dumerili] |
| Toxin | XP_023140794.1 | serine/threonine-protein phosphatase PP1-beta catalytic subunit-like [Amphiprion ocellaris] |
| Arachnida Gland | tr A0A076KU80 | BLTX560 OS=Nephila pilipes OX=299642 PE=2 SV=1 |
| Arachnida Gland | tr A0A076KZF3 | BLTX364 OS=Nephila pilipes OX=299642 PE=2 SV=1 |
| Arachnida Gland | tr A0A076L0D3 | BLTX538 OS=Nephila pilipes OX=299642 PE=2 SV=1 |
| Arachnida Gland | tr A0A0A0V781 | Uncharacterized protein OS=Scytodes thoracica OX=1112478 PE=2 SV=1 |
| Arachnida Gland | tr A0A0A0VC99 | Actin (Fragment) OS=Scytodes thoracica OX=1112478 PE=2 SV=1 |
| Arachnida Gland | tr A0A0A0VCC6 | Actin (Fragment) OS=Scytodes thoracica OX=1112478 PE=2 SV=1 |
| Arachnida Gland | tr A0A0C9QZR2 | Putative 40s ribosomal protein s23 (Fragment) OS=Amblyomma americanum OX=6943 PE=2 SV=1 |
| Arachnida Gland | tr A0A0C9RSV3 | Putative actin-related protein arp2/3 complex subunit arpc4 (Fragment) OS=Amblyomma americanum OX=6943 PE=2 SV=1 |
| Arachnida Gland | tr A0A0C9RV12 | Putative atp-dependent rna helicase (Fragment) OS=Amblyomma americanum OX=6943 PE=2 SV=1 |
| Arachnida Gland | tr A0A0C9S2A3 | Serine/threonine-protein phosphatase (Fragment) OS=Amblyomma americanum OX=6943 PE=2 SV=1 |
| Arachnida Gland | tr A0A0C9S9S6 | Putative 60s ribosomal protein l23 OS=Amblyomma americanum OX=6943 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8R3B2 | Signal recognition particle 9 kDa protein OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8R6X2 | Putative effete OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8R9R4 | AP complex subunit sigma OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8RC33 | Putative 40s ribosomal protein s23 OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8RD36 | Putative adaptor complexes medium subunit family OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8RH12 | Actin-related protein 2/3 complex subunit 4 OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8RIZ9 | AP complex subunit sigma OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8RKQ9 | Tubulin beta chain OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8RL15 | Putative gtp-binding adp-ribosylation factor arf1 OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A0K8RR67 | Putative ubiquitin protein ligase (Fragment) OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr A0A131YHF0 | DNA-directed RNA polymerase subunit beta OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A131YJW2 | Serine/threonine-protein phosphatase OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A131YKP7 | Serine/threonine-protein phosphatase OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A131YTJ4 | Ubiquitin-conjugating enzyme E2 G2 OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |

| | | |
|-----------------|---------------|---|
| Arachnida Gland | tr A0A131YUC8 | Histone H3 OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A131YZ77 | Histone-binding protein RBBP4 OS=Rhipicephalus appendiculatus OX=34631 PE=4 SV=1 |
| Arachnida Gland | tr A0A131YZP9 | Protein transport protein SEC61 subunit alpha OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A131Z1Y2 | ADP-ribosylation factor 6 OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A131Z485 | GABA(A) receptor-associated protein OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A131Z5K9 | AP-2 complex subunit mu-1 OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A131Z5Q6 | ATP-dependent RNA helicase OS=Rhipicephalus appendiculatus OX=34631 PE=3 SV=1 |
| Arachnida Gland | tr A0A1D2AI23 | Putative histone binding protein caf1 like (Fragment) OS=Ornithodoros brasiliensis OX=888526 PE=4 SV=1 |
| Arachnida Gland | tr A0A1D2AIE8 | Serine threonine protein phosphatase 2a catalytic subunit alpha isoform (Fragment) OS=Ornithodoros brasiliensis OX=888526 PE=4 SV=1 |
| Arachnida Gland | tr A0A1D2AIP0 | Tubulin beta chain (Fragment) OS=Ornithodoros brasiliensis OX=888526 PE=3 SV=1 |
| Arachnida Gland | tr A0A1D2AJG7 | Putative histone binding protein caf1 like (Fragment) OS=Ornithodoros brasiliensis OX=888526 PE=4 SV=1 |
| Arachnida Gland | tr A0A1E1XJY2 | AP complex subunit sigma OS=Amblyomma sculptum OX=1581419 PE=2 SV=1 |
| Arachnida Gland | tr A0A1E1XK16 | Putative ubiquitin-protein ligase OS=Amblyomma sculptum OX=1581419 PE=2 SV=1 |
| Arachnida Gland | tr A0A1E1XMN9 | Putative histone H3.3-like isoform x2 OS=Amblyomma sculptum OX=1581419 PE=2 SV=1 |
| Arachnida Gland | tr A0A1E1XN62 | Putative actin-2 (Fragment) OS=Amblyomma sculptum OX=1581419 PE=2 SV=1 |
| Arachnida Gland | tr A0A1E1XNB3 | Putative g protein OS=Amblyomma sculptum OX=1581419 PE=2 SV=1 |
| Arachnida Gland | tr A0A1E1XRZ9 | Putative g-protein beta subunit OS=Amblyomma sculptum OX=1581419 PE=2 SV=1 |
| Arachnida Gland | tr A0A1E1XU50 | Protein yippee-like OS=Amblyomma sculptum OX=1581419 PE=2 SV=1 |
| Arachnida Gland | tr A0A1E1XUH2 | Actin-related protein 2/3 complex subunit 4 OS=Amblyomma sculptum OX=1581419 PE=2 SV=1 |
| Arachnida Gland | tr A6N9Z3 | 60s ribosomal protein L39 OS=Ornithodoros parkeri OX=140564 PE=2 SV=1 |
| Arachnida Gland | tr A9QQ31 | Actin 5C OS=Lycosa singoriensis OX=434756 PE=2 SV=1 |
| Arachnida Gland | tr A9QQ54 | Ribosomal protein l37a OS=Lycosa singoriensis OX=434756 PE=2 SV=1 |
| Arachnida Gland | tr A9QQ76 | 60S ribosomal protein l26 OS=Lycosa singoriensis OX=434756 PE=2 SV=1 |
| Arachnida Gland | tr A9QQB6 | Ribosomal protein l23 OS=Lycosa singoriensis OX=434756 PE=2 SV=1 |
| Arachnida Gland | tr D5J6Z7 | Putative ribosomal protein S23 OS=Pelionobius muticus OX=753628 PE=2 SV=1 |
| Arachnida Gland | tr F0J8Z5 | Vesicle coat complex COPI beta' subunit (Fragment) OS=Amblyomma variegatum OX=34610 PE=2 SV=1 |
| Arachnida Gland | tr F0J9B1 | Serine/threonine-protein phosphatase (Fragment) OS=Amblyomma variegatum OX=34610 PE=2 SV=1 |

| | | |
|-----------------|-----------|--|
| Arachnida Gland | tr F0J9C2 | Transport protein Sec61 alpha subunit (Fragment) OS=Amblyomma variegatum OX=34610 PE=2 SV=1 |
| Arachnida Gland | tr F0J9P6 | Heat shock 70 kDa protein (Fragment) OS=Amblyomma variegatum OX=34610 PE=2 SV=1 |
| Arachnida Gland | tr F0J9U9 | RAS oncogene family member RAB2 OS=Amblyomma variegatum OX=34610 PE=2 SV=1 |
| Arachnida Gland | tr F1CJ93 | Putative zinc-finger containing protein (Fragment) OS=Hottentotta judaicus OX=6863 PE=2 SV=1 |
| Arachnida Gland | tr G3MK37 | Uncharacterized protein OS=Amblyomma maculatum OX=34609 PE=2 SV=1 |
| Arachnida Gland | tr G3MLM7 | Uncharacterized protein OS=Amblyomma maculatum OX=34609 PE=2 SV=1 |
| Arachnida Gland | tr G3MLY4 | Uncharacterized protein OS=Amblyomma maculatum OX=34609 PE=2 SV=1 |
| Arachnida Gland | tr G3MMX5 | Histone H3 OS=Amblyomma maculatum OX=34609 PE=2 SV=1 |
| Arachnida Gland | tr G3MPX6 | Uncharacterized protein OS=Amblyomma maculatum OX=34609 PE=2 SV=1 |
| Arachnida Gland | tr G3MRJ3 | Uncharacterized protein OS=Amblyomma maculatum OX=34609 PE=2 SV=1 |
| Arachnida Gland | tr G3MT67 | Protein yippee-like OS=Amblyomma maculatum OX=34609 PE=2 SV=1 |
| Arachnida Gland | tr L7LY75 | Actin-related protein 2/3 complex subunit 4 OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M2R8 | Putative nucleosome remodeling factor subunit OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M337 | Dynein light chain OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M339 | AP complex subunit sigma OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M345 | Putative ubiquitin-protein ligase OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M547 | Putative ca ²⁺ sensor ef-hand superfamily OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M5U6 | U6 snRNA-associated Sm-like protein LSm2 OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M6G3 | AP complex subunit sigma OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M7P7 | Putative eif4aiii OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M8J7 | Putative actin OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7M958 | Putative g protein beta-subunit 13f OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr L7MK93 | Putative ubiquitin protein ligase (Fragment) OS=Rhipicephalus pulchellus OX=72859 PE=2 SV=1 |
| Arachnida Gland | tr Q09JN6 | ADP ribosylation factor 79F OS=Argas monolakensis OX=34602 PE=2 SV=1 |
| Arachnida Gland | tr Q4PLX6 | ADP-ribosylation factor 1 OS=Ixodes scapularis OX=6945 PE=2 SV=1 |
| Arachnida Gland | tr V5GFU2 | Putative voltage-dependent l-type calcium channel subunit alpha-1d (Fragment) OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr V5HJX6 | Putative ras-related protein rab-14 OS=Ixodes ricinus OX=34613 PE=2 SV=1 |

| | | |
|---------------------|----------------|---|
| Arachnida Gland | tr V5HKJ7 | Putative 40s ribosomal protein OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr V5HQC6 | Putative ubiquitin-conjugating enzyme e2 w (Fragment) OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr V5HR07 | Tubulin alpha chain (Fragment) OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr V5HRZ9 | Putative histone h3.3 (Fragment) OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr V5I100 | Putative 60s ribosomal protein l23 apis mellifera (Fragment) OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr V5IF81 | Putative actin-2 OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Gland | tr V5IK62 | Putative adp-ribosylation factor 1 OS=Ixodes ricinus OX=34613 PE=2 SV=1 |
| Arachnida Muscle | XP_023216322.1 | myosin heavy chain, embryonic smooth muscle isoform-like [Centruroides sculpturatus] |
