



UNIVERSIDADE FEDERAL DO CEARÁ  
FFOE – FACULDADE DE FARMÁCIA, ODONTOLOGIA E ENFERMAGEM  
DEPARTAMENTO DE CIÊNCIAS FARMACÊUTICAS  
PROGRAMA DE PÓS GRADUAÇÃO EM DESENVOLVIMENTO E INOVAÇÃO  
TECNOLÓGICA DE MEDICAMENTOS

**ALBA FABIOLA COSTA TORRES**

**PEÇONHA DE *DINOPONERA QUADRICEPS* COMO FONTE DE SUBSTÂNCIAS  
BIOATIVAS - UMA ABORDAGEM TRANSCRIPTOMICA E PEPTIDOMICA**

FORTALEZA – CE

2013

Dados Internacionais de Catalogação na Publicação

Universidade Federal do Ceará

Biblioteca de Ciências da Saúde

---

T643p Torres, Alba Fabíola Costa.

Peçonha de *dinoponera quadriceps* como fonte de substâncias bioativas. / Alba Fabíola Costa Torres. – 2013.

148 f.: il. color., enc.; 30 cm.

Tese (doutorado) – Universidade Federal do Ceará; Faculdade de Farmácia, Odontologia e Enfermagem; Programa de Pós-Graduação em Desenvolvimento e Inovação Tecnológica de Medicamentos; Doutorado em Desenvolvimento e Inovação Tecnológica de Medicamentos, Fortaleza, 2013.

Área de Concentração: Biotecnologia.

Orientação: Profa. Dra. Alice Maria Costa Martins.

1. Venenos de Formiga. 2. Transcriptoma. 3. Peptídeo. I. Título.

---

CDD 615.94

**ALBA FABIOLA COSTA TORRES**

**PEÇONHA DE *DINOPONERA QUADRICEPS* COMO FONTE DE SUBSTÂNCIAS  
BIOATIVAS – UMA ABORDAGEM TRANSCRIPTOMICA E PEPTIDOMICA**

Tese de Doutorado apresentada ao Programa de Pós-graduação em Desenvolvimento e Inovação Tecnológica de Medicamentos, da Faculdade de Farmácia, Odontologia e Enfermagem da Universidade Federal do Ceará como requisito para obtenção do título de Doutor em Desenvolvimento e Inovação Tecnológica de Medicamentos.

Orientadora: Profa Dra Alice Maria Costa Martins

FORTALEZA – CE

2013

**ALBA FABIOLA COSTA TORRES**

**PEÇONHA DE *DINOPONERA QUADRICEPS* COMO FONTE DE SUBSTÂNCIAS  
BIOATIVAS - UMA ABORDAGEM TRANSCRIPTOMICA E PEPTIDOMICA**

Tese de Doutorado apresentada ao Programa de Pós-graduação em Desenvolvimento e Inovação Tecnológica de Medicamentos, da Faculdade de Farmácia, Odontologia e Enfermagem da Universidade Federal do Ceará como requisito para obtenção do título de Doutor em Desenvolvimento e Inovação Tecnológica de Medicamentos.

Aprovada em: \_\_\_/\_\_\_/\_\_\_

**BANCA EXAMINADORA**

---

Profa. Dra. Alice Maria Costa Martins  
Universidade Federal do Ceará

---

Prof. Dr. Gandhi Rádis-Baptista  
Universidade Federal do Ceará

---

Prof. Dr. Alexandre Havt Bindá  
Universidade Federal do Ceará

---

Dr. Álvaro Rossan de Brandão Prieto da Silva  
Instituto Butantã

---

Dr. Bruno Anderson Matias da Rocha  
Universidade Federal do Ceará

## DEDICATÓRIA

Aos meus pais, pelo exemplo de vida, dedicação e carinho.

Aos meus irmãos, pelo companheirismo e constante apoio.

Aos meus amigos, pela força e incentivo que sempre depositaram em mim.

## AGRADECIMENTOS

À minha família, pelo amor e carinho que sempre me fortalece. Mesmo distante vocês sempre estiveram presentes em minha mente e coração. Vocês são a razão de tudo que sou e de tudo que me tornarei, em vocês me fortaleço e sei também que vocês se fortalecem com o meu sucesso. O meu amor por vocês é indescritível, sem limites e incondicional;

A Profa. Alice Maria Costa Martins, pela aceitação, dedicação e apoio fundamental durante a realização deste trabalho;

Aos Professores Gandhi Rádhis-Baptista, Alexandre Havt Bindá, Yves Patric Quinet e aos colaboradores Álvaro Rossan de Brandão Prieto da Silva, Katsuhiko Konno, Simon Lee e André Junqueira Zaharenko pela fundamental parceria e colaboração;

À secretária Raimundinha pelo constante apoio e amizade;

Aos colegas de laboratório e pesquisa Ticiane, Rodrigo, Ramon, Kamila, Felipe, Gdayllon, Dânya, Clarissa, Izabel, Paloma, Jader, Louise, Patrícia pela constante amizade, apoio e companhia durante a realização deste trabalho.

Ao apoio financeiro da Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico/FUNCAP e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior/CAPES.

“Lembrem-se de que a ciência exige que as pessoas se dediquem a ela durante a vida inteira. E se tivessem duas vidas, ainda assim não seria suficiente. A ciência demanda dos indivíduos grande tensão e forte paixão. Sejam apaixonados por suas ciências e por suas pesquisas.”

Ivan Pavlov

## RESUMO

*Dinoponera quadriceps* é uma formiga gigante, predatória que habita regiões tropicais e submete suas presas injetando um coquetel de substâncias tóxicas. Acidentes em humanos ocorrem ocasionalmente, e podem causar dor e inflamação local, além de sintomas sistêmicos. O estudo detalhado do transcriptoma da glândula de veneno de *D. quadriceps* é requerido para esclarecer sobre o repertório de toxinas que possui, com objetivo de avaliar o potencial biotecnológico destas substâncias, assim como compreender as bases fisiopatológicas do envenenamento por Himenópteros. Realizou-se a análise transcriptômica da biblioteca de cDNA da glândula de veneno de *D. quadriceps* através de sequenciamento pelo método Sanger em combinação com sequenciamento utilizando a tecnologia *Ion Torrent*. A partir da biblioteca de cDNA, 420 clones independentes foram analisados, resultando em 21 contigs e 39 singletos. Embora a proporção de dinoponeratoxinas tenha sido bastante elevada, foi possível descrever a primeira toxina ICK-símile de formiga. Através do sequenciamento *Ion Torrent* um total de 2.514.767 leituras foi alcançado. Estas foram agrupadas em 18.546 contigs. A avaliação dos contigs e singletos utilizando a ferramenta Blast contra banco de dados de proteínas não redundantes e Swiss-Prot mostrou que 73,27% dos transcritos obtidos por NGS e 71% dos obtidos por Sanger apresentaram homólogos, indicando uma interessante diversidade de transcritos relacionados à expressão gênica dos constituintes da peçonha. A maioria das sequências codificam um conjunto de polipeptídios mais abundantes, os quais compreendem substâncias alergênicas, proteínas letais e esterases; e um conjunto de polipeptídios menos abundantes como toxinas estruturalmente conservadas ricas em cisteínas. Uma elevada proporção de transcritos, sejam obtidos pelo método Sanger ou àqueles obtidos pelo método *Ion Torrent*, foram classificados como proteínas hipotéticas ou não encontrou-se homólogos com sequências conhecidas, não sendo possível categorizá-los. Para complementar a investigação realizou-se a análise peptidômica da peçonha bruta de *D. quadriceps*, através da qual foi possível comparar peptídeos presentes na peçonha com as sequências de transcritos. Assim, foi possível confirmar a presença de toxinas maduras, cujos precursores foram revelados através da análise transcriptômica, bem como foi possível prever a forma de processamento dos precursores dos transcritos e modificações pós-traducionais. Este é o primeiro relato do transcriptoma da glândula de veneno da formiga gigante *D. quadriceps*. O sistema glandular foi dissecado e o repertório de toxinas revelado, trazendo à luz novas sequências que incluem toxinas ICK-símile, proteínas alergênicas, esterases, etc. Estes achados contribuem para a compreensão da ecologia, do comportamento e principalmente da venômica de Himenópteros.

**Palavras-chave:** veneno de formiga, transcriptoma, peptídeo

## ABSTRACT

*Dinoponera quadriceps* is a predatory giant ant that inhabits the tropical regions and subdues its prey (insects) with stings that deliver a toxic cocktail of molecules. Human accidents occasionally occur and cause local inflammation, pain and systemic symptoms. A comprehensive study of the *D. quadriceps* venom gland transcriptome is required to advance our knowledge about the toxin repertoire of the giant ant venom to evaluate the presence of substances with biotechnological value and to understand the physiopathological basis of Hymenoptera envenomation. We conducted a transcriptome analysis of a cDNA library from the *D. quadriceps* venom gland with Sanger sequencing in combination with whole-transcriptome shotgun deep sequencing. From the cDNA library using Sanger methodology, a total of 420 independent clones were analyzed resulting in 21 contigs and 39 singlets. Although the proportion of dinoponeratoxin isoform precursors was high, the first giant ant venom inhibitor cysteine-knot (ICK) toxin was found. The deep next generation sequencing yielded a total of 2.514.767 raw reads that were assembled into 18.546 contigs. A Blast search of the assembled contigs against non-redundant and Swiss-Prot databases showed that 73,27% of transcripts from NGS and 71% from Sanger corresponded to hits and indicated an interesting diversity of transcripts related to venom gene expression. The majority of these venom-related sequences code for a major polypeptide core, which comprises venom allergens, lethal-like proteins and esterases, and a minor peptide framework composed of inter-specific structurally conserved cysteine-rich toxins. Both cDNA library and deep sequencing yielded large proportions of contigs that showed no similarities with known sequences or were classified as hypothetical proteins. To improve the investigation, we performed a proteomic analysis with the crude venom of *D. quadriceps*. It was possible to confirm the presence of mature toxins evidenced through the transcriptome and predict the precursors processing and post-translational modifications. To our knowledge, this is the first report of the venom gland transcriptome of the New World giant ant *D. quadriceps*. The glandular venom system was dissected, and the toxin arsenal was revealed; this process brought to light novel sequences that included an ICK-folded toxins, allergen proteins, esterases. These findings contribute to the understanding the venom constituents the ecology, behavior and venomics of hymenopterans.

**Key words:** ant venom, transcriptomic, peptide

## LISTA DE ILUSTRAÇÕES

<b>Figura 1:</b> <i>Dinoponera quadriceps</i> .....	19
<b>Figura 2:</b> Aparato secretório de <i>D. quadriceps</i> .....	20
<b>Figura 3:</b> Esquema da síntese do DNA de <i>D. quadriceps</i> .....	31
<b>Figura 4:</b> Detalhamento do vetor pSMART2IFD.....	33
<b>Figura 5:</b> Sequencia linear do sitio de clonagem do vetor pSMART2IFD.....	34
<b>Figura 6:</b> Esquema simplificado da construção da biblioteca de cDNA.....	36
<b>Figura 7:</b> Eletroforese do RNA total extraído e do cDNA dupla fita.....	44
<b>Figura 8:</b> Perfil de fracionamento do cDNA dupla fita obtido a partir do RNA total extraído da glândula de veneno de <i>D. quadriceps</i> .....	45
<b>Figura 9:</b> Eletroforese em gel de agarose 1,5% para avaliação do tamanho médio dos insertos obtidos a partir de PCR dos clones isolados da biblioteca de cDNA de <i>D. quadriceps</i> .....	46
<b>Figura 10:</b> Exemplo de eletroferograma de alta qualidade do sequenciamento por Sanger de um clone da biblioteca de cDNA da glândula de veneno de <i>D. quadriceps</i> .....	48
<b>Gráfico 1:</b> Proporção relativa de ESTs classificados como precursores de proteínas envolvidas em processo celular (laranja), sem <i>hits</i> significantes (cinza), toxinas (azul) e proteínas hipotéticas (amarelo).....	50
<b>Gráfico 2:</b> Proporção de transcritos de toxinas expressos na glândula de veneno de <i>D. quadriceps</i> .....	51
<b>Tabela 1:</b> Anotação de ESTs agrupados em contigs obtidos a partir da biblioteca de cDNA da glândula de veneno de <i>D. quadriceps</i> .....	52
<b>Tabela 2:</b> Anotação de ESTs individuais obtidos a partir da biblioteca de cDNA da glândula de veneno de <i>D. quadriceps</i> .....	53
<b>Gráfico 3:</b> Distribuição de comprimento dos transcritos da glândula de veneno de <i>D. quadriceps</i> .....	56
<b>Gráfico 4:</b> Proporção relativa dos transcritos da glândula de veneno de <i>D. quadriceps</i> identificados por NGS.....	57

<b>Tabela 3:</b> Anotação de contigs obtidos a partir do sequenciamento de RNA extraído da glândula de veneno de <i>D. quadriceps</i> .....	59
<b>Tabela 4:</b> Visão geral das principais toxinas polipeptídicas de <i>D. quadriceps</i> e de outras formigas.....	63
<b>Figura 11:</b> Alinhamento de múltiplas sequências de aminoácidos do alérgeno de veneno I (Sol i 1) de <i>D. quadriceps</i> com outras sequências de himenópteros.....	65
<b>Figura 12:</b> Comparação entre os polipeptídeos Sol i 3/Alérgeno Antígeno 5 de <i>D. quadriceps</i> com os do outros himenópteros.....	66
<b>Figura 13:</b> Representação esquemática de proteínas da família CRISP ( <i>Cysteine-rich secretory protein Family</i> ).....	71
<b>Figura 14:</b> Organização do cDNA do contig TX02_DIQUA. ....	72
<b>Figura 15:</b> Alinhamento dos precursores dos alérgenos CAP (Ag5). ....	73
<b>Figura 16:</b> Modelo tridimensional de TX02_DIQUA, Alérgeno CAP do veneno de <i>D. quadriceps</i> , utilizando como modelo VA Sol i 3 (2VZN) de <i>S. invicta</i> e VA Ves5 (1QNX) de <i>V. vulgaris</i> .....	74
<b>Figura 17:</b> Organização do cDNA dos contigs 17 (A) e 21 (B).....	80
<b>Figura 18:</b> Alinhamento dos Precursores codificados pelos Contigs TX17_DIQUA e TX21_DIQUA com proteínas da família PBP/GOPB e <i>Venom Allergen 2/4</i> .....	81
<b>Figura 19:</b> Organização do cDNA do contig TX06_DIQUA. ....	83
<b>Figura 20:</b> Alinhamento de TX06_DIQUA com precursores da família das pilosulinas.....	84
<b>Figura 21:</b> Organização do cDNA dos contigs 09 (A) e 01 (B), alinhamento das sequencias nucleotídicas (C) e dos aminoácidos que compõem o peptídeo sinal (D) dos contigs.....	86
<b>Figura 22:</b> Alinhamento de dinoponeratoxinas de <i>D. quadriceps</i> , TX01_DIQUA (A) e TX09_DIQUA (B) com dinoponeratoxinas de <i>D. australis</i> .....	89
<b>Figura 23:</b> Organização do cDNA do contig 05.....	92
<b>Figura 24:</b> Alinhamento do precursor codificado pelo contig TX05_DIQUA.....	94
<b>Figura 25:</b> A) Representação esquemática de uma <i>Knottina</i> . B) Representação esquemática da conectividade um nó de cisteínas presente em fator de crescimento.....	98

<b>Figura 26:</b> Representação esquemática de knottinas de diferentes origens.....	99
<b>Figura 27:</b> Alinhamento de toxinas maduras, com arranjo Konttin e que apresentam estrutura tridimensional.....	101
<b>Figura 28:</b> A) Modelo tridimensional da toxina ICK-símile codificada por TX05_DIQUA. B) Representação esquemática de uma knottina. C) O alinhamento das sequências das duas estruturas que geraram o melhor modelo de homologia tridimensional da toxina de TX05_DIQUA com a estrutura secundária indicada abaixo.....	77
<b>Figura 29:</b> Alinhamento do precursor codificado pelo contig NGS_1144 de <i>D. quadriceps</i> , com homólogos que apresentam domínio Knottina.....	80
<b>Figura 30:</b> Organização do cDNA do contig 12.....	83
<b>Figura 31:</b> Alinhamento de TX12_DIQUA com toxinas de venenos himenópteros com domínio TIL.....	84
<b>Figura 32:</b> Perfil eletroforético do veneno total de <i>Dinoponera quadriceps</i> (DQv) em gel de eletroforese unidimensional SDS-PAGE.....	85
<b>Figura 33:</b> Perfil de cromatográfico da peçonha bruta de <i>D. quadriceps</i> comparado com o perfil cromatográfico da peçonha bruta de <i>D. australis</i> .....	114
<b>Figura 34:</b> Perfil de cromatográfico da peçonha bruta de <i>D. quadriceps</i> .....	115
<b>Figura 35:</b> Perfil de cromatografia líquida acoplada à espectrometria de massa da peçonha bruta de <i>D. quadriceps</i> . .....	117
<b>Tabela 5:</b> Perfil de massas do veneno bruto de <i>D. quadriceps</i> utilizando espectrometria de massas electrospray (ESI).....	119
<b>Figura 36:</b> Moléculas de baixo peso molecular identificadas através de espectrometria de massas.....	120
<b>Tabela 6:</b> Agrupamento dos peptídeos sequenciados através de espectrometria de massas.....	120
<b>Figura 37:</b> Alinhamento do precursor codificado pelo contig TX01_DIQUA com peptídeos cuja sequencia foi determinada pela análise proteômica e sequenciamento de novo MS/MS.....	122

**Figura 38:** Alinhamento do precursor codificado pelo contig TX09\_DIQUA com peptídeos cuja sequência foi determinada pela análise proteômica e sequenciamento de novo MS/MS.....124

**Figura 39:** Alinhamento do precursor codificado pelo contig TX06\_DIQUA com peptídeos cuja sequência foi determinada pela análise proteômica e sequenciamento de novo MS/MS.....127

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
1.1	Toxinas animais como modelos de substâncias bioativas	15
1.2	Biologia de formigas: Subfamília Ponerinae	18
1.3	Aspectos clínicos da picada de formigas	20
1.4	Composição e propriedades farmacológicas da peçonha de formigas	22
1.5	Estudo transcritoômico da composição da peçonha de formigas	24
<b>2</b>	<b>OBJETIVOS</b>	<b>27</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>28</b>
3.1	Autorização para coleta de formigas	28
3.2	Animais	28
3.3	Extração do aparato secretório	28
3.4	Isolamento de RNA total	29
3.5	Construção de biblioteca de cDNA	29
3.5.1	<i>Síntese de cDNA por LD-PCR</i>	29
3.5.2	<i>Purificação e fracionamento do cDNA dupla fita</i>	31
3.5.3	<i>Clonagem no vetor pSMART2IFD</i>	31
3.5.4	<i>Transformação de E. coli com plasmídeo recombinante</i>	34
3.5.5	<i>Amplificação do DNA dos clones obtidos</i>	37
3.5.6	<i>Sequenciamento por Sanger, organização e agrupamento dos clones</i>	37
3.6	Sequenciamento NGS ( <i>Next Generation Sequencing</i> ) utilizando tecnologia Ion Torrent	38
3.6.1	<i>Organização e agrupamento das sequências obtidas por NGS</i>	39
3.7	Análise de sequências específicas de transcritos	39
3.8	Análise proteômica da peçonha total de <i>D. quadriceps</i>	40
3.8.1	<i>Extração da peçonha</i>	40
3.8.2	<i>Dosagem de proteína total pelo Método Bradford</i>	40
3.8.3	<i>Eletroforese SDS-PAGE</i>	41
3.8.4	<i>Purificação por cromatografia líquida de alta eficiência em fase reversa da peçonha bruta</i>	41
3.8.5	<i>Análise por cromatografia a líquido acoplada a espectrometria de massas (LC-MS) da peçonha bruta de D. quadriceps</i>	42

<b>4 RESULTADOS E DISCUSSÃO</b> .....	43
<b>4.1 Extração do RNA total e construção da biblioteca de cDNA da glândula de veneno de <i>D. quadriceps</i></b> .....	43
<b>4.2 Sequenciamento por Sanger e agrupamento das sequências</b> .....	47
<b>4.3 Sequenciamento NGS (<i>Next generation sequencing</i>) por Ion Torrent e agrupamento das sequências</b> .....	56
<b>4.4 Anotação do agrupamento dos precursores de toxinas da glândula de veneno de <i>D. quadriceps</i></b> .....	68
<b>4.4.1 Toxinas alergênicas</b> .....	69
4.4.1.1 <i>Toxina alergênica semelhante à Sol i 3/Ag5</i> .....	69
4.4.1.2 <i>Alérgenos de veneno 2/4 (PBP/GOPB)</i> .....	76
4.4.1.3 <i>Alérgenos Pílosulina-símile</i> .....	82
<b>4.4.2 Dinoponeratoxinas</b> .....	85
<b>4.4.3 Toxina ICK-símile</b> .....	91
<b>4.4.4. Toxinas com domínio TIL</b> .....	108
<b>4.5 Análise proteômica parcial da peçonha bruta de <i>D. quadriceps</i></b> .....	112
<b>4.5.1 Extração, dosagem de proteína e perfil eletroforético SDS-PAGE da peçonha bruta de <i>D. quadriceps</i></b> .....	112
<b>4.5.2 Purificação da peçonha bruta de <i>D. quadriceps</i> por cromatografia em fase reversa (RP HPLC)</b> .....	113
<b>4.5.3 Cromatografia líquida acoplada à espectrometria de massas (LC-MS)</b> .....	116
<b>5 CONSIDERAÇÕES FINAIS</b> .....	129
<b>6 CONCLUSÃO</b> .....	131
<b>7 REFERÊNCIAS</b> .....	132

## 1 INTRODUÇÃO

### 1.1 Toxinas animais como modelos de substâncias bioativas

Há décadas animais venenosos recebem notoriedade, uma vez que sua peçonha contém diversas substâncias bioativas potencialmente úteis (toxinas polipeptídicas), as quais são codificadas por genes ou, no caso de compostos orgânicos, sintetizados por rotas enzimáticas específicas presentes em tecidos especializados onde a biossíntese ocorre – a glândula de veneno.

“A natureza sempre despertou no homem o fascínio não apenas por recursos oferecidos para alimentação mas por ser sua principal fonte de descobertas. A convivência e o aprendizado com os mais diferentes grupos étnicos trouxeram valiosas contribuições para o desenvolvimento da pesquisa com produtos naturais, do conhecimento da relação entre a estrutura química de um determinado composto e suas propriedades biológicas e da inter-relação animais-insetos-plantas. Neste sentido a natureza forneceu muitos modelos moleculares que vêm fundamentando diversos estudos (VIEGAS *et al.*, 2006).”

Este legado ainda permanece, uma vez que continua influenciando o desenho de novas moléculas, no contexto da pressão seletiva pertinente à evolução e, dessa forma, desempenha um papel significativo no desenvolvimento de novas drogas. À exemplo de áreas de intenso estudo como câncer e doenças infecciosas, 60 e 75% das novas drogas, respectivamente, são oriundas de fontes naturais (NEWMAN; CRAGG, 2012; GULLO *et al.*, 2006).

Sabe-se que as estruturas dos produtos naturais possuem características de grande diversidade química, especificidade bioquímica e propriedades moleculares que os tornam favoráveis como modelos que conduzem à descoberta de novas drogas (CLARDY; WALSH, 2004). Assim, a exploração sustentável, com vistas à conservação dos recursos biológicos do planeta terra tem sido frequentemente discutida e estimulou o desenvolvimento de programas de bioprospecção (busca na natureza por substâncias com fins comerciais) em diversos países, integrando universidades, institutos de pesquisa e a indústria farmacêutica, para descobrir e desenvolver novos fármacos. Neste âmbito torna-se importante mencionar o envolvimento de diferentes ramos da ciência como, biologia celular e molecular, bioquímica, farmacologia, química computacional, entre outros.

As toxinas naturais têm sido utilizadas como ferramenta para elucidar mecanismos fisiológicos desde que Claude Bernard em 1850 verificou que o veneno utilizado nas pontas das flechas por índios continham tubocurarinas conhecidas durante muito tempo como um alcaloide letal, mas foi o primeiro relaxante muscular, introduzido na prática anestésica (Atracurium® - Wellcome). Desde então vários outros laboratórios farmacêuticos têm se utilizado de toxinas naturais como modelos de moléculas para o desenvolvimento de seus produtos (PALMA *et al.*, 2007; HARVEY *et al.*, 1998). Pesquisas com toxinas de origem animal têm contribuído muito na compreensão de problemas vasculares, neurológicos, processos inflamatórios, analgesia, processos alérgicos, asma, entre outros (MORTARI *et al.*, 2007). Venenos de invertebrados têm interesse como potencial fonte de substâncias bioativas, como as neurotoxinas, moléculas que têm demonstrado muita utilidade para o entendimento dos eventos de transmissão sináptica e têm contribuído no desenho de novas drogas para o tratamento de desordens neurológicas e da dor (MORTARI *et al.*, 2007; KNAPP *et al.*, 2012).

Do ponto de vista venômico, as toxinas polipeptídicas ganham notoriedade devido à relativa facilidade de síntese, estabilidade estrutural e especificidade farmacológica. A exemplo do Prialt®, analgésico não opióide desenvolvido a partir de  $\omega$ -conotoxina (POPE; DEER, 2013), outros medicamentos, inspirados em peptídeos de veneno, vêm sendo desenvolvidos. Assim, as proteínas terapêuticas têm se tornado uma classe de drogas em crescimento para a indústria farmacêutica, principalmente após o desenvolvimento da tecnologia de DNA recombinante, a qual veio não somente para facilitar o estudo destas substâncias, como também para viabilizar sua produção em maior escala. A diversidade funcional destes peptídeos também impressiona, tendo sido relatados efeitos farmacológicos como antitumoral (HMED *et al.*, 2013), em quadros infecciosos (de OLIVEIRA-JUNIOR *et al.*, 2013) e distúrbios neurológicos (ESSAK *et al.*, 2012).

Tendo em vista que o Brasil possui uma fonte riquíssima de toxinas animais e de microorganismos, um grande esforço vem sendo aplicado na investigação das atividades biológicas destas, e muito tem sido descoberto. Conceição *et al.* (2006) identificaram na secreção cutânea da *Phyllomedusa hypocondrialis*, anfíbio anúrio encontrado no Nordeste e Sudeste brasileiro, peptídeos antimicrobianos (PS-7 e DH-1) capazes de inibir o crescimento de *Staphylococcus aureus*, *Pseudomonas*

*aeruginosa* e *Escherichia coli*. Em estudos posteriores, Conceição et al. (2007) identificaram um peptídeo potenciador de bradicinina (Phypo Xa), com significativa ação hipotensora. Neste sentido o escorpião amarelo *Tityus serrulatus*, comum em grandes cidades, apresenta em seu veneno peptídeos com ação hipotensora, denominados hipotensinas, as quais também exercem seu efeito através da potencialização da ação da bradicinina (VERANO-BRAGA et al., 2008). O carrapato estrela (*Amblyomma cajennense*), encontrado em todo Brasil, possui em sua saliva substâncias anticoagulantes (BATISTA et al., 2008), que inibem o fator 10 da cascata de coagulação, e que pode vir a ser uma droga anti-trombótica. O extrato isolado das cerdas da taturana *Lonomia obliqua*, encontrada na região Sudeste do Brasil, apresenta uma protease ativadora da protrombina (Lopap) (REIS et al., 2001; FRITZEN et al., 2005), e que pode interagir com alguns receptores de células endoteliais, aumentando a expressão superficial de moléculas de adesão (CHUDZINSKI-TAVASSI et al., 2001).

No mesmo âmbito, toxinas derivadas de insetos, principalmente os da ordem Himenóptera, a qual compreende abelhas, vespas e formigas, vem alcançando destaque. Diversas investigações têm sido realizadas visando à descoberta de novas substâncias a partir de venenos de artrópodes. Um componente isolado da peçonha da vespa social *Polybia occidentalis*, uma Thr6-bradicinina, apresenta atividade antinociceptiva central em ratos via receptor B<sub>2</sub> da bradicinina (MORTARI et al., 2007) e peçonha de abelhas apresentam atividade antinociceptiva e antiinflamatória sobre o edema de pata e a hiperalgesia térmica induzidos por carragenina em ratos (LEE et al., 2001). A peçonha da vespa *P. occidentalis* inibiu várias etapas das vias intrínseca, extrínseca e comum da coagulação, bem como a agregação plaquetária e a degradação de fibrinogênio plasmático (CZAIKOSKI et al., 2010). Também foi demonstrado que a peçonha da abelha *Apis mellifera* apresenta efeito antiinflamatório envolvendo a inibição da expressão da enzima ciclooxigenase-2 (COX-2) e o bloqueio de citocinas pró-inflamatórias (TNF- $\alpha$  e IL-1 $\beta$ ) (NAM et al., 2003). No entanto, ainda não foram descritos tais efeitos para peçonha de formigas, embora já tenha sido demonstrado que pequenos peptídeos isolados (poneratoxin ou PoTx) da peçonha de *Paraponera clavata*, pertencente a subfamília Ponerinae, interferem em canais de sódio de fibra muscular esquelética de sapos (DUVAL et al., 1992). Estes canais também poderiam mediar uma resposta antinociceptiva em roedores, visto que

participam tanto de mecanismos nociceptivos centrais quanto periféricos. Nosso grupo observou efeitos renais induzidos pela peçonha da vespa *Polybia paulista* (VINHOTE *et al.*, 2011) e da formiga *Dinoponera quadriceps* (TORRES *et al.*, 2013). Ainda avaliando efeitos biológicos da peçonha de *D. quadriceps* nós observamos, em modelos de convulsão induzidos pelo pentilenotetrazol, uma ação neuroprotetora (LOPES *et al.*, 2011; 2013). Em modelos de ação antimicrobiana a mesma apresentou um pronunciado efeito tanto sobre microrganismos gram positivos como sobre gram negativos (dados não publicados).

## 1.2 Biologia de formigas: Subfamília Ponerinae

As formigas (Vespoidea: Formicidae) pertencem à classe Insecta e ordem Hymenoptera, a qual pertencem também famílias como Apidae (abelhas), e Vespidae (vespas) (GULLAN; CRANSTON, 2010). A família *Formicidae*, apresenta aproximadamente 13.000 espécies de formigas, cuja maioria exibe uma avançada e sofisticada organização social, com colônias compostas de dezenas à milhares de indivíduos e presentes em quase todo tipo de habitat. Assim, as formigas são elementos importantes da biodiversidade terrestre, especialmente em regiões neotropicais, onde cerca de 30% de todas as espécies de formigas conhecidas são encontradas (HOLLDOBLER; WILSON, 1990; FERNANDEZ; OSPINA, 2003). Todas as espécies de formigas possuem hábitos sociais, o mais conhecido sendo a divisão reprodutiva e de trabalho, com uma a muitas rainhas férteis, enquanto as operárias, não férteis, ajudam a(s) rainha(s) na reprodução, e lidam com todas as outras tarefas da colônia, como coleta de alimentos, reparo do ninho e defesa da colônia (WILSON, 1971).

Com mais de 1000 espécies distribuídas em 28 gêneros, como *Dinoponera* e *Paraponera* a subfamília Ponerinae é um grupo primitivo de formigas, principalmente encontrado em áreas tropicais. A Ponerinae é uma subfamília poneromorfa primitiva, que apresenta formigas de grande tamanho (~3,0 cm), e está distribuída pelo mundo inteiro. É representada por vários gêneros, como *Dinoponera*, *Paraponera* e *Diacamma*. O gênero *Dinoponera* abriga 8 espécies (*D. quadriceps*; *D. australis*, *D. gigantea*, *D. hispida*, *D. lucida*, *D. longipes*, *D. mutica*, *D. snellingi*), conhecidas como falsas tocandiras (Figura 1) (SCHOETERS; BILLEN, 1995). Como outras Ponerinae,

as colônias de *Dinoponera* têm uma pobre organização social, com pequeno número de indivíduos e não apresentam uma rainha distinguível. Ao contrário da maioria das espécies de formigas, todas as operárias são potencialmente férteis com espermatecas funcionais. No entanto, somente uma ou poucas operárias são sexualmente ativas e conflitos reprodutivos são característicos da sociedade destas formigas (MONIN; PEETERS, 1998; ARAÚJO; JAISSON, 1994). São geralmente predadoras e possuem ferrão funcional, com o qual submetem suas presas, que compreendem pequenos insetos e larvas (CAETANO *et al.*, 2002). Estão distribuídas, principalmente, ao longo da América do Sul, e algumas dessas espécies são encontradas no Cerrado, na Caatinga e na Mata Atlântica (SCHOETERS; BILLEN, 1995; SIQUIEROLI *et al.*, 2007).

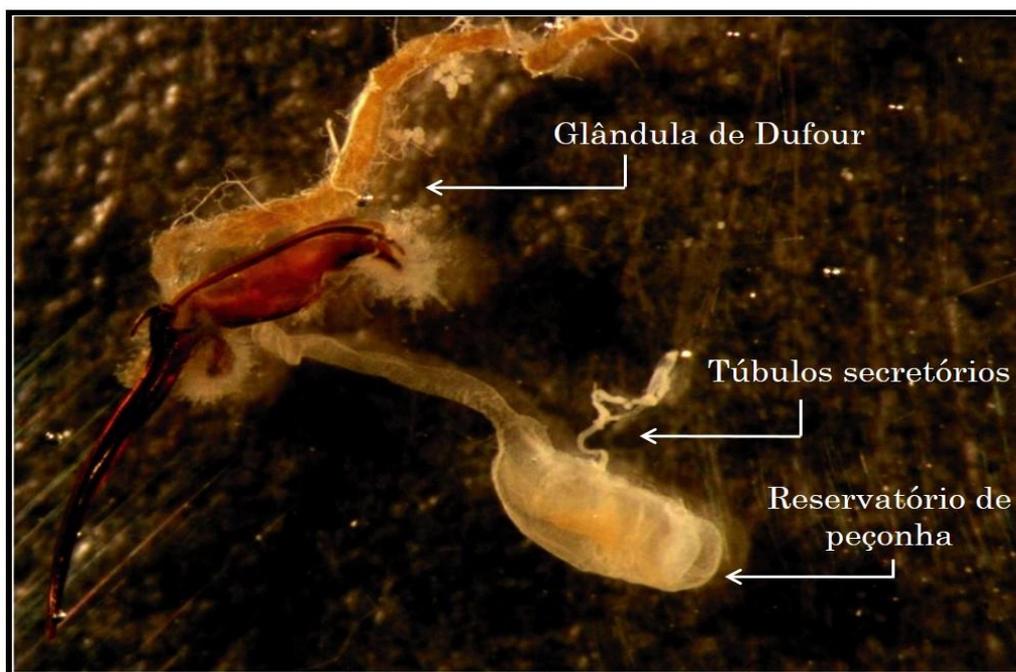
**Figura 1:** *Dinoponera quadriceps* (QUINET, 2010)



Como todos os himenópteros Aculeata (Chrysoidea, Apoidea, Vespoidea), as formigas do gênero *Dinoponera* possuem glândulas especializadas na produção de peçonha. Esta, nos insetos, está localizada no último segmento de seu corpo, o gáster, e é composta pelo ferrão, ao qual estão associadas duas glândulas: a glândula de Dufour e a glândula de veneno (BUSCHINGER; MASCHWITZ, 1984). Em todas as formigas, a glândula de veneno é composta por um par de túbulos

secretórios, os quais convergem para uma glândula contorcida dentro do reservatório de veneno (HOLLDOBLER; WILSON, 1990) (Figura 2). Em *D. australis*, foi demonstrado que a glândula contorcida, assim como os túbulos, são responsáveis pela produção da peçonha, a qual é composta, em sua maior parte, por proteínas (SIQUIEROLI *et al.*, 2007; HÖLLDOBLER; WILSON, 1990; SCHOETERS; BILLEN, 1995).

**Figura 2:** Aparato secretório de *D. quadriceps* (QUINET, 2010)



### 1.3 Aspectos clínicos da picada de formigas

Muitas picadas de insetos estão associadas a eventos patofisiológicos localizados caracterizados por dor, edema e vermelhidão, por 1-2 dias (ELLIS; DAY, 2005). As reações mais severas estão associadas a eventos alérgicos apresentando produção de IgE e infiltrado neutrofílico e eosinofílico (HADDAD, 2009). Estas manifestações são comuns em acidentes com himenópteros. Muitos estudos que descrevem os aspectos clínicos das picadas de formigas reportam acidentes com formigas do gênero *Solenopsis* (Subfamília Myrmicinae), conhecidas como formigas de fogo (LEE *et al.*, 2009; TANKERSLEY, 2008; CHIANURA; POZZI, 2010). Nos casos

mais graves, estes acidentes promovem *rash* cutâneo, convulsões, falha cardíaca, nefrite e, mais raramente, insuficiência renal aguda (KOYA *et al.*, 2007; RACHAIAH *et al.*, 2012).

Acidentes com formigas da subfamília Ponerinae são raros ou raramente reportados. De fato, muitas picadas concomitantes são necessárias para a produção de sintomas clínicos significativos e, para formigas gigantes, esses múltiplos ataques são menos prováveis, uma vez que estas formigas possuem hábitos alimentares, solitários. Tribos indígenas amazônicas utilizam esse tipo de formiga, especificamente a *Paraponera clavata*, em rituais de passagem para a fase adulta. Desda forma índios de aproximadamente 12 anos têm que introduzir sua mão em uma bolsa com um grande número de formigas por cerca de 30 minutos como prova de sua bravura (HADDAD Jr *et al.*, 2005).

Alguns dos acidentes com formigas gigantes podem apresentar importância médica, como os provocados pelos gêneros *Paraponera* e *Dinoponera*, popularmente conhecidas como verdadeira e falsa tocandira, respectivamente. Sua picada é extremamente dolorosa e pode provocar, manifestações sistêmicas como febre, tremores, dano tecidual, sudorese, náuseas, vômitos, linfadenopatia e arritmias cardíacas (HADDAD Jr. *et al.*, 2001, 2005; CRUZ LOPEZ; MORGAN, 1997).

De acordo com Hermann *et al.* (1984) e Orivel e Djean (2001) a peçonha destas formigas pode apresentar efeitos neurotóxicos, causando paralisia rápida e dose dependente, seguida de paralisia permanente que conduz à morte em insetos. Em 1990, Piek relatou novos aspectos sobre o modo como os himenópteros incapacitam suas presas. Segundo o autor, diferentes neurotoxinas bloqueiam a transmissão sináptica no sistema nervoso central ou na junção neuromuscular, alterando fenômenos voltagem-dependentes, a exemplo da Poneratoxina, um polipeptídeo de 25 aminoácidos identificado na peçonha da formiga *Paraponera clavata* (DUVAL *et al.*, 1992), que é capaz de alterar a cinética de canais de sódio voltagem-dependente. Neste sentido Zalat *et al.* (2003) verificaram que a peçonha de abelha, vespas e formigas são capazes de atuar sobre receptores de acetilcolina e receptores ionotrópicos de glutamato.

A peçonha é composta das secreções oriundas das glândulas de veneno e de Dufour e auxilia em funções biológicas como comunicação química, demarcação territorial. Do ponto de vista predatório a peçonha pode estar envolvida na paralisia da

presa, interferindo na transmissão nervosa, bem como no processo de digestão enzimática da mesma. Deste modo a bioatividade dos componentes protéicos pode estar relacionada com estas funções.

#### **1.4 Composição e propriedades farmacológicas da peçonha de formigas**

A peçonha de formigas foi investigada em um número relativamente pequeno de espécies. Estudos mostram que a peçonha de formigas da subfamília *Formicinae* é constituída de misturas simples de ácidos orgânicos, enquanto que a peçonha de formigas das subfamílias *Ponerinae* e *Myrmicinae* é constituída de misturas complexas de proteínas e neurotoxinas (LIMA; BROCHETTI-BRAGA, 2003). Num contexto mais amplo, a peçonha dos himenópteros é constituída de complexos de moléculas orgânicas, proteínas, peptídeos, lipídeos e aminas vasoativas, entre elas a norepinefrina, histamina e dopamina (LIMA; BROCHETTI-BRAGA, 2003). Vários pesquisadores vêm trazendo à luz diversos peptídeos de venenos himenópteros cujas funções variam desde antimicrobiana à neurotóxica (VAN VAERENBERGH *et al.*, 2013; RIFFLET *et al.*, 2012; GAL; LIBERSAT, 2010), como os mastoparanos encontrados no veneno da vespa *Polybia paulista*, que apresentam ação antimicrobiana com propriedade de permear membranas, além de serem indutores da degranulação de mastócitos e causadores de hemólise (YANDEK *et al.*, 2009; ROCHA *et al.*, 2009). Essas substâncias são responsáveis pela sua toxicidade e muitas possuem propriedades farmacológicas e terapêuticas (BLUM, 1978).

As peçonhas mais bem estudadas são das formigas das subfamílias Myrmaciinae, Ponerinae, Pseudomyrmecinae e Myrmicinae. Se apresentam como soluções aquosas ricas em proteínas enzimáticas e não enzimáticas, aminoácidos livres e pequenos compostos biologicamente ativos como histamina, acetilcolina, norepinefrina e dopamina (SCHMIDT *et al.*, 1986; ATTYGALLE; MORGAN, 1984). No entanto, formigas do gênero *Solenopsis* e *Monomorium* produzem peçonhas ricas em alcalóides, com poucas proteínas. Em formigas Formicinae (*Camponotus*, *Formica*), o ferrão não está presente, mas a glândula de veneno produz misturas simples de ácidos orgânicos em solução aquosa. Ácido fórmico está presente em concentrações até 65% (SCHMIDT *et al.*, 1986; ATTYGALLE; MORGAN, 1984).

Como membro do grupo das Ponerinae, espera-se que a glândula de peçonha de *Dinoponera quadriceps* produza peçonha com características proteicas. No entanto, até o presente, poucos estudos foram realizados com a peçonha de formigas do gênero *Dinoponera*. Em dois destes estudos, os quais compararam peçonhas de uma variedade de himenópteros, a presença de proteínas, algumas com ação enzimática (fosfolipase A, hialuronidase e lipase), foi demonstrado para a peçonha de *D. grandis* (*D. gigantea*) (SCHMIDT *et al.*, 1986; LELUK *et al.*, 1989). Em estudo mais recente, compostos peptídicos da peçonha de *D. australis* foram investigados. Mais de 75 componentes proteicos distintos foram relatados, com uma diversidade de propriedades, variando em tamanho, hidrofobicidade e abundância (JOHNSON *et al.*, 2010). Os efeitos biológicos da peçonha de várias formigas têm sido atribuídos ao seu conteúdo proteico. Como demonstrado por Cologna *et al.* (2005), proteínas de alto peso molecular estão presentes na peçonha de *D. australis*. Em uma avaliação comparativa da composição proteica da peçonha de himenópteros, proteínas com peso molecular variando de 24-75kDa foram evidenciadas (LELUK *et al.*, 1989). Adicionalmente, dois peptídeos com menos de 10kDa, bem como proteínas com peso molecular variando de 26-90kDa foram encontradas na peçonha de *Myrmecia pilosula* (WEISE *et al.*, 2006). O perfil eletroforético de vespas também apresenta variação no peso molecular de seus constituintes proteicos variando de 5-200kDa, enquanto que em abelhas esse perfil varia de 2-108kDa (ZALAT *et al.*, 1999).

Himenópteros apresentam-se como uma fonte de moléculas bioativas, as quais, além de estarem envolvidas em causas de alergia e envenenamento, apresentam uso na medicina tradicional e fazem parte de programas de desenvolvimento de novas drogas. O primeiro relato sobre o uso terapêutico dos venenos de formigas foi para tratar casos de artrite reumatoide. De fato, peçonhas de insetos podem apresentar constituintes que justifiquem seu uso na medicina tradicional em países do leste da Ásia, África e América do Sul (CHERNIAK, 2010). Os últimos estudos com peçonhas de formiga ajudaram a demonstrar propriedades benéficas como redução da inflamação, da dor, melhora de resposta imune (KOU *et al.*, 2005; WANG; WU, 2006).

Os venenos de formigas da subfamília *Ponerinae*, compostos principalmente de proteínas, têm também apresentado efeitos farmacológicos. Pequenos peptídeos isolados do veneno da *Paraponera clavata*, chamados

poneratoxin (PoTx) interferem em canais de sódio e têm potencial uso como inseticida biológico (DUVAL *et al.*, 1992; SZOLAJSKA *et al.*, 2004). Badr *et al.* (2012) demonstraram que o veneno total de *Pachycondyla sennarensis* induz a morte por apoptose em células de câncer de mama. O mesmo veneno apresentou ação anti-inflamatória e anti-oxidante (EBAID *et al.*, 2012; DKHIL *et al.*, 2010). Adicionalmente, o veneno da formiga Chinesa, *Polyrhachis lammellidens* apresentou efeitos analgésicos e anti-inflamatórios (KOU *et al.*, 2005). Peptídeos isolados dos venenos de *Pachycondyla goeldii* e *Myrmecia sp.* demonstraram eficácia contra bactérias gram positivas, gram negativas e fungo, o que incluiu *Staphylococcus aureus* metilina resistente e *Candida albicans* (ORIVEL; DJEAN, 2001; ZELEZETSKY *et al.*, 2005). Alguns efeitos biológicos, em diferentes modelos experimentais, provocados pelo veneno de *D. quadriceps* vem sendo investigados por nosso grupo de pesquisa. Em sistema de perfusão de rim isolado de rato o veneno provocou um aumento no fluxo urinário, diminuição da resistência vascular e redução no transporte tubular de sódio, sugerindo um efeito diurético natriurético (TORRES *et al.*, 2013). Ademais, em ensaios de toxicidade sobre células tubulares renais (MDCK-Madin Darbin Canine kidney) o veneno apresentou citotoxicidade de modo dose e tempo dependente. Interessantemente, a maior citotoxicidade foi observada nos menores períodos de incubação, sugerindo uma recuperação da cultura celular após um dado período de tempo (TORRES *et al.*, 2013). Em modelos de convulsão provocados pelo pentilenotetrazol, o veneno demonstrou uma atividade neuroprotetora, a qual se apresentou relacionada à via de administração utilizada (LOPES *et al.*, 2011; 2013). Adicionalmente, em modelos de antinocicepção pode-se observar o efeito analgésico (SOUSA *et al.*, 2013). Em modelos de ação antimicrobiana o veneno de *D. quadriceps* apresentou ação inibitória pronunciada principalmente sobre cepas de *S. aureus*, *P. aeruginosa*, *E. coli* e *Salmonella sp.* (resultados não publicados).

### **1.5 Estudo transcritoômico da composição da peçonha de formigas**

Desde a descrição do DNA dupla hélice por Francis Crick e James Watson (1953), a tecnologia do DNA recombinante e genômica revolucionaram as diferentes áreas das ciências da vida. A compreensão da base bioquímica e molecular da herança gênica tem aperfeiçoado nosso conhecimento sobre a complexidade de todas

as formas de vida e da maneira como os genes e proteínas interagem para compor a diversidade. A revolução genômica foi adicionalmente expandida com o advento da bioinformática e das ciências “ômicas” (transcriptômica, proteômica, peptidômica, metabolômica).

Esforços coletivos têm sido empregado para reunir e anotar a composição gênica de insetos, todavia o genoma de insetos permanece pouco estudado, principalmente se levarmos em consideração sua abundância e diversidade. De aproximadamente 1.000.000 espécies de insetos descritas, estima-se que apenas 0,045% tenham o tamanho do seu genoma descrito. As formigas estão entre os mais abundantes e ecologicamente importantes dos artrópodos. Todavia, ainda existem poucas ferramentas genômicas para a família dos insetos (WILSON, 1971; HÖLLDOBLER; WILSON, 1990; TSUTUI et al., 2008). O primeiro genoma de inseto foi o da mosca doméstica *Drosophila melanogaster*, em 2000. No campo da toxilogia, himenópteros têm recebido especial atenção devido seu comportamento e habilidade de produzir peçonha.

Até o presente, pelo menos dez espécies de formigas tiveram seu genoma analisado e publicado. Estas incluem: a formiga de fogo *Solenopsis invicta*, encontrada na América do Sul, Estados Unidos, China Taiwan, Australia (WURM et al., 2011); *Acromyrmex echinator* (NYGAARD et al., 2011) e *Atta cephalote* (SUEN et al., 2011) encontradas na América do Sul; *Pogonomyrmex barbatus* encontrada na América do Norte e Sul (SMITH et al., 2011b); todas pertencentes à subfamília Myrmicinae; a formiga Argentina *Linepithema humile* (SMITH et al., 2011a), da subfamília Dolichoderinae e as formigas *Camponotus floridanus* dos Estados Unidos e *Harpegnathos saltator* da Índia, Sri Lanka e Sudeste da Ásia (BONASIO et al., 2010), pertencentes às subfamílias Formicinae e Ponerinae, respectivamente.

Além da análise genômica detalhada, a construção de biblioteca de cDNA é uma importante ferramenta para a descoberta de novas moléculas que podem apresentar propriedades biológicas/farmacológicas, proporcionando avaliação gênica tecido específica além do estudo dos níveis de expressão relativa dos transcritos. A análise de peçonhas produzidas em pequena quantidade encontra dificuldades para o uso de técnicas bioquímicas convencionais e são mais facilmente estudadas através de construção de uma biblioteca de cDNA (DNA complementar) (WAGSTAFF et al., 2009). O cDNA pode ser obtido a partir de RNA mensageiro (RNAm) isolado de um

tecido. Como o RNAm representa um conjunto de genes que estão sendo expressos por aquele tecido num dado momento, o cDNA obtido representa o DNA codificante, isto é, sequências que podem ser traduzidas. O conjunto de cDNAs obtidos pode ser inserido em vetores e multiplicado em células hospedeiras, dando origem a bibliotecas de cDNA (DA SILVA, 2004). A leitura e interpretação comparativa das sequências obtidas com sequências conhecidas de diversos tipos de toxinas fornecem um direcionamento sobre a diversidade proteica destas, permitindo, deste modo, a inferência sobre a composição total do veneno produzido e, adicionalmente, a identificação de novas moléculas.

Neste sentido, lançar mão do sequenciamento de peptídeos traz informações muito úteis. De fato, um dos primeiros passos para a investigação da peçonha é a avaliação do perfil cromatográfico seguidos de espectrometria de massas. O sequenciamento de proteínas traz à luz sequências de aminoácidos que compõem estruturas peptídicas presentes na peçonha. Assim, a combinação da análise transcriptômica, a qual fornece informações sobre as sequências nucleotídicas utilizadas na predição dos transcritos; e proteômica, a qual revela sequências de aminoácidos e auxiliam a inferir sobre os processamentos e modificações pós-traducionais, se torna um ponto chave para a detalhada investigação sobre a composição da peçonha, e compõem a estratégia principal da venômica.

## 2 OBJETIVOS

### 2.1 Gerais

Caracterizar os constituintes da peçonha de *Dinoponera quadriceps*, por abordagem transcriptômica e peptidômica.

### 2.2 Específicos

- Estudar a constituição proteica do veneno da *Dinoponera quadriceps*;
- Gerar uma biblioteca de cDNA da glândula do veneno da *Dinoponera quadriceps*;
- Isolar clones e sequenciá-los visando identificar e caracterizar, por similaridade, os transcritos presentes na glândula de veneno da *Dinoponera quadriceps*;
- Comparar estas sequências com outras conhecidas nas diversas bases de dados;
- Complementar a análise transcriptômica através de técnicas de sequenciamento NGS (Next Generation Sequencing);
- Realizar análise proteômica da peçonha bruta de *D. quadriceps*.

## **3 METODOLOGIA**

### **3.1 Autorização para coletas das formigas**

A autorização nº 28794-1, para atividades com finalidades científicas, foi concedida pelo Ministério do Meio Ambiente (MMA)/Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio)/Sistema de Autorização e Informação em Biodiversidade (SISBIO).

### **3.2 Animais**

As formigas foram coletadas na Serra de Maranguape (3°52'51"S e 38°42'39"O), localizada a cerca de 40 km da costa, no Ceará, Nordeste, Brasil. Para atração das formigas e localização dos ninhos foi utilizado atum em conserva. Após a identificação do ninho, o mesmo foi escavado e todos os exemplares encontrados foram coletados com auxílio de pinças e levadas ao Laboratório de Biologia de Insetos Sociais/UECE onde foram acondicionados em ninhos artificiais e alimentados com larva de tenébrio. Após uma semana de adaptação iniciou-se a extração da peçonha, a qual ocorreu semanalmente. Cerca de 20 formigas foram utilizadas para a dissecação do aparato secretório, utilizando um esteromicroscópio e pinças especializadas, para que não houvesse interferência das condições artificiais (alimentação) nos transcritos que seriam pesquisados nos tecidos que compreendem o aparato secretório.

### **3.3 Extração do aparato secretório**

As formigas foram submetidas à refrigeração e em seguida tiveram seu gáster removido e a carapaça rompida para exposição dos tecidos. O aparato secretório foi cuidadosamente separado dos tecidos circunjacentes. Em seguida, separou-se a glândula de Dufour da glândula de veneno e estas estruturas foram rapidamente transferidas para microtubos contendo solução de RNAlater (Life Technologies, USA), e armazenados, à temperatura 2-8°C por 24h, para preservação do RNA. Em seguida foram removidos para um microtubo vazio e armazenados à -80°C até a extração do RNA total (ORIVEL *et al.*, 2001).

### 3.4 Isolamento de RNA total

O RNA total da glândula de veneno foi isolado através o kit RNeasy Mini (Qiagen, Alemanha), o qual é baseado em uma combinação de lise do tecido e denaturação utilizando isotiocianato de guanidina e purificação através de uma membrana de sílica. Os procedimentos foram realizados de acordo com as recomendações do fabricante. A qualidade do RNA foi analisada através de eletroforese de RNA sob condições desnaturantes em gel de agarose (MASEK *et al.*, 2005), sendo a quantificação realizada utilizando NanoDrop (Thermo, USA), o qual também nos fornece informações sobre qualidade do RNA através da razão 260nm/280nm (2,0-2,3).

### 3.5 Construção de biblioteca de cDNA

A biblioteca de cDNA foi construída utilizando o kit IN-FUSION™ SMARTer™ cDNA Library Construction (Clontech, Mountain View, CA-USA). A tecnologia SMARTer (Switching Mechanism At 5' end of RNA Tamplate) possibilita a incorporação de adaptadores sintéticos em ambas as extremidades durante a síntese da primeira fita de DNA, permitindo a completa amplificação e prevenindo a geração de insertos quiméricos. Desta forma temos um cDNA limpo e representativo dos transcritos da glândula de veneno de *D. quadriceps*.

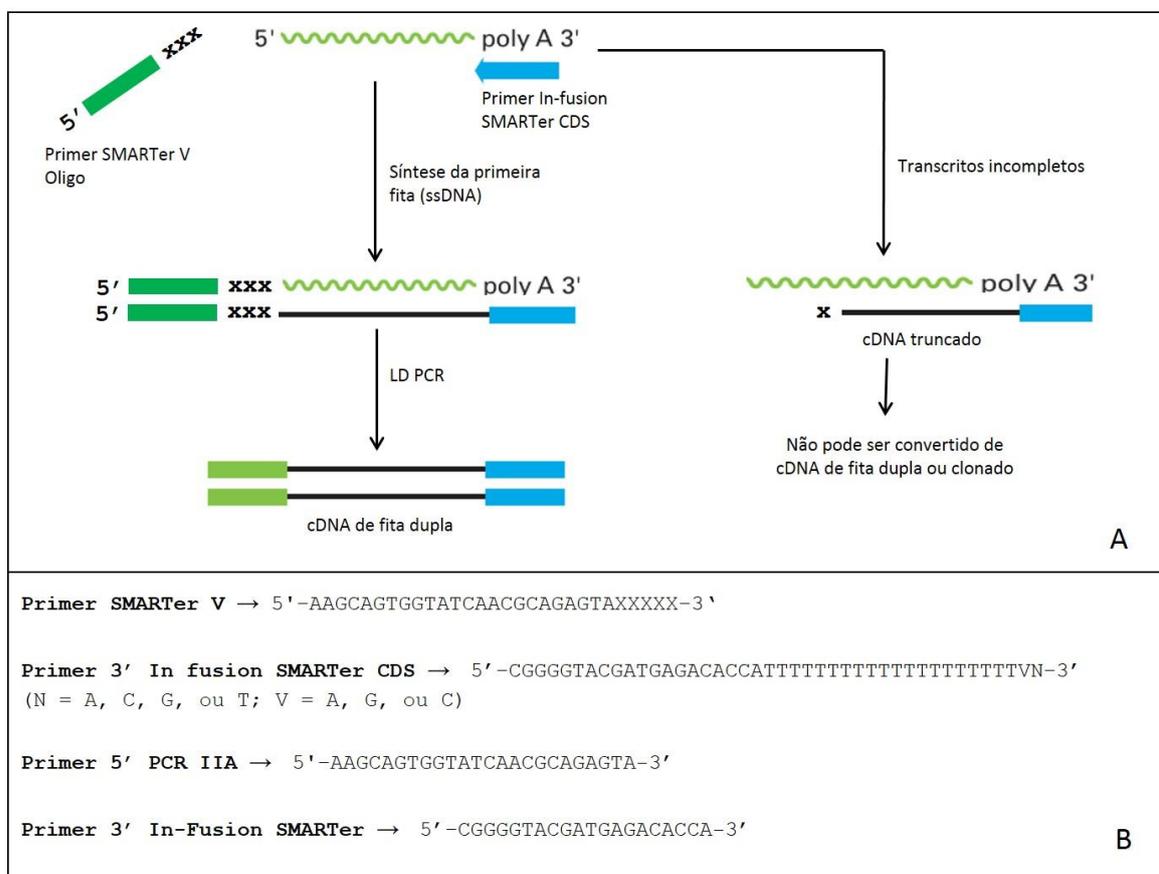
#### 3.5.1 Síntese de cDNA por LD-PCR

O PCR de longa distância (*Long Distance* PCR/LD-PCR) é uma ferramenta útil para amplificação de amostras de DNA de até 100 Kb, permitindo a síntese completa do cDNA (todo o comprimento) (FOORD; ROSE, 1994). A síntese do cDNA foi realizada de acordo com recomendações do fabricante do IN-FUSION™ SMARTer™ cDNA Library Construction Kit. Desta forma foram misturados e homogeneizados à amostra de RNA (1 µg), o primer 3' *In-Fusion SMARTer CDS* (1 µL) e água deionizada para completar um volume de 4.5 µL. O tubo foi incubado a 72°C/3 min e logo após a temperatura foi reduzida para 42°C /2 min. Seguiu-se à adição do tampão para síntese de fita simples (2 µL), DTT (0.25 µL), mistura de dNTP (1 µL), do

oligonucleotídeo (SMARTer II A), do inibidor de RNase (0,25  $\mu$ L), do primer SMARTer V oligonucleotídeo (1  $\mu$ L) e da enzima transcriptase reversa (1  $\mu$ L) (SMARTScribe Reverse Transcriptase) com atividade nucleotídeo (citosinil) transferase. Após homogeneização o tubo foi incubado a 42°C/90 min, e em seguida a 68°C/10 min, para finalizar a síntese da primeira fita.

A primeira fita resultante foi utilizada para síntese do DNA de fita dupla e amplificação. Em um tubo de reação foram adicionados e homogenizados o cDNA (2  $\mu$ L), água mili-Q autoclavada (80  $\mu$ L), tampão para PCR 10x (10  $\mu$ L), mistura de dNTP (2  $\mu$ L), primer 3' In-Fusion SMARTer (2  $\mu$ L), o primer 5' PCR IIA (2  $\mu$ L) e a polimerase Advantage 2 fornecida pelo kit (2  $\mu$ L). Logo após, o tubo foi colocado no termociclador para o primeiro ciclo a 95°C/1 min, o qual foi seguido por 27 ciclos a 95°C/15 seg, 65°C/30 seg e 68°C/6 min. Em seguida 5  $\mu$ L do produto do PCR foi analisado, através de gel de agarose (1,2 %) utilizando um marcador molecular de DNA (1 Kb). O gel foi marcado com brometo de etídio para visualização do cDNA em luz ultravioleta. Esta análise eletroforética serviu para determinarmos a qualidade do cDNA produzido e o perfil do tamanho em pares de base do cDNA sintetizado. Um cDNA de boa qualidade apresenta um arrastado no gel normalmente entre 0,1 a 4 Kb. As amostras foram estocadas a -20°C (Figura 3A e B).

**Figura 3:** Esquema da síntese do DNA de *D. quadriceps*. A) Síntese de cDNA de fita simples e dupla (LD = long distance). B) Primers utilizados para a síntese e amplificação do cDNA (Clontech Laboratories, 2012).



### 3.5.2 Purificação e fracionamento do cDNA dupla fita

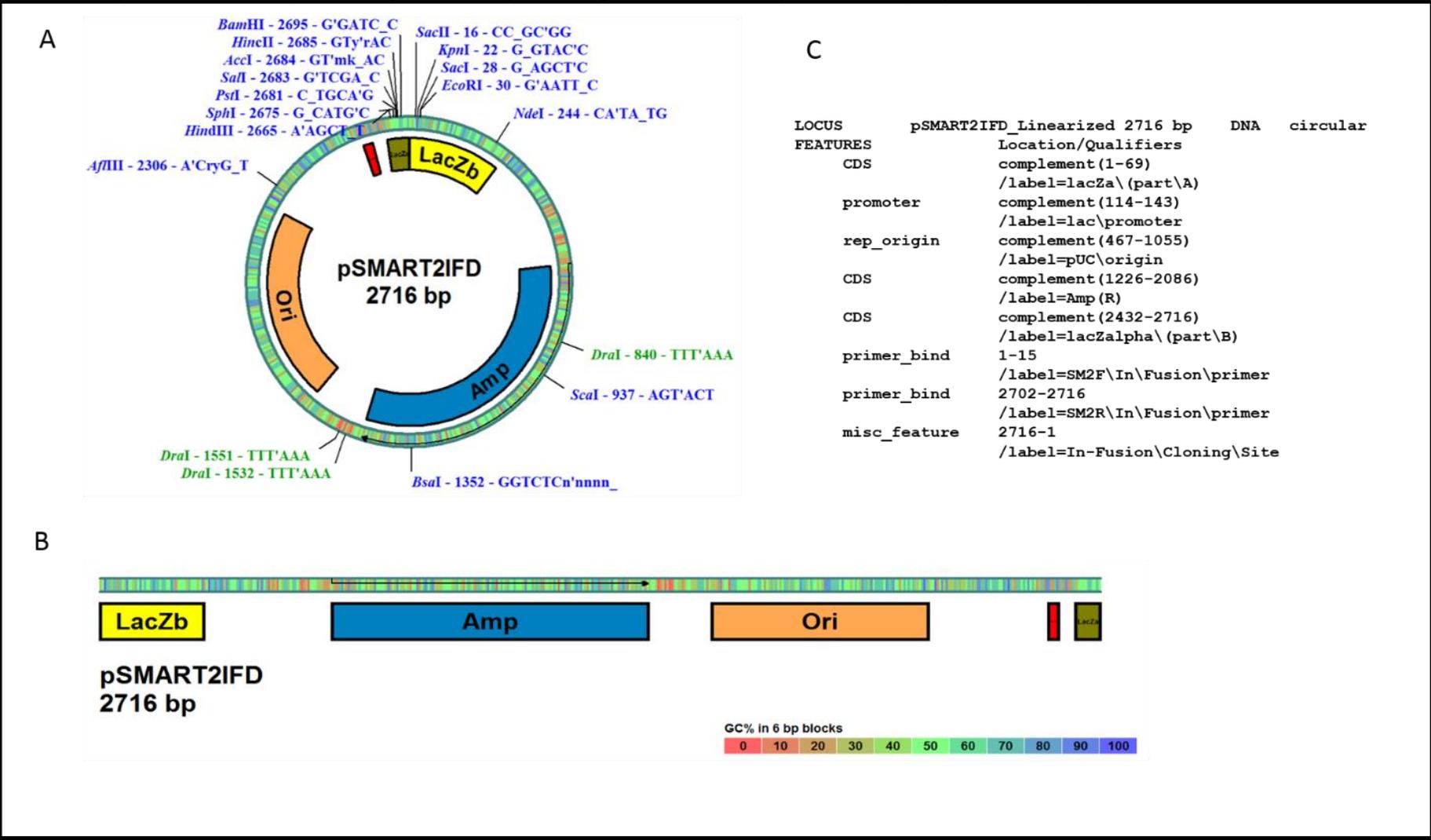
A purificação e fracionamento do cDNA de dupla fita foi realizada com uma coluna cromatográfica CHROMA SPIN™ DEPC (CHROMA SPIN 400, Clontech, USA). A mesma foi equilibrada com tampão e uma amostra de 85 µL do produto de PCR foi adicionada à mesma. As frações foram coletadas, avaliadas através de eletroforese em gel de agarose 1,1%, para estudo da distribuição de tamanho. Em seguida foram combinadas e precipitadas com NaCl/Etanol. O cDNA foi quantificado através de espectrofotometria UV/ 260nm.

### 3.5.3 Clonagem do cDNA no vetor pSMART2IFD

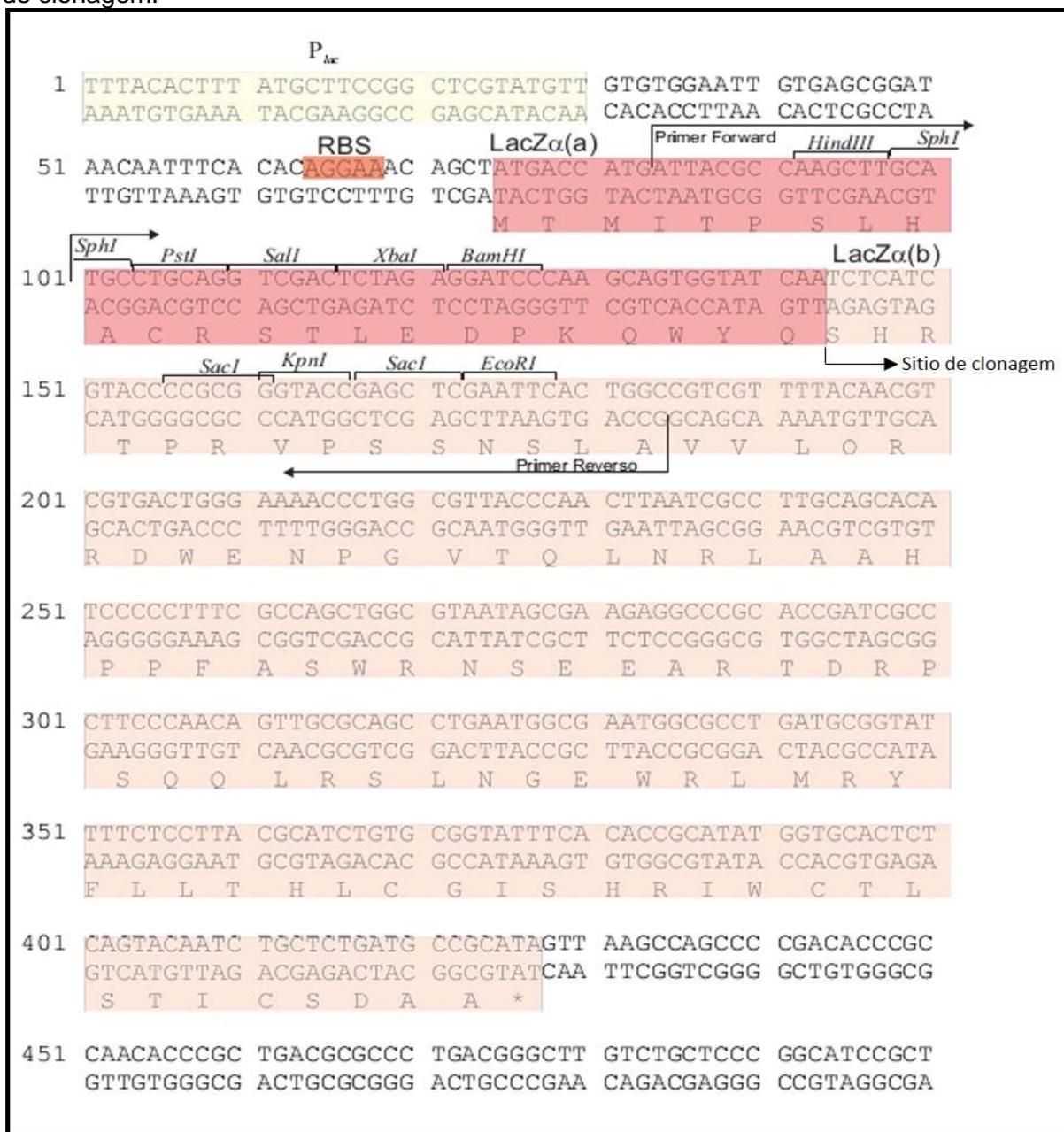
Em tubo de 0,5mL foram adicionados diferentes razões molares entre inserto e vetor (A=2:1; B=2.25:1; C=2.5:1); 2 µL do vetor (pSMART2IFD, 300 ng); 2

$\mu\text{L}$  do tampão de ligação (In-Fusion Reaction Buffer); 1  $\mu\text{L}$  de DNAligase e água deionizada para um volume total de 10  $\mu\text{L}$ . Estes foram incubados por 15 min/50°C. Após este período os tubos foram mantidos em gelo. À estes adicionou-se 90  $\mu\text{L}$  de água estéril e 1.5  $\mu\text{L}$  de glicogênio. Após homogeneização foi adicionado 280  $\mu\text{L}$  de etanol 100%. Os tubos foram incubados a -70°C/1-2 h. Após esse período foram centrifugados a 15.000 rpm/20 min e então, o sobrenadante removido, sendo o pellet ressuspendido em 10  $\mu\text{L}$  de água estéril. As figuras 4 e 5 mostram o vetor pSMART2IFD, o qual apresenta terminações cegas em seu sítio de clonagem, situado no gene *lacZ $\alpha$*  (região que codifica a porção N-terminal para a  $\beta$ -galactosidase, em amarelo escuro). A expressão deste dirige para a região do lac promotor (*P<sub>lac</sub>*, em vermelho). Este sítio de clonagem permite a seleção branco/azul após a clonagem, e desta forma auxilia na identificação de clones cujos vetores de fato receberam o inserto, ver item 3.5.4. O vetor também apresenta a região pUC ori de replicação (em laranja), que permite a propagação em *E. coli* e um gene de resistência a ampicilina (*Amp<sup>r</sup>*, em azul) para seleção dos clones que receberam o vetor.

**Figura 4:** Detalhamento do vetor pSMART2IFD. A) Vetor circular pSMART2IFD. Em amarelo, região que representa o gene lacZ; em azul, gene de resistência à ampicilina; em laranja gene pUC ori de replicação; em vermelho, o gene promotor (P<sub>lac</sub>) de regulação da transcrição. B) Vetor linear pSMART2IFD. C) Localização das sequências genicas de LacZ, pUC ori, Amp e primers de sequenciamento.



**Figura 5:** Sequência linear do sítio de clonagem do vetor pSMART2IFD. Em amarelo claro a região promotora ( $P_{lac}$ ); em rosa escuro região LacZ $\alpha$ (a); em rosa claro região LacZ $\alpha$ (b). Setas indicam regiões de flanquamento dos primers de sequenciamento senso e anti-senso, e sítio de clonagem.



### 3.5.4 Transformação de *E. coli* com plasmídio recombinante

Cepas de *E. coli* TOP10 foram utilizadas para manutenção e propagação da biblioteca de cDNA. Para tanto adicionou-se o volume dos tubos de ligação (A, B e C) ao volume de 25  $\mu$ L de células de *E. coli* TOP10 eletrocompetentes ( $>4 \times 10^9$  ufc/ $\mu$ g). Esta mistura foi eletroporada em cubetas refrigeradas de 0,1 cm. O volume foi

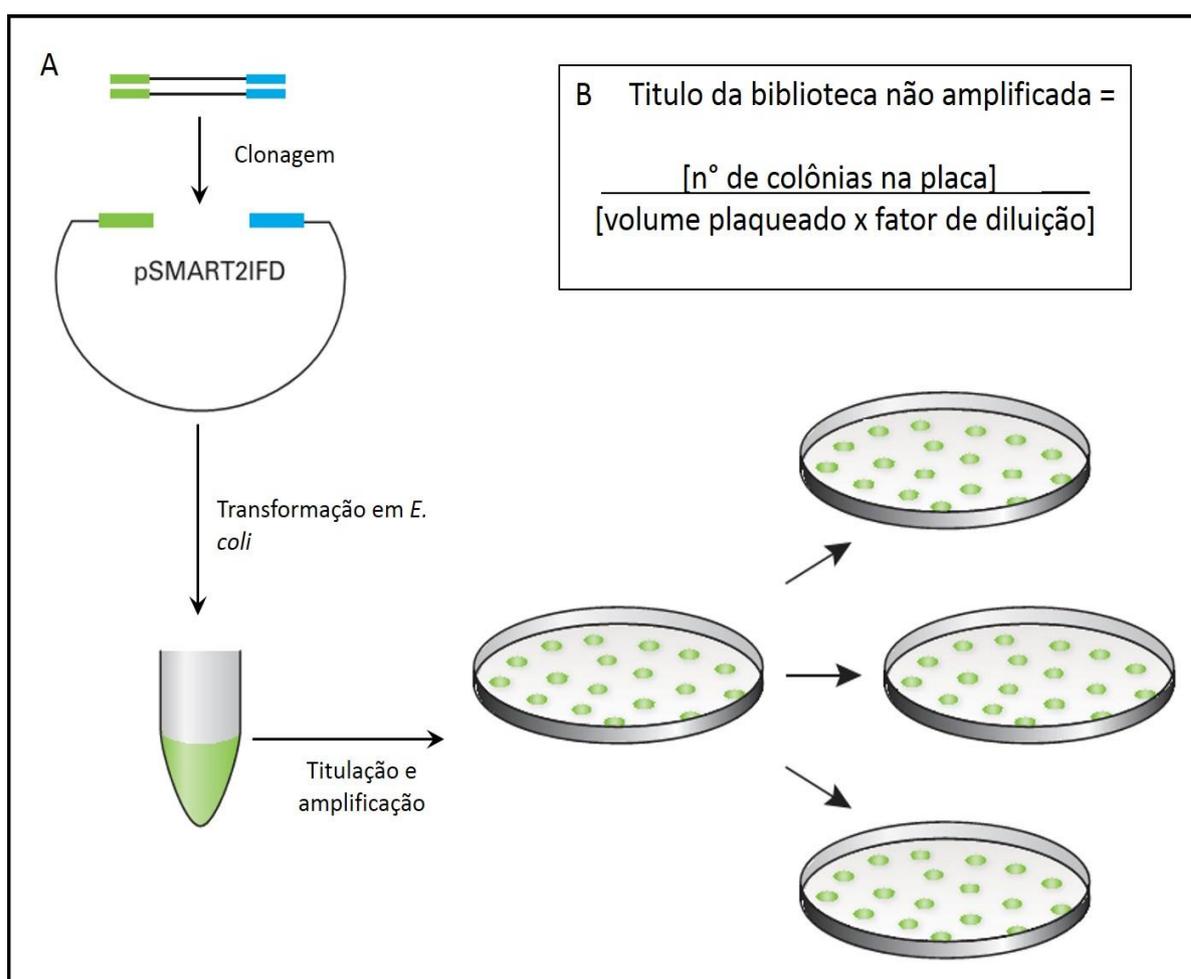
então, transferido para tubos de polipropileno contendo 965  $\mu\text{L}$  de meio LB (Luria-Bertani) e incubados, sob agitação, a  $37^\circ\text{C}/1$  h. Em outros três tubos foram adicionados 50  $\mu\text{L}$  de caldo LB e, ao final da incubação, foi retirado 2  $\mu\text{L}$  da mistura anterior de cada tubo e adicionados à tubos contendo 50  $\mu\text{L}$  de caldo LB. Uma alíquota da mistura foi subcultivada em Ágar LB contendo 100  $\mu\text{g}/\text{mL}$  de ampicilina, 1 mM de IPTG, 75 mg/mL de X-Gal e incubada a  $37^\circ\text{C}/\text{overnight}$ . Esta última etapa nos permitiu avaliar a proporção entre colônias brancas/azuis, método para detecção das bactérias recombinantes. As células transformadas com o vetor contendo o DNA recombinante produzem colônias brancas, enquanto que células transformadas com plasmídeos não recombinantes produzem colônias azuis. Isto é possível devido a presença do gene LacZ, o qual codifica a  $\beta$ -galactosidase. Esta, funcionalmente, apresenta-se na forma de um tetrâmero, composta por peptídeos  $\omega$  e  $\alpha$ . Um mutante  $\beta$ -galactosídeo com o resíduo N-terminal deletado (peptídeo  $\omega$ ) é incapaz de formar o tetrâmero, e portanto é inativo. No entanto, esse mutante forma uma proteína que pode ser tornar ativa novamente em presença do N-terminal (peptídeo  $\alpha$ ). Assim, neste método, a *E. coli* apresenta o gene LacZ mutante, o qual codifica o peptídeo  $\omega$ , enquanto o vetor utilizado apresenta o gene LacZ $\alpha$ , o qual codifica o peptídeo  $\alpha$ . Quando os peptídeos são expressos conjuntamente, formam a  $\beta$ -galactosidase funcional. Estando o sítio de clonagem do vetor pSMART2IFD dentro da sequência LacZ $\alpha$ , os vetores que receberam o inserto rompem o gene e não serão mais capazes de produzir os peptídeos  $\alpha$ . Portanto as células que receberam vetor com inserto não produzem  $\beta$ -galactosidase funcional, sendo representadas pelas colônias brancas. Já as células que receberam vetores sem inserto matêm o gene LacZ $\alpha$  íntegro e portanto produzem  $\beta$ -galactosidase funcional, cuja presença é detectada utilizando X-Gal, um análogo da lactose, que, uma vez clivado pela enzima forma 5-bromo-4-cloro-indoxil, o qual espontaneamente dimeriza e oxida formando um pigmento azul insolúvel (MESSING *et al.*, 1977).

As bactérias que não transformadas também foram impossibilitadas de crescer neste meio devido à presença de ampicilina. As placas que apresentaram as bactérias transformadas confluentes foram comparadas com o controle negativo (sem crescimento, amostra sem DNA) e positivo (crescimento, amostra contendo plasmídeo intacto).

Para garantir uma que uma biblioteca representativa foi obtida, realizou-se a titulação da mesma, a qual deve apresentar-se pelo menos 10x superior ao número de clones independentes. Dez microlitros da biblioteca foram adicionados a 1 mL de caldo LB (Diluição A; fator de diluição  $10^{-2}$ ). Desta primeira diluição foram retirados 10  $\mu$ L e adicionados à 1 mL de caldo LB (Diluição B; fator de diluição  $10^{-4}$ ). Por fim 10  $\mu$ L da diluição A foram misturados à 50  $\mu$ L de caldo LB. Em seguida alíquotas de 50  $\mu$ L foram plaqueadas em ágar LB e incubadas a 37°C/overnight. As placas que apresentaram 30-300 colônias foram quantificadas para determinar o título, conforme a equação na figura 6B.

Em seguida a biblioteca foi amplificada, diversos clones foram isolados randomicamente em placas de 96 poços e estocados a -80°C em caldo LB/25% de glicerol (Figura 6).

**Figura 6:** Esquema simplificado da construção da biblioteca de cDNA (Clontech Laboratories, 2012).



### **3.5.5 Amplificação do DNA dos clones obtidos**

Uma alíquota de ~5 µL do clones isolados randomicamente foi subcultivada em caldo LB, e cerca de 2 µL do subcultivo de cada clone foi submetido à um PCR de colônias, utilizando primers específicos para o vetor empregado na síntese da biblioteca (Figura 5). Utilizamos, para tanto, a enzima GoTaq Green (Promega, USA) e, como protocolo, um primeiro ciclo de 95°C/5 min seguido de 30 a 35 ciclos de desnaturação a 95 °C/30 seg; anelamento a 60 °C/30 seg; e extensão a 72°C/1 min. Após os 30 a 35 ciclos de amplificação realizamos uma extensão final a 72°C/10 min. Todos os produtos obtidos foram avaliados em gel de agarose 1,5% e fotodocumentados através do equipamento Chemidoc XRS (Bio-Rad/USA). Os produtos de PCR de 432 clones independentes foram purificados através de tratamento com EXO/SAP e precipitação com acetato de sódio/Etanol. O DNA purificado foi resuspenso em formamida deionizada e submetido a sequenciamento pelo método de Sanger (1977).

### **3.5.6 Análise, organização e agrupamento dos clones**

Os produtos da PCR purificados foram sequenciados utilizando o sequenciador automático de DNA 3100Avant (Applied Biosystems, USA) e o kit ABI Prism BigDve Terminator v3.0 Cycle Sequencing Ready Reactions (DETWILER *et al.* 2004). As sequências obtidas foram analisadas utilizando o software CLC Bio Main Workbench v6.8.4. O escore de qualidade Phred, comumente utilizado para caracterizar a qualidade das sequências de DNA durante o sequenciamento automático, foi utilizado para selecionar sequências de alta qualidade (Phred >40, Acurácia de 99,99%) (EWING *et al.*, 1998; EWING; GREEN, 1998). Estas foram trimadas para excluir contaminação do vetor e a cauda poliA. Para agrupamento dos clones, os seguintes parâmetros foram utilizados: comprimento mínimo de 50 pb e média estringência de alinhamento (identidade entre as sequências). Os contigs (conjunto de clones cujos segmentos de DNA se sobrepõem e, juntos, representam um consenso de DNA) e singletos (sequências não agrupadas) resultantes foram comparados com outras sequências depositadas em banco de dados de proteínas (UNIPROT) não redundantes (nr)/GenBank (*National Center for Biotechnology*

*Information* (NCBI), <http://www.ncbi.nlm.nih.gov>), utilizando a ferramenta BlastX em banco de dados para todos os organismos e/ou artrópodes, para identificar proteínas que apresentavam alta similaridade com os transcritos e proceder a anotação de função.

### **3.6 Sequenciamento *NGS* (*Next Generation Sequencing*) utilizando tecnologia Ion Torrent**

A tecnologia Ion Torrent baseia-se na detecção de íons hidrogênios liberados a cada ciclo de polimerização do DNA. Neste caso são utilizados chips semicondutores capazes de transformar sinais químicos em informação digital. Estes apresentam centenas de micropoços que abrigam microesferas, às quais estão associadas ao DNA “molde”. Os nucleotídeos são fornecidos em uma sequência conhecida e, uma vez que complementa a sequência da molécula de DNA de um determinado micropoço, ele será incorporado e um íon  $H^+$  será liberado. O pH da solução então muda naquele poço e é detectado pelo sensor.

Foi gerada uma biblioteca de fragmentos a partir de 100ng de RNA, obtido no item 3.4, utilizando o kit Ion Xpress Plus Fragment Library (Life Technologies, Carlsbad, USA). Para tanto, realizou-se fragmentação de DNA, ligação de adaptadores, seleção de fragmentos por tamanho e amplificação da biblioteca. Um total de oito ciclos foram realizados de acordo com recomendações do fabricante. O tamanho dos fragmentos predito foi de aproximadamente 330 pb, sendo o ensaio realizado em cassetes com gel de agarose 2% para o instrumento Pippin Prep (Sage Science, Beverly, USA). A distribuição de tamanhos e a concentração da biblioteca foram determinados utilizando o kit DNA High Sensitivity (Agilent, Santa Clara, CA) e o equipamento Bioanalyzer 2100. A biblioteca de fragmentos foi ajustada para 26pM e amplificada no Ion Sphere particles™ (ISPs) por emulsão de PCR utilizando o equipamento Ion OneTouch™, com o kit Ion OneTouch™ 200 Template (Life Technologies, Grand Island, NY) e amostras ISPs positivas, de acordo com as recomendações do fabricante. As ISPs foram enriquecidas e sequenciadas no chip Ion 316™ utilizando o equipamento Ion Torrent Personal Genome Machine (PGM™) (Life Technologies, San Francisco, CA, USA) por 130 ciclos com o kit Ion PGM™ 200 Sequencing.

### 3.6.1 Organização e agrupamento das sequências obtidas por NGS

De acordo com o sequenciamento através do Ion Torrent Personal Genome Machine (PGMTM), o sequenciamento pode gerar fragmentos de 200-500pb de cada fragmento de cDNA Single-End (SE). As leituras foram processadas de modo a retirar nucleotídeos referentes a sequência dos adaptadores e para remover sequências de baixa qualidade (leituras com mais de 50% de bases com valor de qualidade  $\leq 5$ ). Após esta etapa o agrupamento *de novo* foi realizado utilizando o programa Trinity (GRABHERR *et al.*, 2011).

Todos os transcritos agrupados foram pesquisados no banco de dados para proteínas não redundantes (NCBI), utilizando a ferramenta BlastX para identificar proteínas que apresentam alta similaridade com os transcritos e proceder a anotação de função, assim como realizado com as sequências obtidos por sequenciamento por Sanger.

### 3.7 Análise de sequências específicas de transcritos

Para sequências específicas de transcritos, principalmente àquelas que foram classificadas como toxinas polipeptídicas, análise de similaridade foi realizada utilizando a ferramenta BlastP, sistematicamente contra banco de dados nr e Uniprot, e subsequentemente restritos à Artrópodos, Himenópteros ou Formicidae. Em seguida a análise foi estendida para banco de dados mais específicos como ArachnoServer, Web Konttin Data Base (GELLY *et al.*, 2004). As sequências de aminoácidos preditas foram alinhadas utilizando o software ClustalW (<http://www.ebi.ac.uk>) e BLSP-MACAW, os potenciais sítios de clivagem putativos referentes ao peptídeo sinal foram preditos utilizando o software SignalP (<http://www.cbs.dtu.dk/services/SignalP>). Os pontos isoelétricos e peso molecular foram preditos utilizando a ferramenta EXPASY/Compute pI/Mw tool ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/)). Para a predição estrutural utilizou-se os softwares Chimera, SPDBV e o servidor Modeller do SuussModel. O cDNA foi organizado utilizando BlastX por similaridade conhecida ou detectando-se as extremidades 5' e 3' UTR (Untranslated region), o ORF (Open Reading Frame, utilizando o ORF Finder/NCBI – <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) com os códons de iniciação e parada,

análise da sequência Kozak (representa o ponto de início da tradução, ATG de iniciação) e do sinal de poliadenilação. As figuras dos contigs exibem o peso molecular do peptídeo maduro predito. Os dendogramas de simirialidade entre precursores foram criados utilizando o programa MEGA 5, utilizando o método NJ (Neighbor-joining), o qual constrói uma árvore filogenética utilizando uma matriz de distância gênica de acordo com critérios mínimos de evolução (SAITOU; IMANISHI, 1989).

### **3.8 Análise proteômica da peçonha total de *D. quadriceps***

#### **3.8.1 Extração da peçonha**

A cada duas semanas, cada espécime, mantida em cativeiro, foi imobilizada com uma pinça flexível e a parte posterior de seu gáster foi levemente pressionada por um tubo capilar. A peçonha total de *D. quadriceps* foi extraída, prontamente congelada, liofilizada e armazenada a -20°C, para posterior realização dos ensaios.

#### **3.8.2 Dosagem de proteína total pelo Método Bradford**

Este método é baseado na interação entre o corante Coomassie brilliant blue BG-250 e macromoléculas de proteínas que contêm aminoácidos de cadeias laterais básicas ou aromáticas. No pH de reação, a interação entre a proteína de alto peso molecular e o corante provoca deslocamento do equilíbrio do corante para a forma aniônica, que absorve fortemente em 595nm (BRADFORD, 1976). 100µL da peçonha diluída (2mg/mL) em tampão acetato foi misturado com 2,5mL da solução de Bradford (Coomassie Blue BG-250 a 10,24%; etanol a 5% e ácido fosfórico 10%) e foi realizada a leitura em espectrofotômetro com comprimento de onda 595nm. A concentração proteica foi calculada de acordo com a equação  $C = (F \times A \times D)/100$ , onde:

F = fator de calibração

A = absorvância

D = fator de diluição

### **3.8.3 Eletroforese SDS-PAGE**

A peçonha total de *D. quadriceps* foi submetida à eletroforese em gel de poli(acrilamida), segundo Laemmli (1970). No sistema eletroforético foi utilizado gel de empilhamento (superior) contendo acrilamida a 5% em Tris-HCl 0,5 M; pH 6.8 e SDS a 0,1% e um gel de separação (inferior) contendo acrilamida a 12% em Tris-HCl 2,0M; pH 8.8 e SDS a 0,1%. Os géis foram polimerizados, com auxílio de TEMED (N,N,N',N'-tetramethylethylenediamine), em placas de vidro, as quais foram acopladas a um reservatório contendo tampão composto de Tris (0,036 M), glicina (0,072 M) e SDS a 0,1% em pH 8.3. A amostra foi dissolvida em Tampão Tris-HCl, pH 6.8, contendo glicerol,  $\beta$ -mercaptoetanol (5%), SDS (2%) e azul de bromofenol a 0,001%; e aplicada em poços feitos no gel de acrilamida. Para avaliação da distribuição de peso molecular foi utilizado um padrão de 15-250 kDa (Promega). O gel foi visualizado utilizando Coomassie Blue R-250.

### **3.8.4 Purificação por cromatografia líquida de alta eficiência em fase reversa da peçonha bruta**

A quantidade de peçonha bruta empregada nas purificações por HPLC (High Performance Liquid Chromatography) de fase reversa foi estimada com base no peso seco da mesma, obtido através de liofilização. Inicialmente, 1 mg de peçonha bruta liofilizada foi diluída em 1 mL de uma solução de água: acetonitrila (9:1) contendo 0,1% de ácido trifluoroacético (TFA), de modo a se atingir uma concentração final de 1  $\mu\text{g}/\mu\text{L}$ . O material ressuspenso foi filtrado em membranas de 0,22  $\mu\text{m}$  (Millex, Millipore, Irlanda) e posteriormente submetido a purificação de HPLC de fase reversa.

A peçonha bruta foi injetada (1 mL/min), em quantidades que variaram de 60 a 400  $\mu\text{g}$  por corrida, em um sistema de purificação Äkta Purifier (GE Healthcare, Uppsala, Suécia), tendo como soluções de eluição os denominados solventes A (água contendo 0,1% de TFA) e B (acetonitrila contendo 0,1% de TFA). A eluição e a separação dos componentes foram realizadas empregando-se colunas analíticas de fase reversa C-8 e C-18 (250 X 4.6 I.D., 5  $\mu\text{m}$ , Vydac, EUA) através de gradientes lineares de 10 a 90% de B em 30 minutos, ou de 10 a 60% de B em 40 minutos, de modo a se otimizar a separação dos compostos. As amostras eluídas foram

monitoradas em UV a 214 e 280 nm, sendo integradas e registradas na forma de picos. O programa utilizado para controle do sistema cromatográfico, bem como para a manipulação gráfica dos cromatogramas obtidos foi o Unicorn 7.0, fornecido pelo fabricante.

Todos os picos obtidos foram coletados manualmente e individualmente ou agrupados em sub-frações, tendo sido secos em speed-vac para serem posteriormente submetidos à análise por espectrometria de massas e repurificações.

### ***3.8.5 Análise por cromatografia a líquido acoplada a espectrometria de massas (LC-MS) da peçonha bruta de *D. quadriceps****

A peçonha bruta foi analisada em um sistema de LC-MS Thermo Scientific, equipado com sistema de HPLC acoplado a um detector Accela de arranjo de fotodiodo em um instrumento LTQ-Orbitrap XL (Thermo Fisher Scientific, USA). As amostras de peçonha foram eluídas em um gradiente de 0-65% de solvente B (acetonitrila contendo 0,1% de ácido fórmico) contra solvente A (água desionizada contendo 0,1% de ácido fórmico) durante 20 minutos, elevando-se o gradiente a 85% de B por 5 minutos. O fluxo de eluição empregado foi de 0,2 mL/min usando-se uma coluna Capcell Pak C18 UG-120 (1.5 mm i.d. x 150 mm; Shiseido Co.) mantida aquecida a 35°C. O eluato da coluna foi aplicado diretamente a uma interface de ionização electrospray (ESI) do espectrômetro de massa. As condições de ionização e detecção foram as seguintes: modo positivo ESI foi aplicado com uma voltagem de spray de 5,0 kV/m temperatura capilar de 350°C, fluxo de gás de arraste de 40 (unidades arbitrárias), fluxo de gás auxiliar de 10 (unidades arbitrárias), voltagem capilar de 10 V e voltagem de tubos de lentes de 20 V. Um conjunto de varredura de espectros simples (MS) bem como espectros subsequentes de íons precursores fragmentados (MS/MS) foi aplicado através das análises, sendo que espectros de alta qualidade foram gravados com um modo de varredura completa (full scan) em uma resolução de 60,000 (em m/z 400) com uma faixa de detecção de m/z 50-2000; Os dois sinais mais abundantes no espectro de varredura completa foram selecionados para prover alta resolução (resolução de 30.000) de espectros de MS/MS.

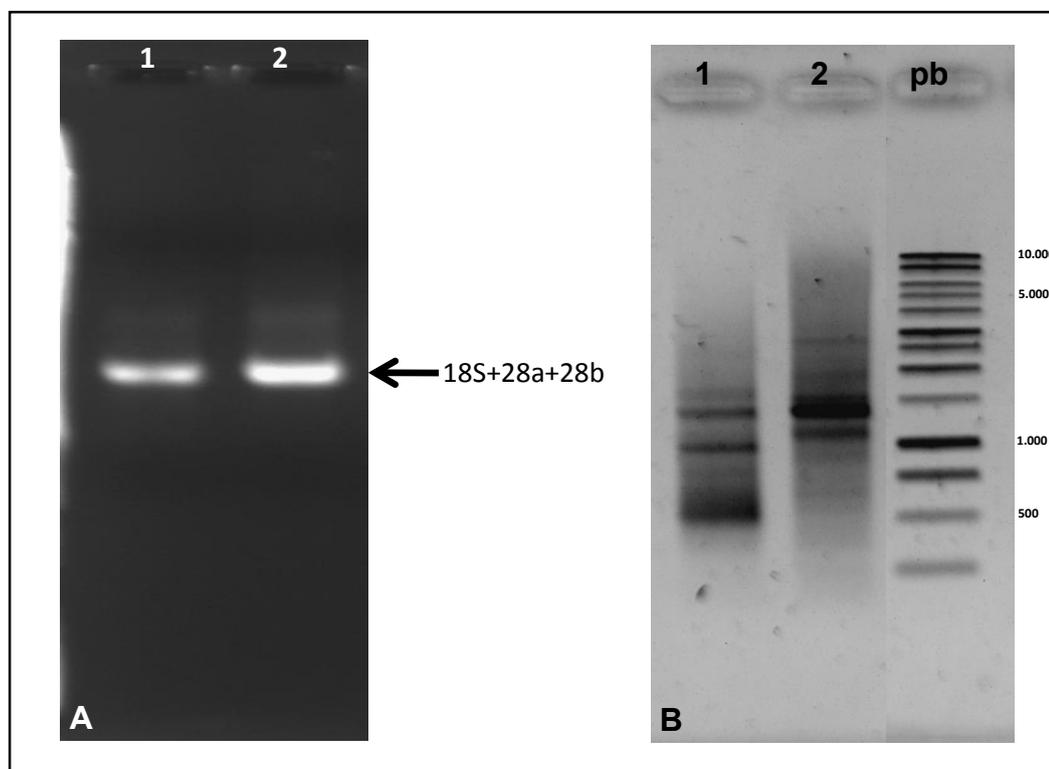
## 4 RESULTADOS E DISCUSSÃO

### 4.1 Extração do RNA total e construção da biblioteca de cDNA da glândula de veneno de *D. quadriceps*

A biblioteca de DNA complementar (cDNA) foi preparada a partir de 20 glândulas de veneno de *D. quadriceps* para identificar os mais importantes transcritos expressos na peçonha. A partir das glândulas foi extraído RNA total e a qualidade do mesmo foi avaliada através de eletroforese de RNA em condições desnaturantes (Figura 7A), onde observamos que o RNA extraído apresentou um perfil característico de migração de RNA de insetos, com banda única decorrente da separação das subunidades  $\alpha$  e  $\beta$  que compõem a porção 28S do RNA de insetos, e que migram juntamente com a porção 18S (WINNEBECK *et al.*, 2010). Assim, pudemos concluir que o RNA estava em condições adequadas para seguirmos com a síntese do cDNA. Obteve-se aproximadamente 1.374 ng/ $\mu$ L de RNA total.

A biblioteca foi construída utilizando o kit In-Fusion SMARTer (Clontech, USA). Utilizamos 1  $\mu$ g do RNA total para a síntese do cDNA. Esta baseou-se na ação do primer SMARTer 3'CDS complementar à cauda poliA do RNA mensageiro (mRNA) de eucariotos, e adição de uma extensão molde independente à 5' em decorrência da ação citosinil transferase terminal da enzima transcriptase reversa, permitindo, desse modo, adaptar um primer 5' específico (SMARTer V oligo). Assim, a subsequente síntese do cDNA de fita dupla ocorre a partir das sequências de cDNA de fita simples que apresentam o sítio de ligação proporcionado pela ação citosinil tranferase da enzima transcriptase reversa e o adaptador a 3' proporcionado pelo primer SMARTer 3'CDS, utilizando os primers 5' PCR primer IIA, 3' In-Fusion SMARTer (Figura 3B) e LD-PCR, ferramenta útil para amplificação de amostras de DNA de até 100Kb (FOORD; ROSE, 1994). O cDNA dupla fita obtido apresentou uma variação de tamanho de 500 a 1.500pb (Figura 7B).

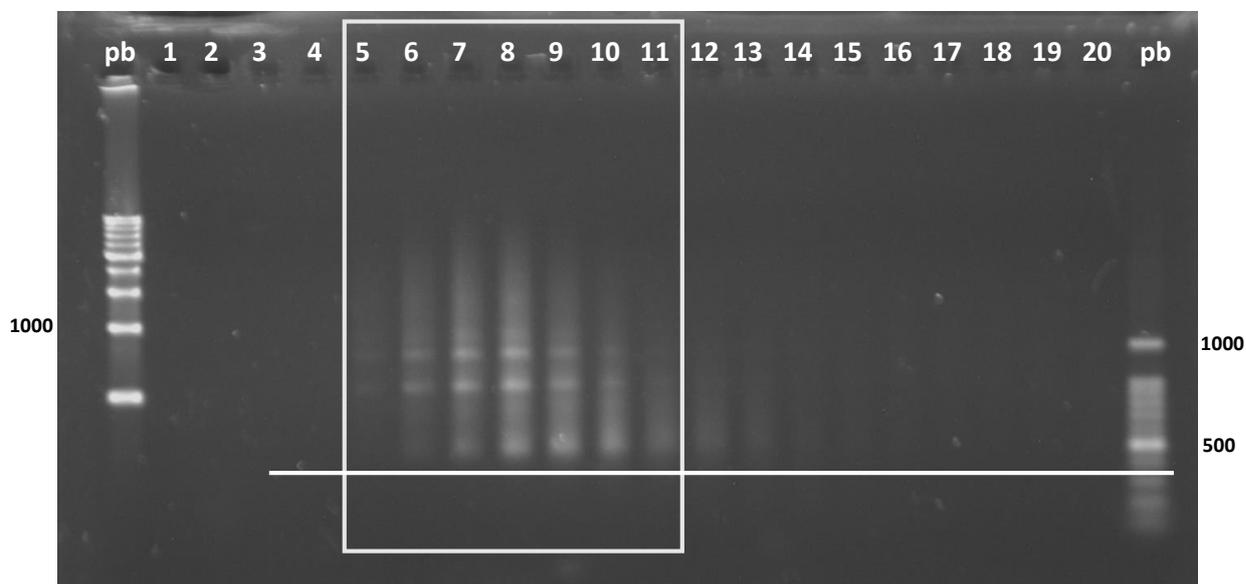
**Figura 7:** Eletroforese do RNA total extraído e do cDNA dupla fita. (A) Eletroforese de RNA em tampão TAE. Avaliação do RNA total extraído da glândula de veneno (1) e da glândula de Dufour (2) de *D. quadriceps*. (B) Eletroforese, em gel de agarose 1,5%, do cDNA dupla fita obtido a partir do RNA total extraído da glândula de veneno (1) e da glândula de Dufour (2) de *D. quadriceps*.



18S, 28a e 28b representam subunidades do RNA de artrópodes.

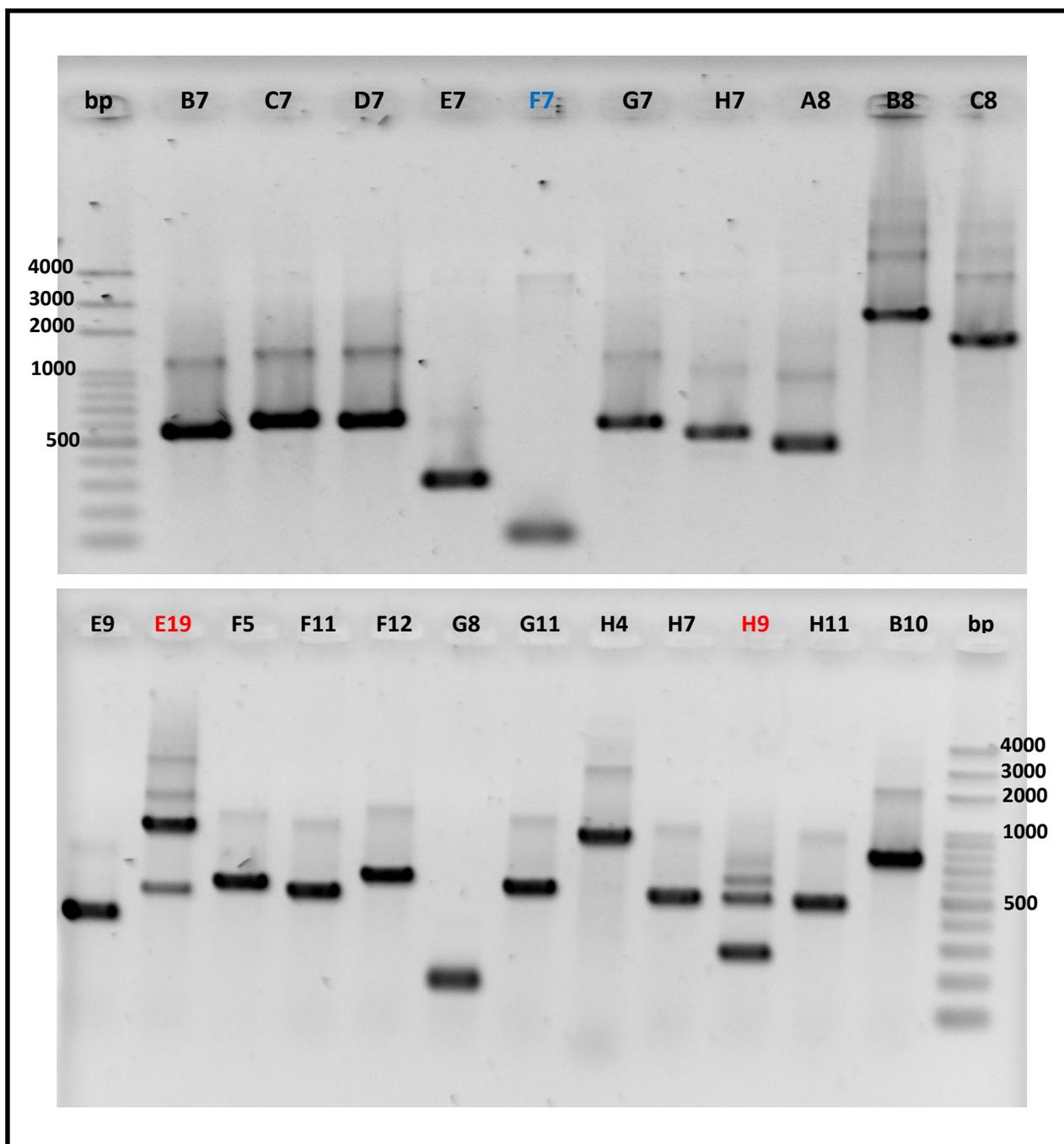
O cDNA dupla fita da glândula de veneno foi fracionado de acordo com o tamanho para a remoção de fragmentos de baixo peso molecular, pequenos DNAs contaminantes e nucleotídeos não incorporados (Figura 8). As frações 5-11, correspondentes aos cDNAs enriquecidos em alto peso, foram misturadas em um único microtubo e utilizadas para a clonagem.

**Figura 8:** Perfil de fracionamento do cDNA dupla fita obtido a partir do RNA total extraído da glândula de veneno de *D. quadriceps*.



O cDNA da glândula de veneno foi clonado no vetor linearizado pSMART2IFD fornecido pelo kit (Ver figura 5), com auxílio de uma DNAligase. A proporção ótima de cDNA e vetor definida foi de 2:1 em razão molar. Após a clonagem, *E. coli* TOP10 eletrocompetentes foram transformadas, através de eletroporação, com os plasmídeos recombinantes, e subsequentemente cultivadas em ágar LB. Obteve-se uma eficiência de clonagem de  $1 \times 10^5$  ufc/ $\mu$ g de DNA. Aproximadamente 500 clones independentes foram, aleatoriamente, isolados e amplificados, tendo apresentado um tamanho médio dos insertos de 563pb (Figura 9). Ainda na figura 9 observamos em vermelho clones não eletivos para sequenciamento pois apresentaram várias bandas no gel; e a amostra em azul representa o vetor pSMART2IFD sem inserto.

**Figura 9:** Eletroforese em gel de agarose 1,5% para avaliação do tamanho médio dos insertos obtidos a partir de PCR dos clones isolados da biblioteca de cDNA de *D. quadriceps*.

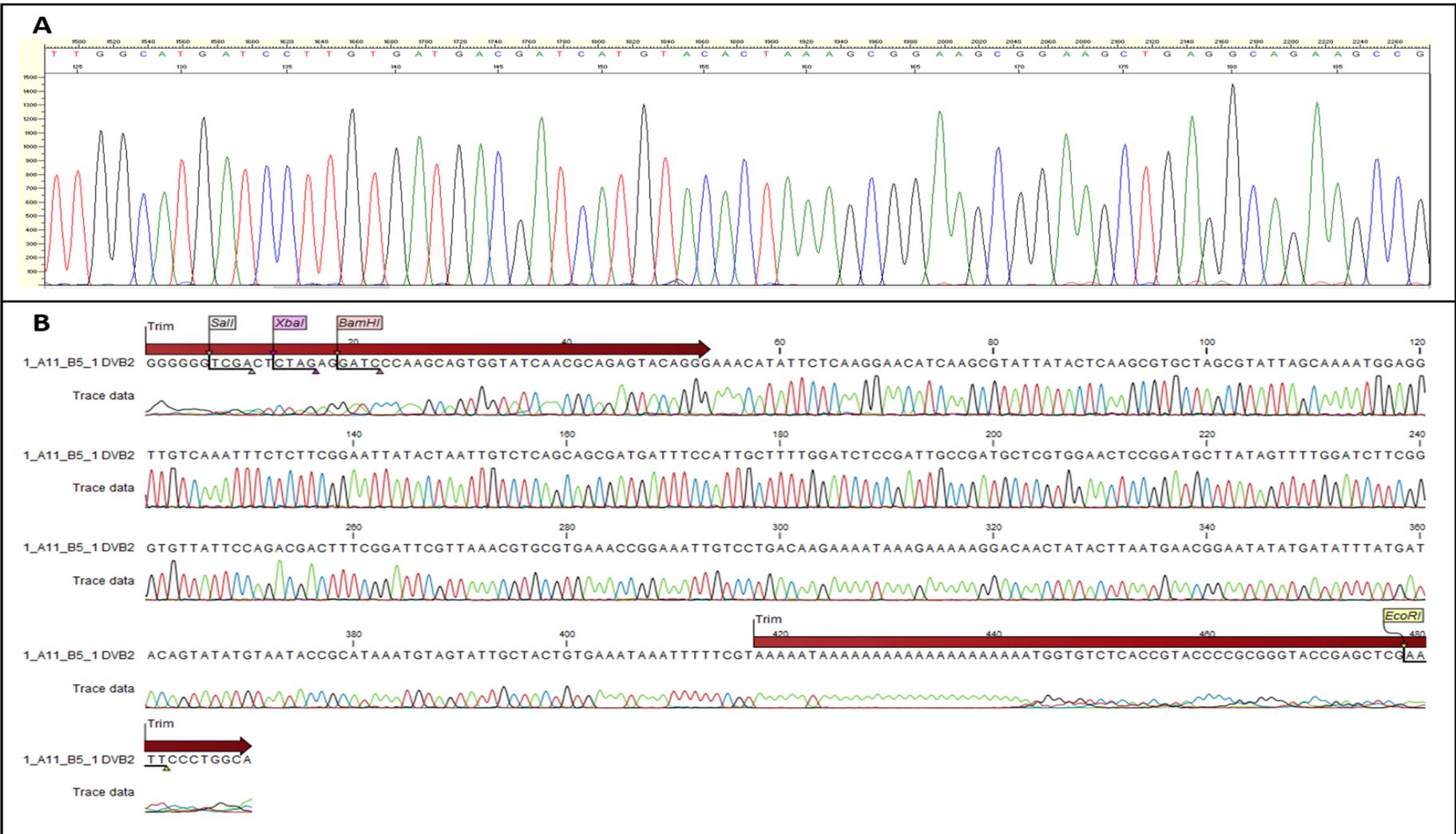


Em vermelho estão alguns clones que não foram clones aprovados para sequenciamento, em azul representa o vetor sem inserto de DNA.

## 4.2 Sequenciamento por Sanger e agrupamento das sequências

Um total de 432 clones selecionados aleatoriamente foram sequenciados pelo método enzimático utilizando didesoxirribonucleotídeos trifosfato, descrito por Sanger (1977). Os ESTs foram analisados utilizando o software CLC Main Workbench versão 6.8.4 (CLC Bio, USA). 420 sequenciamentos de alta qualidade, mensurada através do score Phred (Figura 10A e B), foram trimados para retirada de nucleotídeos que correspondiam ao vetor e à cauda poliA (Figura 10B). As sequências foram agrupadas por homologia e agrupadas em 21 contigs e 39 singletos. A seguir cada consenso foi analisado quanto à sua organização identificando o ORF, as UTRs e os sinais funcionais que identificam o precursor.

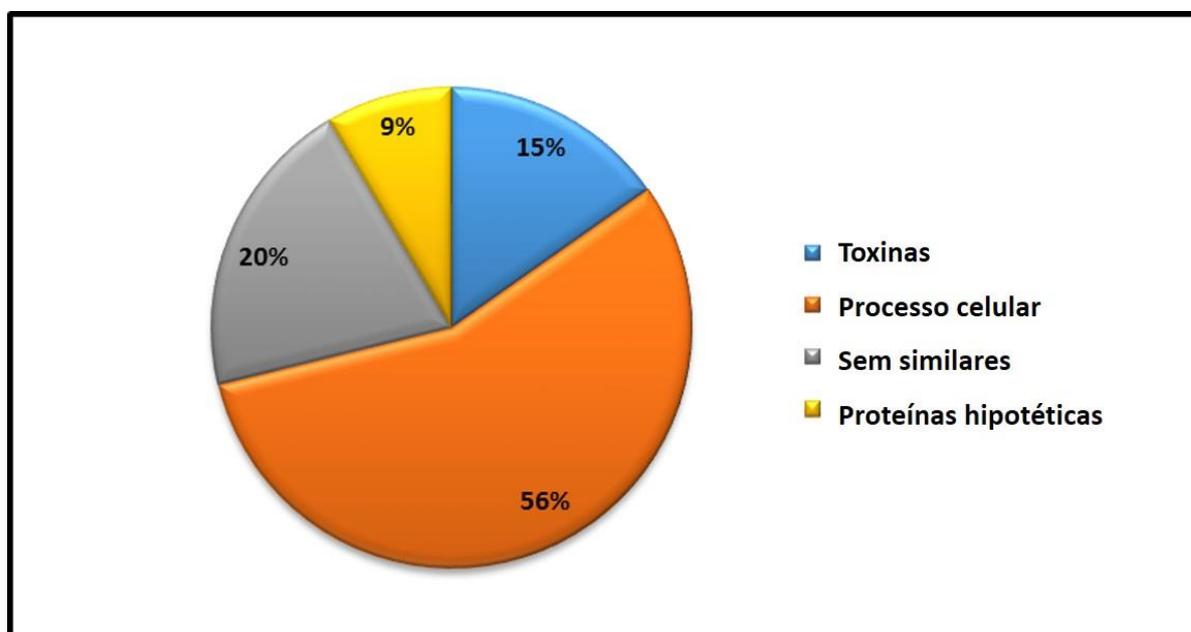
**Figura 10:** Exemplo de eletroferograma de alta qualidade do sequenciamento por Sanger de um clone da biblioteca de cDNA da glândula de veneno de *D. quadriceps*. (A) Eletroferograma de score phred >40 gerado no software Sequence scanner do sequenciador ABI 3100 Avant, Applied Biosystems. (B) Eletroferograma de score phred>40, trimado quanto ao vetor, adaptadores e cauda poliA, gerado no software CLC Bio  
Main Workbench v6.8.3.



A comparação com sequências traduzidas em bancos de dados de proteínas e de nucleotídeos foi realizada utilizando a ferramenta BlastX. A maioria das sequências nucleotídicas não apresentaram similaridade com as depositadas no banco de dados. Deste modo nossa classificação foi fundamentalmente baseada em similaridades encontradas nos bancos de dados de proteínas, o que nos permitiu a classificação e identificação dos agrupamentos em (1) proteínas envolvidas em processo celular (22,72% contigs; 69,23% singletos), por exemplo, transferases, ATP sintases, desidrogenases, proteínas ribossômicas, citocromo c, proteínas envolvidas na maturação e diferenciação de fêmeas férteis; (2) polipeptídeos associados à peçonha (42,85% contigs), devendo-se destacar a abundância das dinoponeratoxinas, peptídeos alergênicos, toxinas contendo domínios ICKs (*Inhibitor cystein knot*); (3) sequências precursoras de cDNA sem similares (23,8% contigs; 17,94% singletos) e (4) ESTs que representam proteínas hipotéticas com função desconhecida (4,76% contigs; 10,25% singletos) (Gráfico 1) (Tabelas 1 e 2).

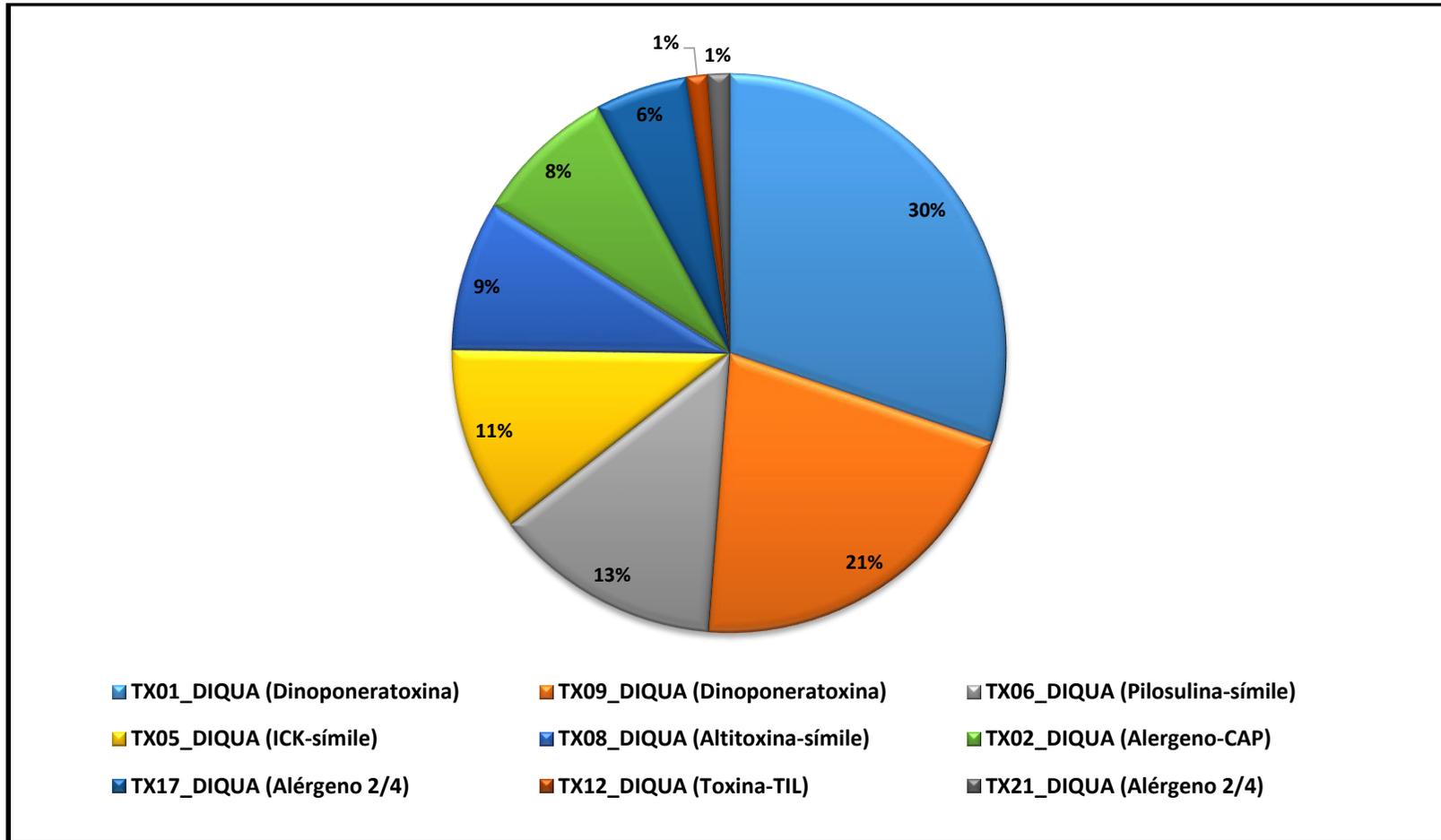
De modo interessante, através da comparação de sequências de aminoácidos e nucleotídeos de proteínas hipotéticas contra banco de dados específicos de artrópodes, os ESTs de *D. quadriceps* revelam certo grau de homologia com polipeptídeos de escorpiões (*Androctonus amoreuxi*) e outras formigas, como *Harpegnathos saltator*, *Solenopsis invicta* e *Camponotus floridanus*, refletindo sua relação filogenética. Os transcritos que não apresentaram homologia nos bancos de dados avaliados tiveram seus ORFs (Open Reading Frame) definidos e podem representar diferentes proteínas com novas possibilidades de funções biológicas. As sequências de ESTs de *D. quadriceps* estão sendo depositadas em banco de dados do GenBank.

**Gráfico 1:** Percentual total de ESTs da glândula de veneno de *D. quadriceps* classificados como precursores de proteínas envolvidas em processo celular (laranja), sem similares significantes (cinza), toxinas (azul) e proteínas hipotéticas (amarelo).



Os precursores de toxinas evidenciados foram classificados conforme apresentado no Gráfico 2. Os transcritos mais abundantes foram agrupados em dois novos grupos de precursores de dinoponeratoxinas (TX01\_DIQUA e TX09\_DIQUA). Juntos representam aproximadamente 50% do total de transcritos de toxinas. O segundo e o terceiro grupo de toxinas mais abundantes (TX06\_DIQUA e TX05\_DIQUA) foram classificados como pilosulina-símile e ICK-símile. A toxina TX08\_DIQUA, classificada como altitoxina, uma neurotoxina, também apresenta-se como um transcrito abundante, todavia esse sequenciamento apresentou-se incompleto, não sendo possível identificar o peptídeo sinal, devendo, portanto, ser confirmado. Os precursores das toxinas TX02\_DIQUA, TX17\_DIQUA e TX21\_DIQUA representam substâncias alergênicas, compreendendo cerca de 15% da toxinas. A TX12\_DIQUA, menos abundante, porém não menos importante, apresentam um domínio TIL (*Trypsin inhibitor-like*).

**Gráfico 2:** Proporção de ESTs relativas a possíveis toxinas expressas na glândula de veneno de *D. quadriceps*.



**Tabela 1:** Anotação de ESTs agrupados em contigs obtidos a partir da biblioteca de cDNA da glândula de veneno de *D. quadriceps*.

Contig	Comprimento (bp)	Definição proteica	Espécies com maior percentual de identidade (%)	E-value	Acesso Uniprot /Gi
Contig1	426	Dinoponeratoxin Da-2501	<i>Dinoponera australis</i> [formiga]	$7.4 \times 10^{-02}$	294863158
Contig2	919	Allergen Sol r III	<i>Solenopsis Invicta</i> [formiga]	$2 \times 10^{-80}$	6136163
Contig3	351	CCP	<i>Medicago truncatula</i> [legume]	4.4	XP_003606771
Contig4	767	Proteína hipotética	<i>Harpegnathos saltator</i> [formiga]	$5 \times 10^{-58}$	307212163
Contig5	485	ICK-like	<i>Androctonus amoreuxi</i> [escorpião]	$1.4 \times 10^{-03}$	PS01138
Contig6	430	Pilosulin	<i>Myrmecia banski</i> [formiga]	$6 \times 10^{-3}$	Q68Y23
Contig7	395	Sem similar			
Contig8	184	Altitoxin (?)	<i>Parabuthus transvaalicus</i> [escorpião]	1.8	PC1B5
Contig9	578	Dinoponeratoxin Da-3177	<i>Dinoponera australis</i> [formiga]	$4 \times 10^{-07}$	294863162
Contig10	576	Daughter sevenles protein	<i>Drosophila melanogaster</i> [mosca]	1.7	Q9VZZ9
Contig11	807	Cytochrome c oxidase	<i>Camponotus vafer</i> [formiga]	$2 \times 10^{-53}$	359828649
Contig12	585	Chymotrypsin inhibitor	<i>Acromyrmex echinator</i> [formiga]	$1 \times 10^{-19}$	EG167859
Contig13	954	PREDICTED: Histon-lisine N-methyltransferase	<i>Hydra magnipapillata</i> [cnidario]	$5 \times 10^{-3}$	449685068
Contig14	357	Sem similar			
Contig15	388	Sem similar			
Contig16	804	ATP syntase	<i>Vollenhovia benzai</i> [formiga]	$2 \times 10^{-19}$	373432543
Contig17	640	Allergen 2/4	<i>Solenopsis saevissima</i> [formiga]	$1 \times 10^{-4}$	A5X2H7

Contig18	436	Sem similar			
Contig19	424	Sem similar			
Contig20	717	NADH dehydrogenase	<i>Solenopsis geminata</i> [formiga]	$6 \times 10^{-03}$	312233093
Contig21	708	Allergen 2/4	<i>Solenopsis invicta</i> [formiga]	$7 \times 10^{-5}$	P35777

(?) Sequenciamento incompleto

**Tabela 2:** Anotação de ESTs individuais (singletos) obtidos a partir da biblioteca de cDNA da glândula de veneno de *D. quadriceps*.

Singlete	Comprimento (bp)	Definição proteica	Espécies com maior percentual de identidade (%)	E-value	Acesso Uniprot /Gi
A1_DVB2	504	NADH dehydrogenase	<i>Camponotus chromaoides</i> [formiga]	$7 \times 10^{-10}$	429535911
B4_DVB2	774	Chaperonin-like	<i>Harpegnathos saltator</i> [formiga]	$2 \times 10^{-118}$	EFN85211
B11_DVB2	443	Odorant-binding protein	<i>Drosophila melanogaster</i> [mosca]	6.6	Q9VR94
C9_DVB2	652	Adenosine monophosphate-protein transferase FICD homolog	<i>Drosophila willistoni</i> [mosca]	1.4	B4MUQ2
G8_DVA2	498	60S ribosomal protein	<i>Camponotus floridanus</i> [formiga]	$5 \times 10^{-54}$	307181094
B10_DVB1	826	Small nuclear ribonucleoprotein	<i>Harpegnathos saltator</i> [formiga]	$5 \times 10^{-05}$	307200838
C8_DVC1	707	Cytochrome b	<i>Formica gagatoides</i> [formiga]	$3 \times 10^{-53}$	317411386
D1_DVB2	785	60S ribosomal protein	<i>Nasoina vitripennis</i> [formiga]	$6 \times 10^{-101}$	156544263
B10_DVB1.2	212	DCN1-like	<i>Drosophila melanogaster</i> [mosca]	2.5	Q9VUQ8
D10_DVA2	906	Proteína hipotética	<i>Solenopsis invicta</i> [formiga]	$1 \times 10^{-63}$	322786660

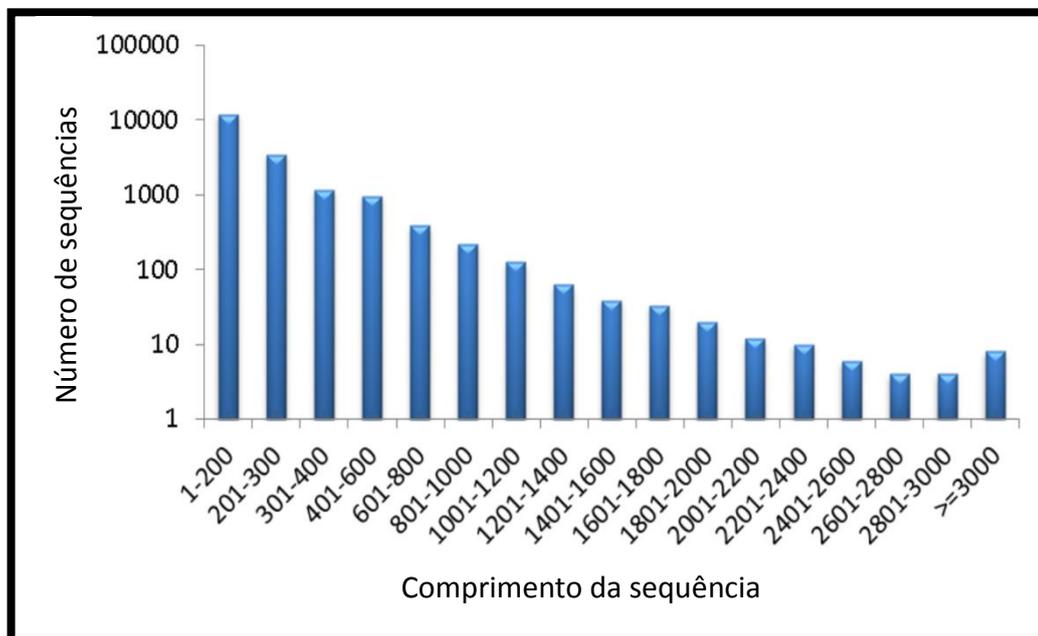
C3_DVB1.2	918	60S ribosomal protein L7a	<i>Drosophila melanogaster</i> [mosca]	$2 \times 10^{-70}$	P46223
B8_DVB1.2	261	Proteína hipotética	<i>Eremothecium cymbalariae</i> [fungo]	$1 \times 10^{-01}$	363755426
C1_DVB2	186	Sem similar			
B2_DVB2	638	Sem similar			
C12_DVB1.2	724	Sem similar	<i>Proformica mongolica</i> [formiga]	$2 \times 10^{-08}$	383281861
H6_DVB1.2	162	Sem similar			
E5_DVB2	578	Sem similar			
E7_DVA2	420	NADH dehydrogenase	<i>Nasonia longicornis</i> [formiga]	2.0	198114541
A3_DVB2	309	40S ribosomal protein	<i>Lysiphlebus testaceipes</i> [vespa]	0.024	Q56FH2
D8_DVB1.2	409	Sem similar			
F9_DVB1.2	585	Sem similar			
C8_DVB1.2	658	40S ribosomal protein	<i>Harpegnathos saltator</i> [formiga]	$2 \times 10^{-23}$	307214989
E6_DVC1	663	40S ribosomal protein	<i>Camponotus floridanus</i> [formiga]	$2 \times 10^{-69}$	307184786
F2_DVC1	490	40S ribosomal protein	<i>Camponotus floridanus</i> [formiga]	$3 \times 10^{-50}$	307180729
A9_PM	840	Iron-sulfur cluster	<i>Acromyrmex echinator</i> [formiga]	$2 \times 10^{-91}$	332020272
H6_DVB1	714	60S acidic ribosomal protein	<i>Harpegnathos saltator</i> [formiga]	$2 \times 10^{-141}$	307199404
E6_DVB1	830	PDZ and LIM domain protein Zasp	<i>Drosophila melanogaster</i> [mosca]	$6 \times 10^{-83}$	A1ZA47
G2_DVC1	843	Cytochrome oxidase subunit I	<i>Dinoponera gigantea</i> [formiga]	$6 \times 10^{-02}$	87046831
H4_PM	821	Aubergine	<i>Drosophila melanogaster</i> [mosca]	0.14	O76922
B9_DVC2	452	Cytochrome B	<i>Dinoponera lucida</i> [formiga]	$5 \times 10^{-02}$	307696607
E7_DVC2	716	Heavy neurofilament NF-H	<i>Culex quinquefasciatus</i> [díptero]	$2 \times 10^{-04}$	170037708

F8_DVA2	755	Proteína hipotética	<i>Harpegnathos saltator</i> [formiga]	$5 \times 10^{-32}$	307196685
F11_DVC2	584	15-Hydroxyprostaglandin dehydrogenase	<i>Harpegnathos saltator</i> [formiga]	$2 \times 10^{-62}$	307206934
D11_DVA2	207	Fit family protein	<i>Drosophila melanogaster</i> [mosca]	1.4	Q9VRJ2
G1_DVA2	955	Yellow protein (Geleia real)	<i>Harpegnathos saltator</i> [formiga]	$5 \times 10^{-47}$	307210802
F10_DVA2	847	Proteína hipotética	<i>Harpegnathos saltator</i> [formiga]	$5 \times 10^{-05}$	307210524
A9_DVA2	849	Protein Krh2	<i>Drosophila melanogaster</i> [mosca]	3.9	Q9V447
B3_DVC2	575	Sem simile			

### 4.3 Sequenciamento NGS (*Next generation sequencing*) por Ion Torrent e agrupamento das sequências

Mediante a redundância dos transcritos evidenciados no sequenciamento por Sanger, decidiu-se realizar o sequenciamento NGS, utilizando tecnologia Ion Torrent, com objetivo de evidenciar a presença de transcritos mais raros. O sequenciamento foi realizado a partir da mesma amostra utilizada na construção da biblioteca de cDNA e obteve-se um total de 2.514.767 novas leituras como tamanho médio de 230pb. Os dados foram submetidos à análise de bioinformática utilizando o software Trinity. Inicialmente as leituras foram filtradas para a remoção das sequências de baixa qualidade, utilizando o programa Perl Script, o que resultou em um total de 18.546 transcritos agrupados, cujos tamanhos dos contigs variaram de 101-8.114pb e um tamanho médio de 232pb. A distribuição de tamanho dos transcritos agrupados é mostrada na Gráfico 3.

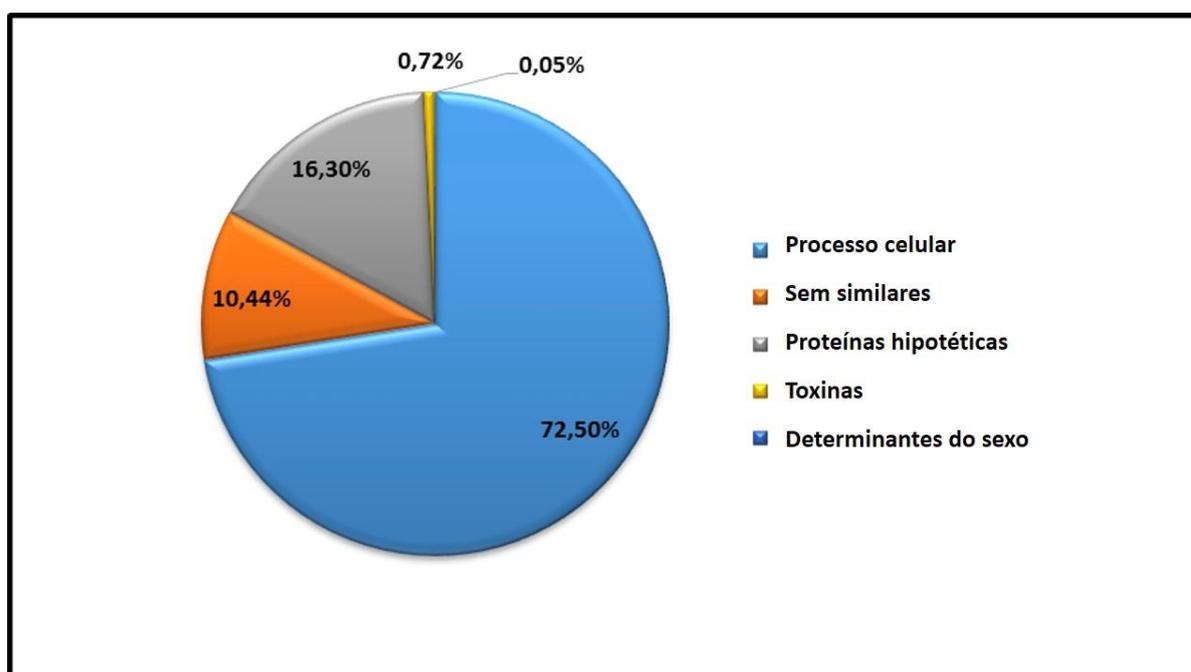
**Gráfico 3:** Distribuição de comprimento dos transcritos da glândula de veneno de *D. quadriceps*.



As sequências foram comparadas com outras em bancos de dados de proteínas não redundantes utilizando a ferramenta BlastX e 6.429 contigs com homologia gênica foram anotados (Gráfico 4). A maioria das sequências de transcritos

de *D. quadriceps* (3.807 contigs) apresentaram homologia com outras espécies de formigas (*H. saltator*), confirmando sua relação filogenética. Interessantemente, os ORFs de alguns contigs apresentaram elevada homologia com outros Himenópteros como *A. florae* e *Megachile rotundata*. Aproximadamente 98% dos ESTs anotados a partir da biblioteca de cDNA sequenciada por Sanger apresentaram homologia com os contigs obtidos a partir do sequenciamento Ion Torrent. Estes contigs foram categorizados em famílias de proteínas envolvidas, por exemplo, em processos celulares, polipeptídeos relacionados à peçonha e determinantes do sexo (Gráfico 4). Este transcriptoma está sendo depositado no DDBJ/EMBL/GenBank.

**Gráfico 4:** Proporção relativa dos transcritos da glândula de veneno de *D. quadriceps* identificados por NGS.



Os transcritos relacionados à toxinas e componentes da peçonha representam menos de 1% de todos os transcritos (18.524) e algumas foram correspondentes àquelas encontradas no sequenciamento por Sanger, como as representadas pelos contigs 01 e 02 (Dinoponeratoxina), contig 12 (alérgeno de veneno) (Tabela 3). Notavelmente, cerca de 15% dos transcritos evidenciados por Sanger e menos de 1% dos transcritos evidenciados por NGS correspondem à toxinas ou polipeptídeos associados à peçonha. Como o universo de transcritos sequenciados por NGS e a representação de proteínas envolvidas em processos celulares é deveras

maior, essa diferença percentual é compreensível. No universo de peptídeos associados à toxinas, as dinoponeratoxinas também se apresentaram bastante abundantes. A Tabela 4 apresenta uma visão geral dos polipeptídeos da peçonha de *D. quadriceps* e de outras espécies de formigas, evidenciando que compartilham semelhanças em sua composição, principalmente no que concerne ao conteúdo enzimático e de substâncias com potencial alergênico. Entretanto, dado que os componentes mais comuns no veneno apresentam elevado nível de similaridade, reações cruzadas podem ocorrer após acidentes com Himenópteros.

**Tabela 3:** Anotação de transcritos agrupados em contigs da glândula de veneno de *D. quadriceps*, obtidos a partir do sequenciamento Ion Torrent.

Contig	Comprimento (bp)	Definição proteica	Espécies com maior percentual de identidade (%)	E-value	Acesso Gi
*Contig2	483	Dinoponeratoxin Da-3177	<i>Dinoponera australis</i> [formiga]	8x10 <sup>-07</sup>	294863162
*Contig1	464	Dinoponeratoxin Da-2501	<i>Dinoponera australis</i> [formiga]	3x10 <sup>-06</sup>	294863158
*Contig12	1051	Venom allergen III, Sol i 3; Flags: Precursor	<i>Solenopsis invicta</i> [formiga]	2x10 <sup>-72</sup>	14424466
*Contig27	1656	Venom allergen I, Sol i 1; Flags: Precursor	<i>Solenopsis invicta</i> [formiga]	1x10 <sup>-44</sup>	74827072
Contig141	3365	Protein lethal (2) essential for life Chaperon heat shock protein like	<i>Acromyrmex echinator</i> [formiga]	2x10 <sup>-33</sup>	332022897
Contig351	3845	Protein lethal (2) essential for life Chaperon heat shock protein like	<i>Harpegnathos saltator</i> [formiga]	1x10 <sup>-37</sup>	307197364
Contig455	1497	Probable multidrug resistance- associated protein lethal (2)03659 ABC transporter	<i>Harpegnathos saltator</i> [formiga]	4x10 <sup>-169</sup>	307196484
Contig857	966	Probable multidrug resistance- associated protein lethal (2)03659 ABC transporter	<i>Harpegnathos saltator</i> [formiga]	3x10 <sup>-131</sup>	307196484
Contig1700	826	Probable multidrug resistance- associated protein lethal (2)03659	<i>Harpegnathos saltator</i> [formiga]	3x10 <sup>-36</sup>	307196484

ABC transporter					
Contig5586	139	Lethal(2)neighbour of tid protein	<i>Harpegnathos saltator</i> [formiga]	$5 \times 10^{-19}$	307207923
Contig8914	165	Lethal(2)neighbour of tid protein	<i>Acromyrmex echinator</i> [formiga]	$9 \times 10^{-19}$	332018132
Contig11936	176	Lethal(2)neighbour of tid protein	<i>Harpegnathos saltator</i> [formiga]	$6 \times 10^{-23}$	307207923
Contig12935	189	Lethal malignant brain tumor like 3 protein	<i>Harpegnathos saltator</i> [formiga]	$2 \times 10^{-09}$	307202837
Contig17436	120	Pre-mRNA-splicing regulator female-lethal	<i>Harpegnathos saltator</i> [formiga]	$2 \times 10^{-11}$	307211429
Contig17643	159	Probable multidrug resistance- associated protein letha (2)03659	<i>Harpegnathos saltator</i> [formiga]	$6 \times 10^{-10}$	307207958
ABC transporter					
Contig18270	128	Probable multidrug resistance- associated protein lethal (2)03659	<i>Harpegnathos saltator</i> [formiga]	$4 \times 10^{-14}$	307196485
ABC transporter					
Contig187	1410	PREDICTED: phospholipase A2 isoenzyme PA3A/PA3B/PA5	<i>Apis mellifera</i> [abelha]	$9 \times 10^{-109}$	328778177
Contig662	1603	Putative phospholipase B-like	<i>Harpegnathos saltator</i> [formiga]	$8 \times 10^{-26}$	307214880
Contig4464	699	Phospholipase A2 isoenzyme PA4	<i>Harpegnathos saltator</i> [formiga]	$2 \times 10^{-35}$	307198397
Contig5423	186	Phospholipase A2 isoenzyme PA3A/PA3B/PA5	<i>Acromyrmex echinator</i> [formiga]	$2 \times 10^{-18}$	332029748

Contig6851	234	Lysophospholipase-like protein 1	<i>Harpegnathos saltator</i> [formiga]	$5 \times 10^{-17}$	307200032
Contig9856	272	Phospholipase A2 activating protein	<i>Harpegnathos saltator</i> [formiga]	$2 \times 10^{-22}$	307196404
Contig10011	233	Group X11A secretory phospholipase A2	<i>Camponotus floridanus</i> [formiga]	$1 \times 10^{-34}$	307187781
Contig12747	163	Phospholipase B1 membrane associated	<i>Acromyrmex echinator</i> [formiga]	$9 \times 10^{-06}$	332028550
Contig13100	134	Phospholipase B1 membrane associated	<i>Harpegnathos saltator</i> [formiga]	$3 \times 10^{-11}$	307193256
Contig17494	145	Phospholipase D3	<i>Camponotus floridanus</i> [formiga]	$6 \times 10^{-11}$	307176418
Contig18196	185	Phospholipase DDHD2	<i>Harpegnathos saltator</i> [formiga]	$8 \times 10^{-28}$	307211672
Contig1144	495	PREDICTED: U8-agatoxin-Ao1a-like isoform 1	<i>Bombus impatiens</i> [abelha]	$6 \times 10^{-39}$	350400650
Contig326	451	PREDICTED: venom carboxylesterase-6-like	<i>Bombus terrestris</i> [abelha]	$2 \times 10^{-31}$	340722695
Contig391	921	PREDICTED: venom carboxylesterase-6-like	<i>Apis florea</i> [abelha]	$3 \times 10^{-102}$	380027409
Contig584	933	PREDICTED: venom carboxylesterase-6 isoform 1	<i>Nasonia vitripennis</i> [abelha]	$1 \times 10^{-103}$	345497204
Contig1381	150	PREDICTED: venom carboxylesterase-6-like	<i>Apis mellifera</i> [abelha]	$8 \times 10^{-13}$	48097744
Contig1632	149	PREDICTED: venom carboxylesterase-6-like	<i>Megachile rotundata</i> [abelha]	$1 \times 10^{-16}$	383858940
Contig3130	165	PREDICTED: venom carboxylesterase-6-like	<i>Megachile rotundata</i> [abelha]	$2 \times 10^{-12}$	383858940

---

Contig3290	474	PREDICTED: venom carboxylesterase-6 isoform 1	<i>Nasonia vitripennis</i> [formiga]	$2 \times 10^{-44}$	345497204
------------	-----	---	--------------------------------------	---------------------	-----------

---

\* Representam contigs que se assemelham aos da biblioteca de cDNA.

**Tabela 4:** Visão geral das principais toxinas polipeptídicas de *D. quadriceps* e de outras formigas.

Subfamília	Gênero e espécie	Principais polipeptídeos da peçonha	Referência
<b>Ponerinae</b>	<i>Dinoponera quadriceps</i>	Dinoponeratoxina, alérgenos (Sol i 1/PLA <sub>1</sub> B-like, Sol i 3/Ag 5-like, CAP allergen 2/4), esterases (fosfolipases A and B, fosfatase ácida, carboxilesterases), ICK-símile, Pilosulina-símile, Inibidor de quimiotripsina-símile	Este trabalho
	<i>Pachycondyla sp.</i>	16 and 24 kDa (AlérgenoPac c 3/Sol i 3 s) Proteínas ligadores de IgE	HOFFMAN, 2010
	<i>Harpegnathos saltator</i>	Esterases	BONASIO <i>et al.</i> , 2010
<b>Myrmicinae</b>	<i>Solenopsis invicta</i>	Alérgenos (Sol i 1, 2, 3 and 4), Fosfolipases (incluindo Sol i 1/PLA <sub>1</sub> B)	HOFFMAN, 2010
	<i>Tetramorium bicarinatum</i>	Peptídeos pilosulina-símile (1,3, 5), Alérgenos (Sol i 3/Ag 5), precursor de citocinas de insetos uENF2, metaloproteases	BOUZID <i>et al.</i> ,2013
	<i>Acromyrmex echinator</i>	Carboxilesterase-3, Esterase FE4, Fosfolipase A1,	GenBank GI68854, GI68747
<b>Formicinae</b>	<i>Camponotus floridanus</i>	Carboxilesterase-3	BONASIO <i>et al.</i> , 2010
<b>Myrmeciinae</b>	<i>Myrmecia sp.</i>	Fosfolipase A <sub>2</sub> , Fosfolipase B, Hialuronidase, Fosfatase ácida e alcalina	MATUSZEK <i>et al.</i> , 1992

Como mencionado anteriormente, substâncias alergênicas são bastante abundantes em peçonhas de himenópteros, como os alérgenos 1 (Sol i 1), contig 27 do NGS, não identificado por Sanger; e 3 (Sol i 3/Ag5) (identificado por Sanger, contig 02, e por NGS, contig 12), para os quais o alinhamento com diferentes sequências de outras espécies de himenópteros apresentou significativa homologia (Figura 11). Os resíduos de aminoácidos idênticos e conservados de diferentes espécies de formigas e vespas estão destacados em preto e cinza, respectivamente. Sequências de proteínas Sol i 3 de outras espécies de himenópteros, especialmente abelhas, foram submetidas à análise filogenética e, como esperado, observou-se que as sequências de *D. quadriceps* se mostraram estreitamente relacionadas quando comparadas com *S. invicta* e diferentes espécies de vespas, por exemplo (Figura 12).

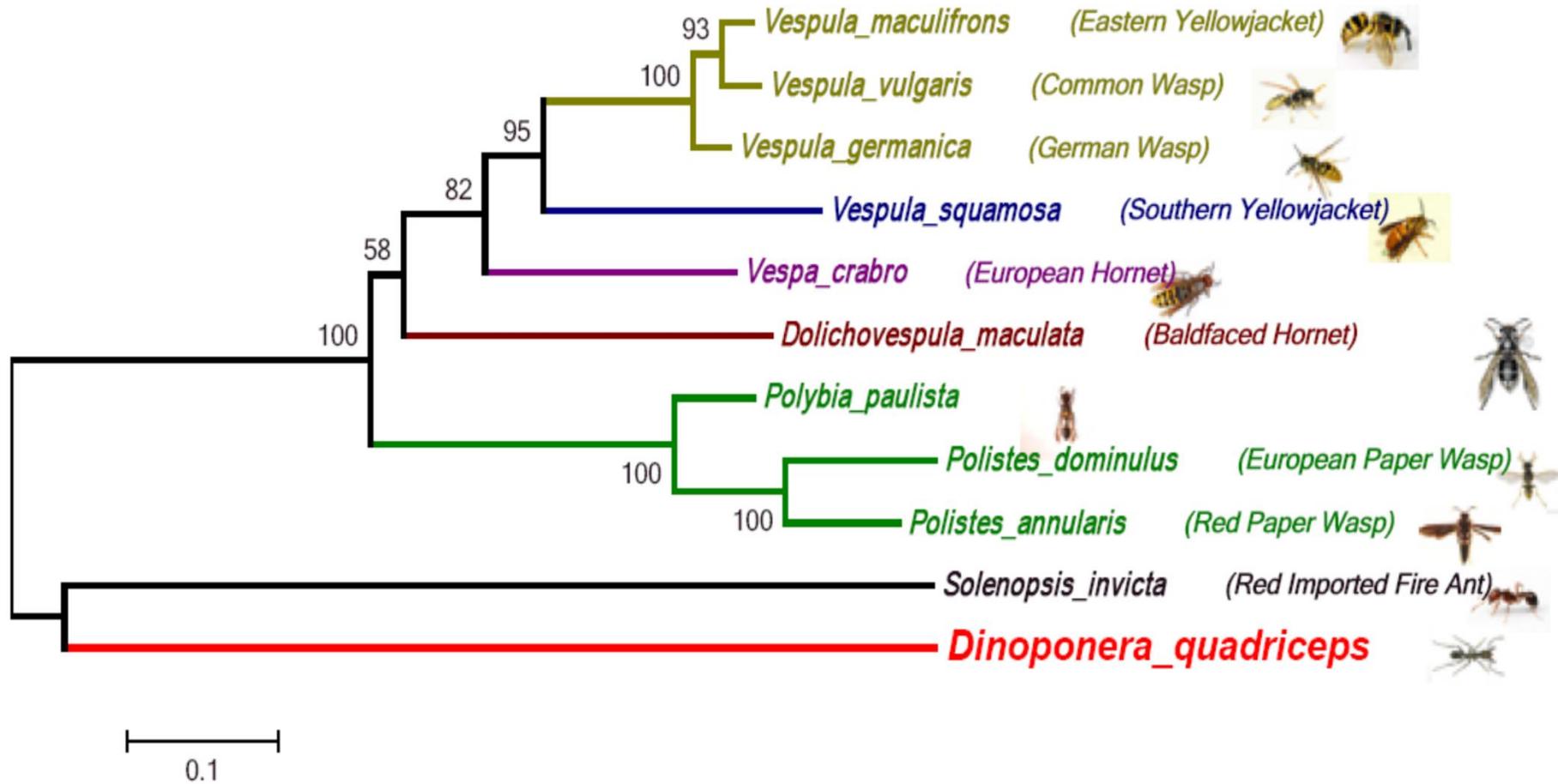
Identificamos dentre os transcritos da glândula de veneno de *D. quadriceps* homólogos aos quatro mais bem conservados alérgenos de veneno encontrados em formigas do gênero *Solenopsis* (Subfamília Myrmicinae): Sol i 1, Sol i 2, Sol i 3 e Sol i 4. A toxina alergênica do grupo Sol i 1 (Contig 27, NGS), está relacionada ao grupo das fosfolipases A<sub>1</sub> (PLA<sub>1</sub>). Estas proteínas podem ser classificadas em dois grupos, baseado em sua estrutura e localização celular. Um grupo compreende enzimas intracelulares e incluem PLA<sub>1</sub> seletivas para ácido fosfatídico e o outro enzimas extracelulares seletivas por fosfatidilserina (RICHMOND; SMITH, 2011). Membros deste grupo de enzimas tem sido encontrados em peçonha de himenópteros, principalmente nas de vespas. Sua estrutura é conservada em uma variedade de organismos e em humanos, e são capazes de induzir agregação plaquetária, reações alérgicas e hemólise (SONODA *et al.*, 2002; AOKI *et al.*, 2002; SEISMANN *et al.*, 2010). As demais toxinas alergênicas serão discutidas adiante neste trabalho.

**Figura 11:** Alinhamento da sequência de aminoácidos, evidenciada através da tradução de sequência nucleotídica obtida por Ion Torrent, do precursor alérgeno de veneno I (Sol i 1) da glândula de veneno de *D. quadriceps* com múltiplas sequências de aminoácidos dos precursores do alérgeno de veneno I de outros himenópteros.

PA12_DOLMA	-----	
PA1_VESGE	-----	
DIQUA_NGS27	PSTSRIVKVKLDNIGRRFRPLSRYSLLCTMKFITGILVIFCVYLLSTGGDSKILPLKKEP	60
PA1_SOLIN	-----MRKFAAIFVVFVQCTHLYSLAQARAEP	28
PA12_DOLMA	-----GILPECKLVPEEISFVLSTRENR	23
PA1_VESGE	-----GPKCPFNDDTVSMIETRENR	21
DIQUA_NGS27	SKIFHLKSHVDNTVKKPLKVFHLKSHVENSVGFPLRMNKLTPNCFVGVKSMVMVLFKKNIP	120
PA1_SOLIN	-----DEGVVEYLKQSCVYGNSSYINVYLYNSRF	57
PA12_DOLMA	DGVYLTLOKLNKNGKMFKNSDLSSKVPFLIHGFISSATNKNYADMTRALLDKDDIMVIST	83
PA1_VESGE	NRDLYTLQTLQNHPEFKEKTIITRPVVFITHGFTSSASETNFINLSKALVDKDNVMVISI	80
DIQUA_NGS27	DGKYISLSDSLG----RDLDLT-KTIYFTGHGFTSNVNHLSNRLSRALVEKD-YTVFSL	174
PA1_SOLIN	QG----KNLGNQQSCQDINAS-LFVVFITHGFTSSAQVSTFKDLANAFVQKGG-HTAFIV	110
PA12_DOLMA	DWRDGCSENEFALLKFIGYPKAVENTRAVVKYIADFSKILIQKYKVLLENIRLIGHSLGA	143
PA1_VESGE	DWQTAECTNEAAGLKYLYPTAASNTRLVGOYIATITQKLVKQYKISMANIRLIGHSLGA	140
DIQUA_NGS27	DWSDAACTTGGLPLVKLLYPSGVQNTREIGNLMADYVMSLIDHG-ASLRMGFIGHSLGS	233
PA1_SOLIN	DWSEAACTDGLPGVQFAEYNAAASNTYDIGOLMAKYTVDMNKCKIPLNNIQYVGHSLGS	170
PA12_DOLMA	QIAGFAGKEFORFKLGKYPEIIGLDPAGPSFKKDCPERICETDAHYVQILHTSSNLGTE	203
PA1_VESGE	HVSGFAGKVKQELKLGKYEIIGLDPAGPSFSSNKCSDRLCETDAEYVQILHTSNHLGTE	200
DIQUA_NGS27	HVC-FASKKIYESGYGKVPLLFAADPAQPLFQLKCCPDRLCDTDGKLVITLHTS-QIGLG	291
PA1_SOLIN	HVCGFAAHVKKLINKTMPYLLALDPADESEFGSNKGERICKSDAKRIVVFKTSILGIGE	230
PA12_DOLMA	RTLGTVDYFINDGSENOFGTYIIGETCSHTRAVKYLTECIRRECCLIGVF-----	253
PA1_VESGE	RILGTVDYFMNNGKNOPGGRFFTEVCSHSRAVIYMAHCIKHECCCLIGVF-----	250
DIQUA_NGS27	YPIGGLDLYFN-GFVQPKCH--LDITCGHRSVLYLINMVEKKCSFPGIPGTYYKQILNPF	348
PA1_SOLIN	NIIEHLLIVEDGCKSOPAPS-WYDVPCHSSESIVWATGMVSGRCQHLAVE-----	279
PA12_DOLMA	-----QSKNPQ-----FVSKCTRNECVVVG-----	273
PA1_VESGE	-----KSKSSQ-----PISSCTKQECVAVG-----	270
DIQUA_NGS27	SKFPYPNSKTTFDCFVMDDSIENPRKSLQNLAGIYYMFVDEDTFCTRKNFNCRORQQQLR	408
PA1_SOLIN	-----WTAQQRINPIQWKFWRVFTSNIIPAYPTSDTTNIVV-----	314
PA12_DOLMA	-----LNAKEYPKKGSFYVPVEAKAPFCNNNGKII-	303
PA1_VESGE	-----LNAKKYPSRGSEYVPVESTAPFCNNKGGKII-	300
DIQUA_NGS27	NVNFMTSQSKWLPGLSLIEYRSLKYEYDSFGVQKLAHFMQLEVADDLNVLLIDNNNSMY	468
PA1_SOLIN	-----LNTNVFKNDNTFEGEYHAFPCARNLFFKCRQ	345
PA12_DOLMA	-----	
PA1_VESGE	-----	
DIQUA_NGS27	FQMLNVYQYLDIIIDAVKTNINYNQVNVCGKCGSNKRCFVIIIDAVACKKKKKKKNV	527
PA1_SOLIN	Q-----	346

PA12\_DOLMA: Fosfolipase 1 2 da peçonha da vespa *Dolichovespula maculate* (P53357); PA1\_VESGE: Fosfolipase A1 da peçonha da vespa *Vespula germanica* (Q3ZU95); DIQUA\_NGS27: precursor do alérgeno de veneno I (Sol i I) da peçonha de *D. quadriceps*; PA1\_SOLIN: Fosfolipase A1 da peçonha de *S. invicta* (Q68KK0). Os resíduos de cisteína conservados estão destacados por uma caixa em negrito.

**Figura 12:** Comparação entre os polipeptídeos Sol i 3/Alérgeno Antígeno 5 de *D. quadriceps* com homólogos de diferentes himenópteros.



Árvore filogenética baseada em análise neighbor-join de alinhamento do Alérgeno 5/Sol i 3 e sua relação ortológica com diferentes himenópteros. A escala 0,1 indica substituições por sítio.

A família das fosfolipases é um grupo importante de enzimas, em sua maioria com papel intracelular. Porém é amplamente descrita em peçonha de serpentes e outros organismos como toxinas secretadas. Este grupo de toxinas também é descrito em veneno himenópteros. Em nossa análise transcriptômica evidenciamos precursores de diferentes isoformas de fosfolipases como fosfolipase A<sub>1</sub> (discutida anteriormente), fosfolipase A<sub>2</sub> (contigs 187, 4464, 5423, 10011) fosfolipase B<sub>1</sub> (contigs 12747 e 1310) e fosfolipase D (contigs 17494 e 18196) (Tabela 3). As fosfolipases A<sub>1</sub> e A<sub>2</sub> hidrolisam pontes fosfodiéster de fosfolipídeos nas posições *sn-1* e *sn-2*, respectivamente. As fosfolipases A<sub>2</sub>, subdivididas em mais de doze subtipos, é a mais abundante fosfolipase expressa na peçonha de himenópteros. Abelhas e vespas apresentam elevado nível de atividade fosfolipase em suas peçonhas, e são descritas como principal classe de enzimas responsáveis pelas reações alérgicas e dano tissular (KING; SPANGFORT, 2000; dos SANTOS PINTO *et al.*, 2012; HUANG *et al.*, 2011). Schmidt e Blum (1978) primeiramente descreveram a presença de fosfolipase A<sub>2</sub> e B na peçonha da formiga *Pogonomyrmex badius*. Desde então, a atividade fosfolipase vem sendo descrita na peçonha de várias outras espécies de formigas como *Neoponera apicalis*, *Paraponera clavata*, *Pachycondyla cressiondes*, *Pogonomyrmex occidentalis* e *Dinoponera grandis* (ZALAT *et al.*, 2003). As fosfolipases B, as quais apresentam atividade fosfolipase-A/lisofosfolipase (CD01824) estão envolvidas na conversão da fosfatidilcolina em ácidos graxos e glicerofosfocolina, provavelmente auxiliando na captação de lipídeos da dieta. Já as fosfolipases D, catalisam a hidrólise de pontes fosfodiéster de glicerofosfolipídeos gerando ácido fosfatídico, um sinalizador intracelular envolvido no transporte vesicular (LISCOVITCH *et al.*, 2000; MCDEMOTT *et al.*, 2004). Dentre os transcritos de *D. quadriceps*, encontramos a primeira fosfolipase D descrita em formigas gigantes (esfingomielinase D, E.C. 3.1.4.41), uma toxina dermonecrótica encontrada na peçonha de aranhas, particularmente da aranha marrom *Loxocles gaucho*, responsável pelo dano necrótico e a intensa resposta inflamatória durante o loxoscelismo (CHAIM *et al.*, 2011; MAGALHÃES *et al.*, 2013). A expressão de diferentes isoformas e subgrupos de fosfolipases assinala a homeostase lipídica e a sinalização mediada por compostos lipídicos como alvo para o dano tissular à presa ou vítima.

Transcritos para carboxilesterases também foram identificados. São enzimas que hidrolisam ésteres carboxílicos em seus correspondentes ácido e álcool (GO0004091). Enzimas desta classe são consideradas moléculas protetoras, uma vez que atuam na detoxificação celular, inativando substâncias tóxicas e carcinógenos. Como descrito por Hatfield e Potter (2011), diversos fármacos, pesticidas e produtos veterinários contém estruturas éster suscetíveis à conversão catalítica por estas enzimas. Elas têm sido encontradas em peçonhas de diversos himenópteros, como *Apis mellifera* (CUI *et al.*, 2011), *N. vitripennis* (WERREN *et al.*, 2010), *H. saltator*, *C. floridanus* (BONASIO *et al.*, 2010) e *A. echinator* (GI332028825). Esta classe de componentes pode estar envolvida na detoxificação de xenobióticos, conferindo resistência à inseticidas (WERREN *et al.*, 2010; BONASIO *et al.*, 2010). Além do papel fisiológico, as carboxilesterases de *A. mellifera* têm sido atribuídas à indução de reações alérgicas em humanos (BLANK *et al.*, 2008).

Contigs que não apresentaram nenhuma correlação com as sequências conhecidas no banco de dados de proteínas não redundantes representam 65% do total de transcritos. Este elevado percentual pode ser explicado pela elevada stringência utilizada para o Blast. Adicionalmente, o número de sequências anotadas relativas à peçonha de outras espécies de formigas é limitado. Deste modo, esse universo de transcritos sem homologia podem representar mRNA que carregam informações de novos genes. Para identificar e caracterizar esses contigs no banco de dados de proteínas não redundantes, os mesmos foram submetidos à avaliação utilizando BlastX no banco de dados UniProtKB. Esta análise resultou em 34 novas classificações. Para os demais contigs, realizou-se BlastN contra os ESTs obtidos a partir do sequenciamento por Sanger e vários contigs foram mapeados através dos ESTs (Tabela 3). Em adição, a predição dos ORFs foi realizada utilizando a ferramenta OrfPredictor (MIN *et al.*, 2005).

#### **4.4 Anotação do agrupamento dos precursores de toxinas da glândula de veneno de *D. quadriceps***

Os dados dos sequenciamentos foram utilizados para classificar funcionalmente os transcritos agrupados, baseado na homologia das sequências com outras depositadas em bancos de dados como UniProt e GenBank. Os resultados

permitiram a classificação funcional conforme descrito nos itens 4.2 e 4.3. Assim, de acordo com as espécies e as informações das anotações de proteínas dos bancos de dados, os transcritos relacionados à toxinas foram agrupados em algumas categorias: (1) Toxinas alergênicas, (2) ICK-símile, (3) Dinoponeratoxinas, (4) Toxinas com domínio TIL. Apresentaremos em detalhes, nesta seção, as toxinas mais bem caracterizadas.

#### **4.4.1 Toxinas alergênicas**

Proteínas alergênicas são abundantemente encontradas em venenos de himenópteros e estão entre os alérgenos mais potentes, com doses na ordem de nanogramas causando sensibilização e anafilaxia (HOFFMAN, 1993). As toxinas classificadas como alergênicas representam o segundo grupo mais abundante de toxinas presentes no veneno de *D. quadriceps*. Estas foram subclassificadas em toxinas alergênicas semelhantes à Sol i 1/PLA<sub>1</sub> (Contig 27, NGS, discutida no item 4.3), Sol i 3/Ag5/CAP domain (Contig TX02\_DIQUA do sequenciamento por Sanger e contig12 do NGS), pertencente à família da toxina alergênica antígeno 5, encontrada em vespas; e toxinas alergênicas da família 2/4, caracterizadas no veneno da formiga *Solenopsis invicta* (Contigs TX17\_DIQUA e TX21\_DIQUA, do sequenciamento por Sanger). Outro grupo de toxinas alergênicas é representado pelo contig TX06\_DIQUA, sendo caracterizado como Pilosulina-símile, primeiramente identificadas na peçonha de formigas do gênero *Myrmecia*. Estas toxinas podem estar envolvidas na produção de imunoglobulina E específica e degranulação de mastócitos durante a resposta alérgica em seres humanos.

##### **4.4.1.1 Toxina alergênica semelhante à Sol i 3/Ag5/CAP domain**

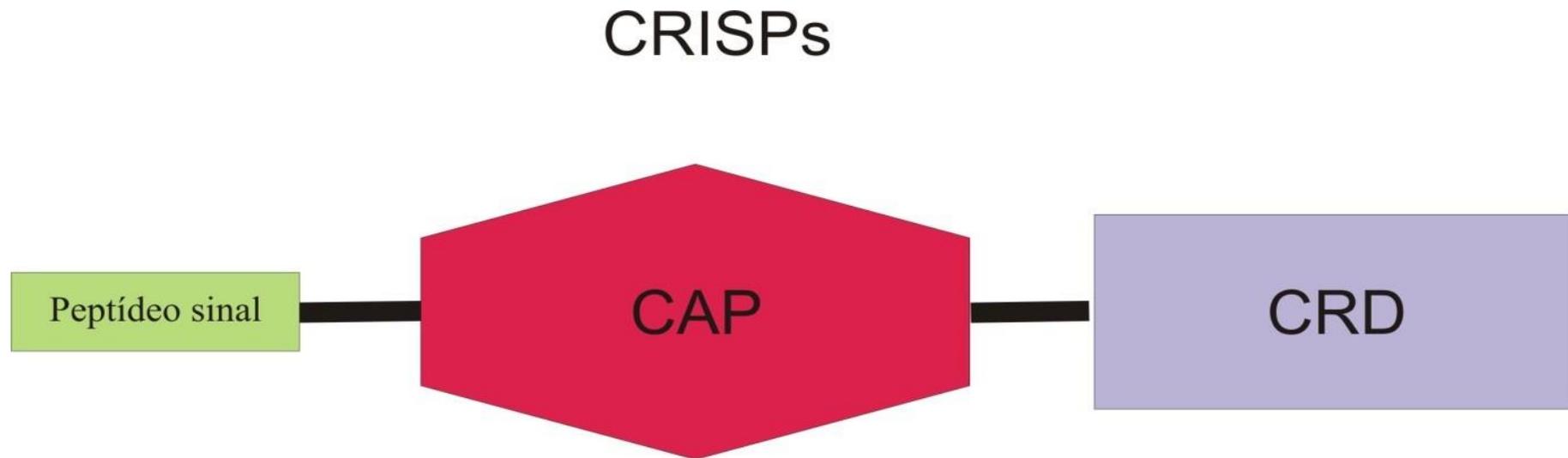
Os contigs TX02\_DIQUA (Sanger) e Contig 12 (NGS) codificam o grupo de toxinas alergênicas que apresentam um domínio CAP na extremidade N-terminal e portanto pertence à superfamília CAP (**C**ysteine-rich secretory protein family - CRISPs, **I**nsect venom **A**llergen antigen 5, and **P**athogenesis-related 1 proteins). As CRISPs são compostas de um domínio CRD (Cystein-Rich Domain) C-terminal e um domínio CAP conservado N-terminal (Figura 13). Os alérgenos Sol i 3/Ag 5 e o

precursor codificado pelo contig TX02\_DIQUA contam apenas com o domínio CAP isolado N-terminal, não apresentando, portanto, a estrutura característica CRISP.

O cDNA de TX02\_DIQUA apresenta 921pb e é composto por uma região 5'UTR de 79pb, seguida de uma região codificante (80-775pb) e uma região 3'UTR que se estende da posição 776-921. O códon de inicialização encontra-se na posição 80 e está precedido pelo consenso Kozak (KOZAK, 1987). Esta organização resultou em um ORF de 693pb que se estende até um códon de terminação na posição 775. Submetendo-se a proteína precursora à ferramenta SignalP determinou-se o ponto de clivagem entre as posições 21-22 (Aminoácidos TIA-TN) do precursor (Figura 14). Sua toxina madura apresenta pI de 9.18 e massa de 23,6 KDa preditos utilizando a ferramenta ExPASy/Compute pI/Mw.

O consenso do contig TX02\_DIQUA traduzido gerou uma proteína precursora com homologia com proteínas alergênicas VA3\_SOLIN (P35778) e VA5\_VESVU (Q05110) de *Solenopsis invicta* e de *Vespula vulgaris*, respectivamente, o que pode ser avaliado através do alinhamento das sequências como mostrado na figura 15. Ainda por homologia com as proteínas de formiga e vespa o peptídeo sinal e a organização das pontes dissulfeto foram preditas. Os elementos da estrutura secundária ( $\alpha$ -hélice e folha- $\beta$ ), os segmentos que conferem assinatura aos subdomínios CAP, bem como os resíduos de aminoácidos (histidina, glutamina/ácido glutâmico) que compõem o sítio ativo puderam ser indicados (Figura 15). A superfamília CAP é caracterizada por uma sequência de subdomínios conservada, quais sejam CAP1, CAP2, CAP3 e CAP4 (Vide figura 15). Embora muitas proteínas desta superfamília apresente um domínio CAP isolado (Ag5 e Pr-1, encontradas em insetos e plantas, respectivamente), um grande número delas apresenta uma extensão a C-terminal, sugerindo uma arquitetura multidomínio, e portanto proteínas com funções distintas (GIBBS *et al.*, 2008).

**Figura 13:** Representação esquemática de proteínas da família CRISP (*Cysteine-rich secretory protein Family*)



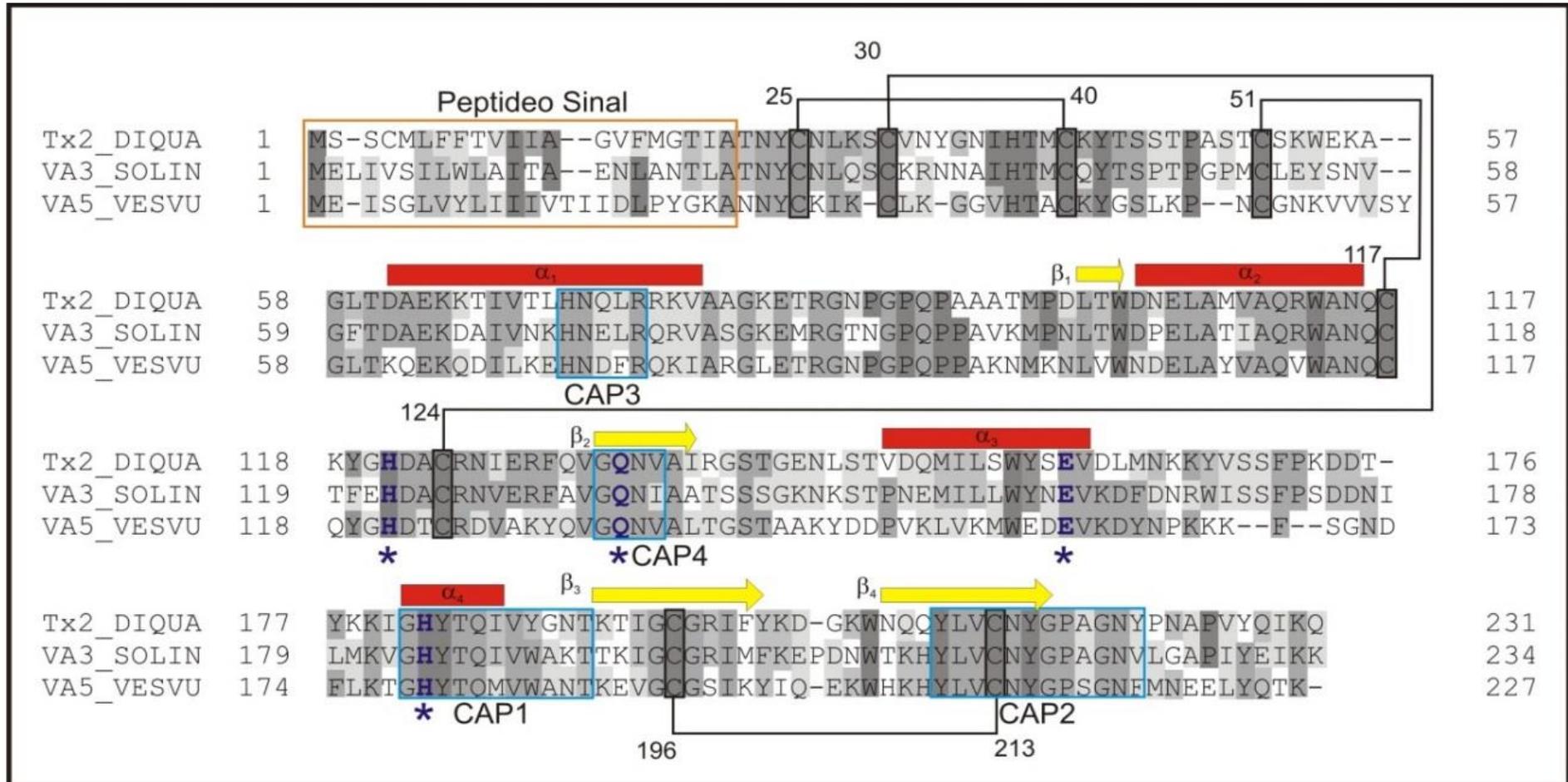
CAP: *Cysteine-rich secretory protein family - CRISPs, Insect venom Allergen antigen 5, and Pathogenesis-related 1 proteins*. CRD: *Cystein-Rich Domain*

Figura 14. Organização do cDNA do contig TX02\_DIQUA.

Contig TX02_DIQUA CAP-Alérgeno Dq-23660				
1	5'UTR			
	TAGGATAGGA	ATAGGCGAGA	AGTTACTGAC	AACAGTTT TAG AAAACGGAAA
51	AAATCCATGA	AATATATATC	ACGACGAAAA	<b>ATG</b> TCGAGTTG TATGTTGTTT
			M	S S C M L F
	Peptídeo sinal			
101	TTTACTGTAA	TCATAGCGGG	GGTGTTTATG	GGCACAATCG CGACAAATTA
	F T V I	I A G	V F M	G T I A T N Y
151	TTGCAATTTG	AAGTCATGCG	TAAACTATGG	AAATATTCAT ACGATGTGCA
	C N L	K S C V	N Y G	N I H T M C K
201	AGTATACTTC	ATCTACACCG	GCGTCAACAT	GTAGCAAATG GGAAAAAGCA
	Y T S	S T P	A S T C	S K W E K A
251	GGTTTAACCG	ATGCAGAAAA	GAAGACAATA	GTGACTCTAC ACAACCAGCT
	G L T D	A E K	K T I	V T L H N Q L
301	GAGGCGAAAA	GTTGCAGCGG	GTAAAGAAAC	GAGAGGAAAT CCGGGCCCGC
	R R K	V A A G	K E T	R G N P G P Q
351	AACCAGCGGC	AGCTACAATG	CCAGACTTGA	CTTGGGATAA TGAATTAGCA
	P A A	A T M	P D L T	W D N E L A
401	ATGGTTGCAC	AGAGATGGGC	TAATCAATGC	AAATACGGCC ATGATGCTTG
	M V A Q	R W A	N Q C	K Y G H D A C
	Toxina madura			
451	TAGAAATATA	GAAAGATTCC	AAGTTGGCCA	GAACGTAGCC ATAAGAGGCA
	R N I	E R F Q	V G Q	N V A I R G S
501	GCACTGGTGA	AAACCTTTTCG	ACTGTGGATC	AAATGATTCT TTCGTGGTAC
	T G E	N L S	T V D Q	M I L S W Y
551	AGTGAAGTGG	ATTTGATGAA	CAAAAAATAT	GTTTCCTCAT TTCCGAAAGA
	S E V D	L M N	K K Y	V S S F P K D
601	TGACACGTAT	AAGAAAATAG	GTCATTATAC	TCAAATAGTT TATGGTAATA
	D T Y	K K I G	H Y T	Q I V Y G N T
651	CGAAGACTAT	TGGTTGTGGA	CGAATATTTT	ACAAGGATGG CAAATGGAAC
	K T I	G C G	R I F Y	K D G K W N
701	CAACAATATT	TGGTTTGTA	CTACGGTCCA	GCTGGAAATT ATCCAAACGC
	Q Q Y L	V C N	Y G P	A G N Y P N A
751	ACCAGTATAC	CAAATTAAGC	AA <b>TAA</b> ATTTA	ACCATTAAAT GTATAAATGA
	P V Y	Q I K Q	*	Códon de terminação
801	TCCATAAGCA	ATGTGAAAAC	TATAACGAGA	TACAATAAGA GGGAAAACAT
	3'UTR			
851	ATTTCTATAT	CTCTCTTCTT	TATGAAATAA	TTCGAAACAT TCTGCGATAA
	Sinal de poliadenilação			
901	AA <b>TAAA</b> AGCA	ATAGCAATCA	TAAAAAAAAA	AAAAAAAAA

5'UTR (lilás), Peptídeo sinal (amarelo), Proteína madura (vermelho), 3'UTR (verde), sinal de poliadenilação (caixa amarela). O peptídeo sinal foi predito utilizando a ferramenta SignalP.

Figura 15: Alinhamento dos precursores dos alérgenos CAP (Ag5).

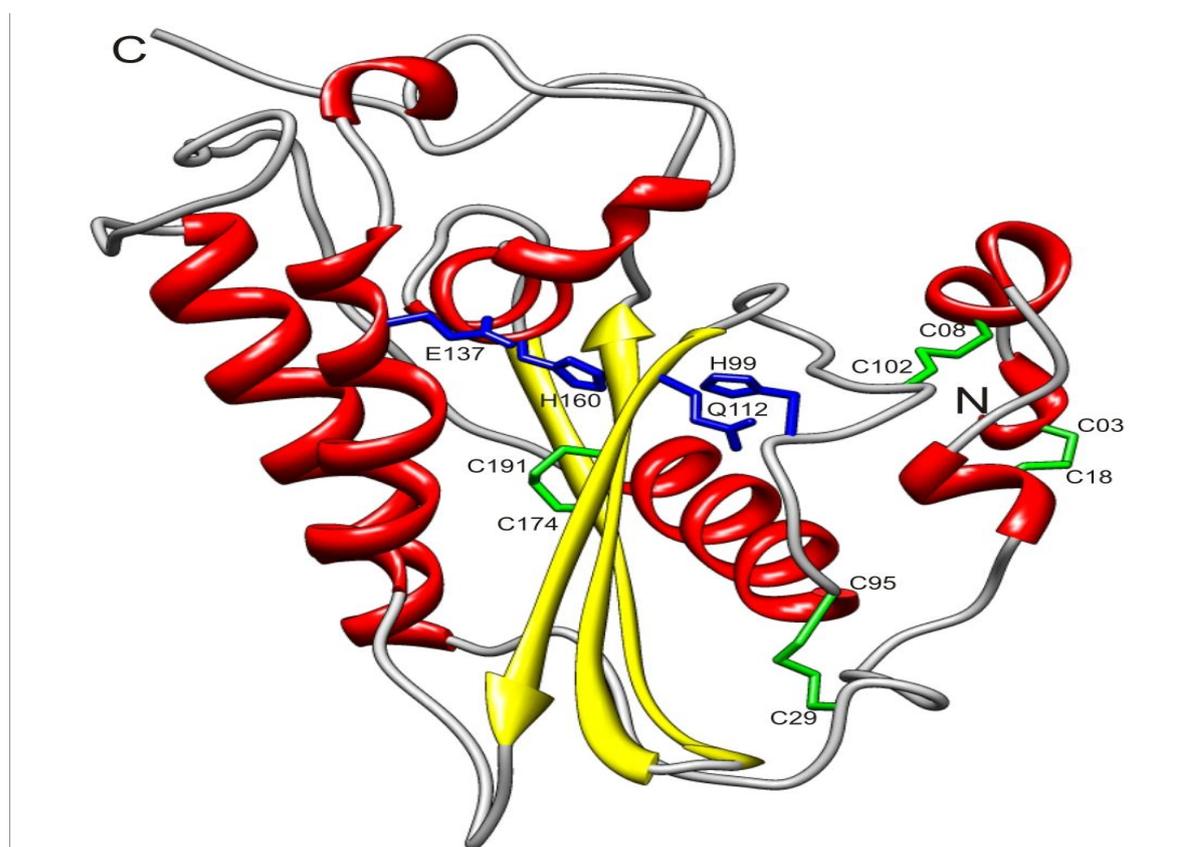


**TX02\_DIQUA:** Precursor de *D. quadriceps* (Dq CAP VA) alinhado com os precursores **VA3\_SOLIN** (UNIPROT P35778) de *S. invicta* e **VA5\_VESVU** (UNIPROT Q05110) de *V. vulgaris*. Peptídeo sinal predito por homologia (Caixa laranja). Pontes dissulfeto indicadas pelas linhas de conexão numeradas. Elementos da estrutura secundária conservados estão indicados em vermelho ( $\alpha$ -hélice) e amarelo (folha- $\beta$ ). Segmentos que caracterizam o domínio CAP (Caixa azul). \*resíduos de histidina e glutamina/ácido glutâmico que compõem o sítio ativo.

O modelo tridimensional (Figura 16) para o alérgeno CAP de *D. quadriceps* foi predito utilizando como molde os CAP alérgenos Sol i 3 de *Solenopsis invicta* (2VZN) e Ves5 de *Vespula vulgaris* (1QNX). A estrutura secundária do alérgeno CAP de *D. quadriceps* apresenta uma arquitetura em sanduiche consistindo de folhas beta antiparalelas envolvidas por alfa hélices. Estrutura semelhante à dos homólogos anteriormente mencionados (HENRIKSEN *et al.*, 2001; PADAVATTAN *et al.*, 2008).

As proteínas Ag5 são abundantes em venenos de vespas e estão associadas à resposta alérgica em humanos. Em formigas Sol i 3 de *Solenopsis invicta* e Pac 3 de *Pachycondyla sp.*, apresentam elevada homologia com Ag5 de vespas, sendo relatados como degranuladores de mastócitos e como antígenos reconhecidos por imunoglobulina E (HOFFMAN, 2010).

**Figura 16:** Modelo tridimensional de TX02\_DIQUA, Alérgeno CAP do veneno de *D. quadriceps*, utilizando como modelo VA Sol i 3 (2VZN) de *S. invicta* e VA Ves5 (1QNX) de *V. vulgaris*



Folha- $\beta$  (amarelo);  $\alpha$ -hélice (vermelho); Cisteínas (Verde); Sítio ativo (Azul), N-terminal (N); C-terminal (C). Os números indicam as posições das cisteínas na sequência do peptídeo maduro de *Dinoponera quadriceps*.

A conservação estrutural do Antígeno 5/Sol i 3 (TX02\_DIQUA) não é relacionada unicamente a outras espécies de formigas, mas também a espécies de vespas e abelhas. De fato, as proteínas denominadas CAP, compreendem uma superfamília de polipeptídeos encontrados em uma ampla variedade de organismos, incluindo procariotos. É interessante notar que polipeptídeos do trato reprodutivo de machos de mamíferos e em ductos secretórios de venenos, por exemplo *Stecrisp*, *PSTx*, *Triflin*, *Helothermine*; de diferentes espécies de serpentes e lagartos são proteínas com domínio CAP, e estão subclassificadas no grupo CRISP. Como relatado anteriormente, este grupo apresenta dois domínios: um domínio CAP N-terminal e um domínio C-terminal rico em resíduos de cisteínas (CRD). Do ponto de vista venômico, o domínio CRD apresentam ponto de dobra comum com inibidores de canais iônicos de anêmonas marinhas (toxinas Bgk- e Shk-símile), as quais estão envolvidas em bloqueio de canais de  $Ca^{+2}$  e  $K^{+}$ , bem como modulam receptores de rianodina. De modo interessante, estas atividades biológicas são observadas somente se estes polipeptídeos estão associados a um domínio CAP (GIBBS *et al.*, 2006; GIBBS *et al.*, 2008). Embora as proteínas que apresentam um domínio CAP isolado necessitem ser melhor caracterizadas funcionalmente, sabe-se que os homólogos à Sol i 3/Ag5 de venenos de himenópteros formam uma maior e distinta classe. Os homólogos a Sol i 3/Ag5, embora inicialmente caracterizados no veneno de formigas do gênero *Solenopsis* e de vespas distintas, foram adicionalmente identificados no intestino de *Drosophila*, na saliva de carrapatos, de flebotomíneos e mosquitos. Na secreção de insetos que se alimentam de sangue, as proteínas Ag5 fazem parte de um coquetel de proteínas salivares que podem se apresentar como inibidores do sistema imune do hospedeiro ou na inibição da formação do coágulo para alimentação prolongada (GIBBS *et al.*, 2008; KING; SPANGFORT, 2000). Assim a conservação estrutural destes alérgenos no veneno de *D. quadriceps* (Ponerinae), e outras espécies de formigas (*Pachycondyla sp.* e *Solenopsis sp.*) e vespas, é um aspecto interessante da venômica de Himenópteros, desde que os grupos de insetos têm uma longa e natural história de divergência de mais de 150 milhões de anos (WARD, 2006).

#### 4.4.1.2 Alérgenos de veneno 2/4 (PBP/GOPB)

O consenso do contig TX17\_DIQUA (640pb) apresenta uma região 5'UTR de 74pb e um quadro de leitura de tradução em proteína que se inicia na posição 75 (ORF) e se estende até a posição 468 (Códon de terminação), seguido de uma região 3'UTR que se estende da posição 468-640pb. Já o consenso do contig TX21\_DIQUA apresenta 708pb, uma região 5'UTR de 74pb, seguida de uma região codificante que se estende da posição 75-467pb e uma região 3'UTR entre as posições 468-708pb. Estes consensos traduzidos geraram proteínas precursoras de 132 (TX17\_DIQUA) e 131 (TX21\_DIQUA) aminoácidos, com peptídeo sinal de 19 aminoácidos e possível ponto de clivagem entre as posições 19-20 (Aminoácidos GYG-FD). Adicionalmente, os pontos isoelétricos e a massas preditos das toxinas maduras é de 9,8 e 12,6 KDa para TX17\_DIQUA e 9,18 e 13,2 KDa para TX21\_DIQUA (Figura 17 A e B).

Estes contigs apresentam relação com os alérgenos Sol i 2 e Sol i 4 da peçonha de *S. invicta*, e tratam-se de sequências de cDNA que codificam precursores pertencentes à mesma superfamília de proteínas de insetos chamada *Allergen 2/4 superfamily of Pheromone binding proteins-PBP/General-odorant binding proteins-GOBP* (InterPro IPR023316). Sol i 4 é encontrada nas rainhas e Sol i 2 é encontrado nas operárias.

Utilizando-se a ferramenta BlastP, ambos os transcritos apresentam alinhamentos de alta significância de similaridade com sequências hipotéticas e não caracterizadas do genoma de formigas da espécie *Harpegnathos saltator* (p.ex. UNIPROT E2BBA6). Entretanto, por se tratar de uma dedução do dado genômico, estas sequências parecem ser parte de um fragmento gênico que codifica para um segmento de 63 aminoácidos que se alinham com a região C-terminal dos precursores das proteínas codificadas pelos contigs TX17\_DIQUA e TX21\_DIQUA.

Os dois contigs compartilham 64% de identidade ou 83% de similaridade entre si, evidenciando que se trata de isoformas pertencentes à mesma família (Figura 18 A). Os peptídeos sinais preditos pela ferramenta SignalP são muito similares e o ponto de clivagem ocorre uma posição imediatamente após uma tirosina conservada em todas as sequências alinhadas, indicando que este grupo de proteínas sofre um processamento similar durante a secreção para o lúmen do epitélio da glândula de peçonha. Por outro lado, ambas as sequências também apresentaram alinhamentos

completos de alta significância de similaridade com alguns dos principais alérgenos presentes na peçonha das formigas *Solenopsis invicta*, denominados de alérgeno Sol i 2 e alérgeno Sol i 4, que são a causa mais frequente das reações de hipersensibilidade a Himenópteros no sul dos EUA (HOFFMAN, 1993). Esses alérgenos constituem uma família denominada “*Ant Venom Alergens 2/4*” (InterPro IPR020181) e se caracterizam por apresentarem 6 cisteínas conservadas que formam 3 pontes dissulfeto e se associarem em homodímeros ligados por uma ponte dissulfeto intercadeia via uma cisteína livre (BORER *et al.*, 2012) (Figura 18 A). Como mencionado anteriormente, homólogos para os quatro subtipos de alérgenos mais conservados de *S. invicta* foram encontrados nos transcritos de *D. quadriceps*. Os alérgenos 2 e 4 são estreitamente relacionados e, para *S. invicta* apresentam aproximadamente 35% de identidade.

O contig TX17\_DIQUA apresenta 21% de identidade com *Venom allergen 2* de *Solenopsis invicta* (Uniprot P35775) e 22% de identidade com *Venom allergen 2* da formiga *Solenopsis saevissima* (Uniprot A5X2H7). Já o contig TX21\_DIQUA apresenta 27% de identidade com *Queen venom protein Sol i 4* de *Solenopsis invicta* (Uniprot B1WA78). Podemos observar no alinhamento da figura 18A que os precursores codificados pelos contigs TX17\_DIQUA e TX21\_DIQUA apresentam quatro de seis cisteínas conservadas nas sequências de *S. invicta* envolvidas na formação de pontes dissulfeto intracadeia. As cisteínas envolvidas na formação da segunda ponte estão ausentes. Assim como também está ausente a cisteína livre que é conservada nas sequências da família “*Ant venom allergen 2/4*” e seria responsável pela formação do dímero. Podemos concluir que os precursores codificados pelos contigs TX17\_DIQUA e TX21\_DIQUA geram monômeros livres secretados na peçonha de *D. quadriceps* que não se associam covalentemente por meio de pontes dissulfeto intercadeia e que apenas duas pontes SS se conservam. É importante mencionar que o precursor de *Queen venom protein Sol i 4* de *Solenopsis invicta* (Uniprot B1WA78) também não forma dímeros ligados por ponte SS uma vez que não existem cisteínas livres na molécula. Sendo que este precursor, por sua vez é mais similar à sequência codificada pelo contig TX21\_DIQUA (Figura 17 B).

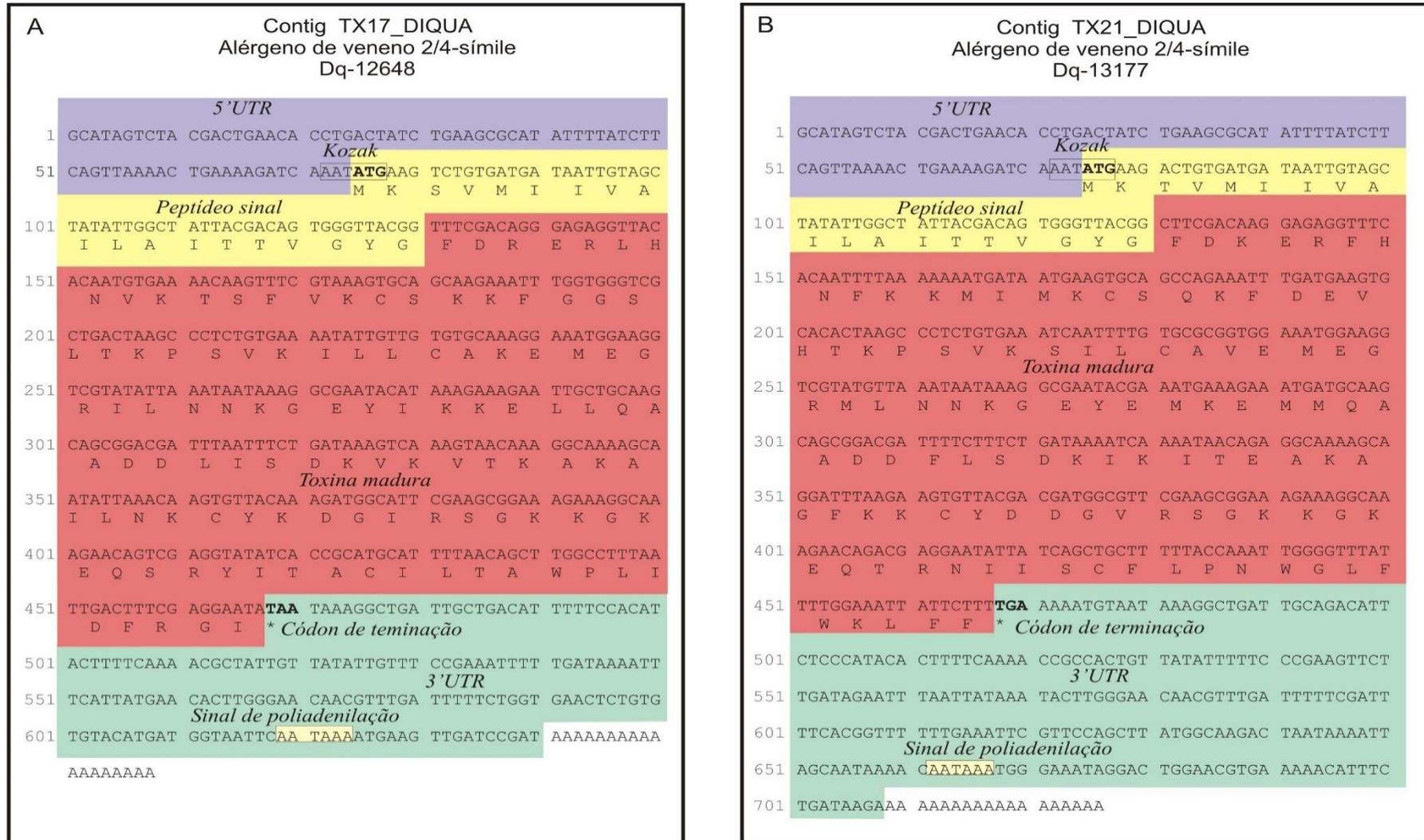
Por outro lado, ambos os contigs apresentam alinhamento de similaridade altamente significante com proteínas pertencentes à família das proteínas de inseto ligadoras de feromônios e/ou proteínas ligadoras de odorantes em geral.

Curiosamente os receptores olfativos dos animais terrestres existem em um ambiente particularmente aquoso apesar de detectarem moléculas de odor que são essencialmente hidrofóbicas. A solubilidade desses odorantes hidrofóbicos em meio aquoso é grandemente facilitada através da sua ligação a proteínas de ligação a odorantes que existem no fluido extracelular que rodeia os neurônios olfativos (VOGT *et al.*, 1991). Esta família é composta por proteínas ligantes de feromônios de insetos (PBP), que em geral são machos específicas e associadas a neurônios sensíveis a feromônios. Mas também fazem parte desta família diversas proteínas ligadoras de moléculas odoríferas de maneira geral (GOBP). Este domínio pode estar associado a diversas arquiteturas constituindo proteínas monoméricas ou multi-domínio. Um exemplo de proteína desta classe é a *Odorant binding protein* da borboleta monarca *Danaus plexippus* (Uniprot G6DFG7), este domínio único de 145 aminoácidos apresenta 21% de identidade com o contig TX17\_DIQUA e 28% de identidade com o contig TX21\_DIQUA. Caracteristicamente, a segunda ponte dissulfeto dos alérgenos da família 2/4 de *S. invicta* também está ausente na proteína da borboleta monarca. Enquanto que a primeira e a terceira ponte dissulfeto são formadas por cisteínas em posições conservadas em todas as sequências da família. Da mesma forma, também está ausente a cisteína livre que participa da formação da ponte intercadeia do dímero dos alérgenos Sol i 2 de *S. invicta*. Sob este ponto de vista, os contigs TX17\_DIQUA e TX21\_DIQUA são mais similares aos precursores de proteínas da família PBP/GOPB.

A Figura 18 B, apresenta um dendrograma da similaridade dos precursores alinhados, criado através do programa MEGA 5 (TAMURA *et al.*, 2011), permitindo agrupar os Contigs TX17\_DIQUA e TX21\_DIQUA com o precursor da *Odorant binding protein* da borboleta monarca. Enquanto que os alérgenos Sol i 2 das formigas *S. invicta* e *S. saevissima* se agrupam com o alérgeno 4 de *S. invicta* (*Queen venom protein* Sol i 4). Este, por sua vez, se agrupa com a PBP/GOPB da borboleta monarca indicando que o ancestral comum dos dois grupos é, muito provavelmente, um precursor monomérico do tipo PBP/GOPB com apenas duas pontes dissulfeto. Portanto, a classe de alérgenos 2/4 da peçonha de *Solenopsis invicta* é uma subfamília do grupo PBP/GOPB. Deste modo, é preferível classificar os precursores codificados pelos contigs TX17\_DIQUA e TX21\_DIQUA como toxinas da peçonha de *Dinoponera quadriceps* similares à família PBP/GOPB (*Dinoponeratoxin vPBP/GOPB-*

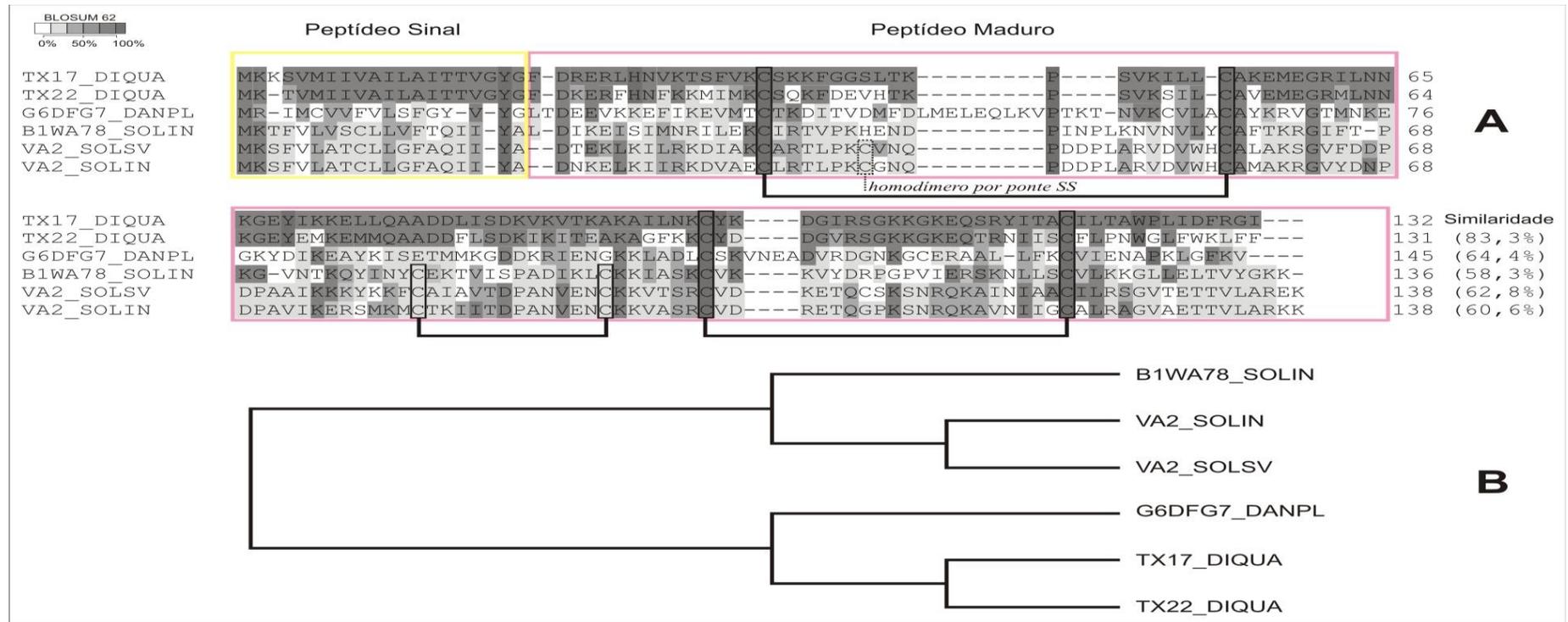
*like: venom like pheromone binding proteins/general-odorant binding proteins*). De maneira análoga, pode-se imaginar que estas sequências de *Dinoponera* possam desempenhar importante papel na transdução de sinal e/ou comunicação química destas formigas, interagindo seletivamente e não-covalente com substâncias odorantes capazes de estimular o olfato e potencialmente desempenharem importantes atividades biológicas seja na comunicação social, seja na peçonha, tal como hipersensibilidade alérgica ou outra atividade tóxica quando injetada na presa em especial se for outro inseto.

Figura 17: Organização do cDNA dos contigs 17 (A) e 21 (B).



5'UTR (lilás), Peptídeo sinal (amarelo), Proteína madura (vermelho), 3'UTR (verde), sinal de poliadenilação (caixa amarela). O peptídeo sinal foi predito utilizando a ferramenta SignalP.

**Figura 18:** Alinhamento dos Precusores codificados pelos Contigs TX17\_DIQUA e TX21\_DIQUA com proteínas da família PBP/GOPB e *Venom Allergen 2/4*.



A) **TX17\_DIQUA** e **TX22\_DIQUA**: indicam respectivamente as sequências dos precusores codificados pelos contigs de mesmo nome identificados na biblioteca de cDNA da glândula de veneno de *Dinoponera quadriceps*. **G6DFG7\_DANPL**: *Odorant binding protein* da borboleta *Danaus plexippus* (Uniprot G6DFG7). **B1WA78\_SOLIN**: *Queen venom protein Sol i 4* da formiga *Solenopsis invicta* (Uniprot B1WA78). **VA2\_SOLSV**: *Venom allergen 2* da formiga *Solenopsis saevissima* (Uniprot A5X2H7). **VA2\_SOLIN**: *Venom allergen 2* da formiga *Solenopsis invicta* (Uniprot P35775). As cisteínas conservadas, envolvidas na formação de pontes dissulfeto intracadeia, estão indicadas por caixas conectadas por linhas cheias. As cisteínas livres envolvidas na formação de ponte dissulfeto intercadeia do homodímero são indicadas por uma caixa pontilhada. O peptídeo sinal está indicado por uma caixa de contorno amarelo, enquanto que o peptídeo maduro é indicado por uma caixa de contorno rosa. Os números indicam a quantidade de aminoácidos na cadeia alinhada. O grau de similaridade com o precursor codificado pelo contig TX17\_DIQUA é indicada em porcentagem. A matriz utilizada no alinhamento foi BLOSUM62. B) Utilizando o programa MEGA 5, foi criado um dendrograma representado o grau de similaridade das sequências alinhadas e que separa as sequências em dois grupos principais (clusters).

#### 4.4.1.3 Alérgenos *Pilosulina-símile*

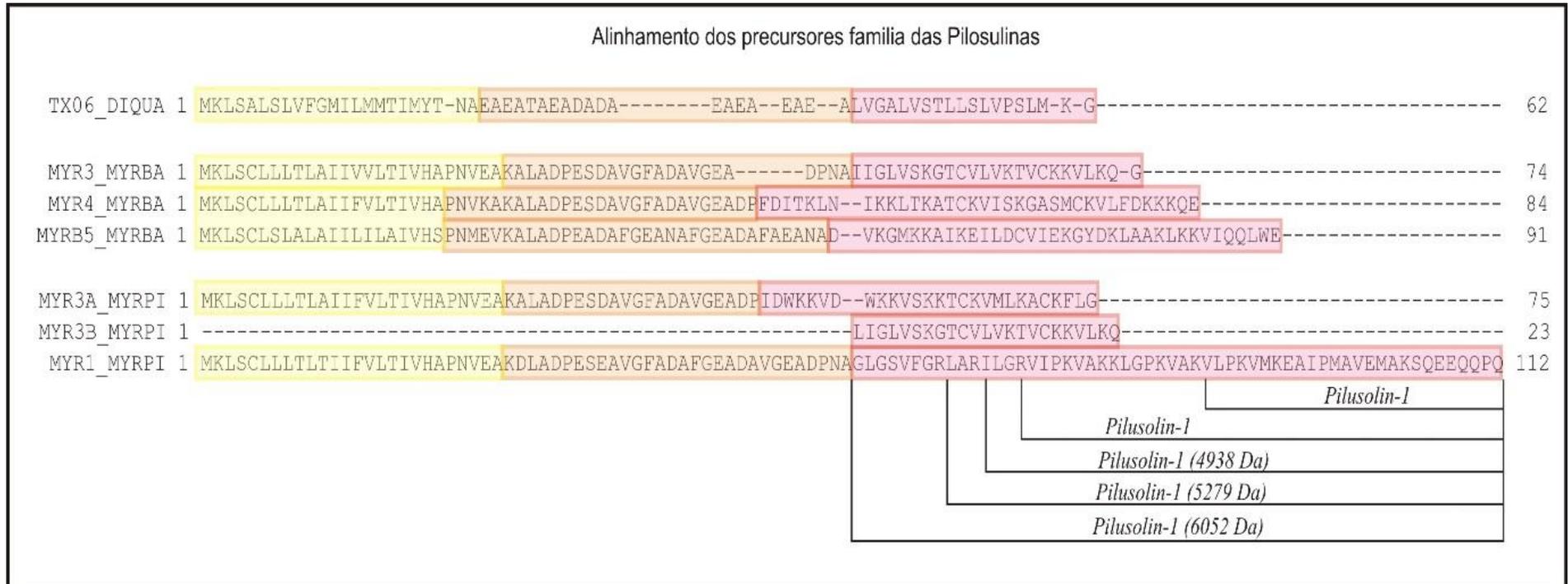
As toxinas alergênicas semelhantes à *Pilosulina* são codificadas pelo consenso do contig TX06\_DIQUA. Este apresenta 430pb, uma região 5'UTR de 51pb seguida de um ORF de 186pd, que se inicia na posição 52 com quadro de leitura de tradução que se estende até a posição 238pb. Em seguida observa-se a região 3'UTR que se estende da posição 239 a 430pb. O consenso do contig TX06\_DIQUA traduzido gera a uma proteína precursora de 62 aminoácidos, cujo peptídeo sinal predito é clivado entre as posições 23 e 24 (Aminoácidos TNA-EA) (Figura 19). Por predição, a TX06\_DIQUA pode apresentar um pró-peptídeo que se estende da posição 121-180pb da sequência de nucleotídeos (Figura 19), apresentando a toxina madura com pI de 8.71 e massa de 1898Da.

A TX06\_DIQUA apresentou maior homologia com MYR3A\_MYRPI de *Myrmecia pilosula* (Q26464) e com MYR4\_MYRBA de *M. banski* (Q68Y22). O alinhamento de TX06\_DIQUA com as demais proteínas da família pilosulina mostra que a homologia entre essas proteínas se deve ao peptídeo sinal e não à cadeia madura. Em geral apresentam um propeptídeo ácido, sendo que a cadeia madura das pilosulinas apresentam uma característica anfipática. Já a cadeia madura de TX06\_DIQUA apresenta uma característica hidrofóbica (Figura 20). Ainda analisando a figura 20 pode-se observar que as pilosulinas podem ser clivadas em diversos peptídeos menores, o que pode lhes conferir diferentes atividades biológicas.

Pilosulinas é um termo genérico para peptídeos de diversas origens de *Myrmecia pilosula*, e compreendem uma família de pequenos peptídeos mono- e diméricos (INAGAKI *et al.*, 2004; INAGAKI *et al.*, 2008). São os alérgenos mais abundantes no veneno de formigas do gênero *Myrmecia*. A Myr p 1 (Pilosulina 1 de *M. pilosula*), compreende uma sequência de 112 aminoácidos e pode ser clivada em cinco diferentes cadeias. A Myr p 2 (Pilosulina 2 ou Pilosulina 3a) forma um heterodímero com a pilosulina 3b através de uma ponte dissulfeto (DAVIES *et al.*, 2004). Além do seu papel alergênico estas proteínas têm sido relatadas por sua ação hemolítica, citotóxica e antimicrobiana (INAGAKI *et al.*, 2004; ZELEZETSKY *et al.*, 2005), efeitos variáveis de acordo com o subtipo. Por exemplo, a pilosulina 4 apresenta um potente efeito antimicrobiano e não possui efeito hemolítico ou citolítico; já a pilosulina 1, além do potente efeito antimicrobiano apresenta acentuada ação



**Figura 20:** Alinhamento de TX06\_DIQUA com precursores da família das pilosulinas.

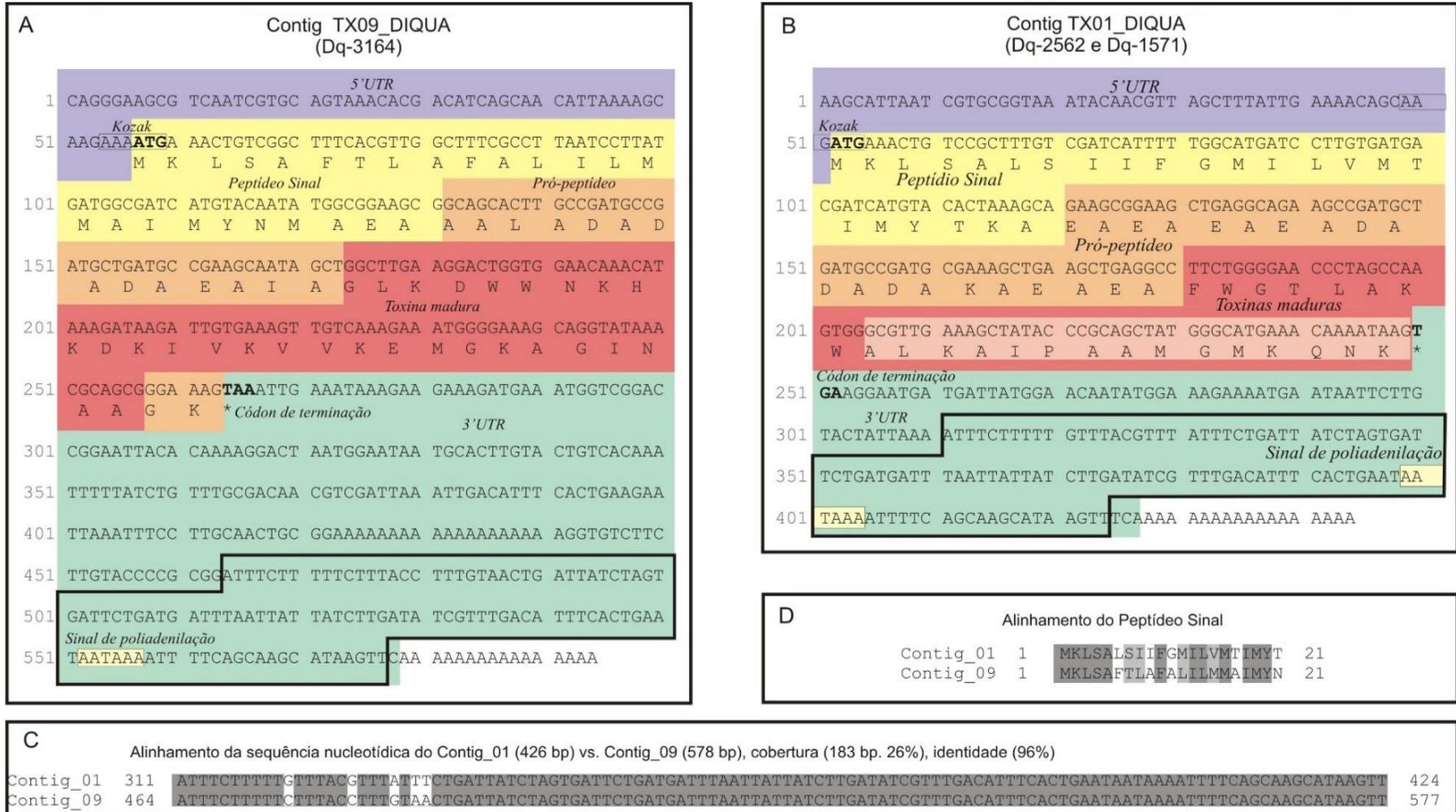


**TX06\_DIQUA:** precursor codificado pelo contig de mesmo nome identificado na biblioteca de cDNA da glândula de veneno de *Dinoponera quadriceps*. **MYR3\_MYRBA:** Pilosulin 3 da formiga *Myrmecia banksi* (Uniprot Q68Y23). **MYR4\_MYRBA:** Pilosulin 4 da formiga *Myrmecia banksi* (Uniprot Q68Y22). **MYR3A\_MYRPI:** Pilosulin 3a da formiga *Myrmecia pilosula* (Uniprot Q26464). **MYR3B\_MYRPI:** Pilosulin 3a da formiga *Myrmecia pilosula* (Uniprot P0C023). **MYR1\_MYRPI:** Pilosulin 1 da formiga *Myrmecia pilosula* (Uniprot Q07932). Os subprodutos de clivagem de MYR1\_MYRPI estão indicados. O peptídeo sinal está indicado por uma caixa amarela, o pró-peptídeo está indicado por uma caixa laranja, enquanto que o peptídeo maduro é indicado por uma caixa de contorno rosa.

#### **4.4.2 Dinoponeratoxinas**

As dinoponeratoxinas foram identificadas tanto através do sequenciamento por Sanger quanto por sequenciamento NGS. Estas toxinas estão representadas pelos contigs TX01\_DIQUA e TX09\_DIQUA do agrupamento das sequências obtidas por Sanger; e pelos contigs 01 e 02 das sequências obtidas por NGS. Essas duas novas dinoponeratoxinas representam 51% do total de precursores de toxinas. A organização do cDNA para as toxinas TX01\_DIQUA e TX09\_DIQUA, evidencia em lilás a extremidade 5'UTR, o sinal consenso de Kozak é evidenciado por uma caixa, o peptídeo sinal em amarelo, em laranja a região do pró-peptídeo, em vermelho a região que codifica a toxina madura e em verde a extremidade 3'UTR (Figura 21 A e B). Comparando os dois grupos de dinoponeratoxinas podemos observar que apresentam maior identidade na região não traduzida da extremidade 3'UTR, região evidenciada pela caixa em negrito (Figura 21 C). Alinhando as sequências de aminoácidos dos precursores traduzidos podemos observar homologia na região do peptídeo sinal (Figura 21 D). Não há similaridade na região da cadeia madura.

**Figura 21:** Organização do cDNA dos contigs 09 (A) e 01 (B), alinhamento das sequências nucleotídicas das extremidades 3' UTR (C) e dos aminoácidos que compõem o peptídeo sinal (D) dos contigs.



5'UTR (lilás), Peptídeo sinal (amarelo), Proteína madura (vermelho), 3'UTR (verde), sinal de poliadenilação (caixa amarela). O peptídeo sinal foi predito utilizando a ferramenta SignalP.

Avaliando cada contig individualmente observamos que o contig TX01\_DIQUA apresenta um códon de inicialização (ATG), precedido de uma sequência consenso Kozak clássica, que se inicia na posição 52. Esta organização resulta num ORF de 198 pb até um códon de terminação na posição 250. O cDNA apresenta 426pb composto por uma região 5'UTR de 51pb, seguida de uma região codificante (52-250pb) e uma região 3'UTR que se estende da posição 250 a 426pb. A proteína precursora foi submetida à ferramenta SignalP e foi determinado um ponto de clivagem entre as posições 23-24 do precursor (Aminoácidos TKA-EA), gerando uma toxina que provavelmente apresenta um pró-peptídeo (Figura 21 B).

O consenso do contig TX01\_DIQUA traduzido gerou uma proteína precursora que corresponde a um único fragmento homólogo (TX01\_DINAS, UniProtKB/Swiss-Prot:P0CF01) utilizando a ferramenta BlastX. A toxina resultante é composta por um pró-peptídeo, seguido pela toxina madura de caráter anfipático com *pl* predito 10.48 e massa provável 2562 Da. A sequência TX01\_DINAS (P0CF01) gera duas cadeias de Dinoponeratoxinas, a Da-2501 é clivada gerando a Da-1585, ambas estão presentes na peçonha de *D. australis* (JOHNSON *et al.*, 2010). Por similaridade, TX01\_DIQUA produz um precursor que sofre processamento semelhante, resultando num peptídeo maduro Dq-2562, o qual provavelmente é processado gerando Dq-1571(Figura 22A). De fato, a partir da análise proteômica (resultados apresentados com mais detalhes adiante neste trabalho), observamos a presença de quatro subprodutos sugerindo processamento adicional e distinto do que foi evidenciado para *D. australis*.

O contig TX09\_DIQUA apresenta um quadro de leitura de tradução em proteína que se inicia na posição 57, apresentando uma sequência consenso Kozak clássica resultando num ORF de 207pb até um códon de terminação na posição 264. O cDNA sequenciado apresenta 578pb composta por uma região 5'UTR de 56pb, seguida de uma região codificante (posição 57-264pb) e uma região 3'UTR mais longa que se estende da posição 265-578. A proteína precursora resultante apresenta cerca de 59 aminoácidos sendo composta por um peptídeo sinal com 25 aminoácidos preditos pela ferramenta SignalP, com possível ponto de clivagem entre as posições 25-16 (Aminoácidos AEA-AA) (Figura 21A).

O consenso do contig TX09\_DIQUA traduzido gerou uma proteína precursora com homologia com TX77\_DINAS Da-3177 e TX05\_DINAS Da-3105

(UniProtKB/Swiss-Prot:P0CF05 e P0CF02, respectivamente). A toxina resultante é composta de um pró-peptídeo hidrofóbico rico em resíduos de alanina, ácido aspártico com carga líquida negativa, seguido por uma toxina madura rica em resíduos aromáticos e básicos de caráter anfipático com *pI* predito de 9.88 e massa provável de 3164Da. Por similaridade com *D. australis*, o ponto de clivagem entre o pró-peptídeo e a toxina é na posição 39-40 (Aminoácidos AIA-GLK). Ainda em comparação com os peptídeos homólogos de *D.australis* o precursor também apresenta dois aminoácidos extras no C-terminal (GK) que provavelmente são excisados durante a secreção ou podem sofrer modificações (Figura 22 B).



Como observado por Johnson *et al.* (2010), estas toxinas compartilham similaridade variável com peptídeos antimicrobianos de outras formigas da subfamília Ponerinae (ponericin G e W3), assim como com peptídeos de outros organismos, como sapos (gaegurin-5, brevinin-1 PTa). De fato, o veneno total de *D. quadriceps* apresentou efeito antimicrobiano sobre *S. aureus*, *P. aeruginosa*, *E. coli* e *Salmonella sp.* (resultados não publicados). Os peptídeos classificados como dinoponeratoxinas por Cologna *et al.* (2013) apresentaram amplo espectro antimicrobiano, incluindo ação antifúngia. A comparação de dinoponeratoxinas com outros peptídeos antimicrobianos esclarece sobre a função desses peptídeos no veneno de *D. quadriceps* e de outras formigas ponerinae, embora a função destas toxinas ainda mereça ser mais profundamente investigada. Outro ponto interessante sobre os transcritos de dinoponeratoxinas é que o precursor apresenta certo grau de similaridade com o precursor de TX06\_DIQUA (Pilosulina símile) à extremidade 5'UTR principalmente no que concerne à região que codifica o peptídeo sinal e o pró-peptídeo. Assim, à luz da evolução das toxinas e sua diversificação, o compartilhamento de sequências de nucleotídeos entre dinoponeratoxinas e pilosulina é um indicativo de um gene parental comum e seu recrutamento como molde envolvendo variantes bioativas de peptídeos na peçonha.

Os peptídeos antimicrobianos (*Antimicrobial peptides* – AMPs) têm atraído o interesse de pesquisadores, principalmente mediante à necessidade de desenvolvimento de novos agentes antimicrobianos. Os AMPs são bem descritos em uma variedade de organismos e estão envolvidos na modulação de células do sistema imune (GANZ, 2003) e na ação antimicrobiana direta (WIESNER; VICINSKAS, 2010; POWERS; HANCOCK, 2003). Estes, em insetos, já compreendem um grupo de mais de 170 peptídeos (BULET *et al.*, 1999). Geralmente apresentam baixo peso molecular e são ricos em aminoácidos hidrofóbicos. Em formigas já foram descritos peptídeos antimicrobianos em *Myrmecia gulosa* (MACKINTOSH *et al.*, 1998), *Formica rufa* (TAGUSHI *et al.*, 1998), *Pachycondyla goeldi* (ORIVEL *et al.*, 2001), *Camponotus floridanus* (RATZKA *et al.*, 2012), *Tetramorium bicarinatum* (RIFFLET *et al.*, 2012), *Formica aquilonia* (VILJAKAINEN; PAMILO, 2005). Muitos deles apresentam características anfipática e hidrofóbicas, o que favorece sua ligação a fosfolípidios aniônicos, rompendo ou permeando a membrana bacteriana (BERTHOLD *et al.*, 2013). Embora alguns peptídeos antimicrobianos de insetos ainda não sejam

passíveis de classificação, algumas classes têm sido descritas, tais como: peptídeos alfa-hélice (Cepropin), peptídeos ricos em cisteínas (defensinas), peptídeos ricos em prolina (drosocin), peptídeos ricos em glicina (himenoptatecin) (REDDY *et al.*, 2004; ZHANG; ZHU, 2012). Alguns dos AMPs de formigas além de suas propriedades antimicrobianas, também apresentam efeitos alergênicos, hemolíticos e citotóxicos. Assim, muitos esforços têm sido aplicados no sentido de promover modificações em suas estruturas, com o objetivo de melhorar o efeito antimicrobiano e diminuir os efeitos tóxicos, como no caso das pilosulinas (*M. pilosula*) (ZELEZETSKY *et al.*, 2005), temporin (*Rana temporaria*) (SIMMACO *et al.*, 1996) e jelein 1 (*Apis mellifera*) (CAPPARELLI *et al.*, 2012).

Embora os peptídeos antimicrobianos sejam principalmente avaliados como potenciais alternativas para tratar doenças infecciosas o seu uso como peptídeos anticâncer tem sido ressaltado como uma estratégia terapêutica a ser explorada, principalmente porque o câncer apresenta elevada taxa de morbi/mortalidade e, portanto, também necessita do desenvolvimento de drogas mais seletivas e eficazes (GASPAR *et al.*, 2013). Sob este ponto de vista os CPPs (Cell-penetrating Peptides) recebem destaque por sua capacidade de penetrar a membrana das células em concentrações micromolares sem a necessidade de receptores e sem causar danos significativos à mesma (MADANI *et al.*, 2011). Outra importante característica dessas substâncias é sua associação covalente à outras moléculas, como fármacos anticâncer, podendo direcioná-los especificamente para às células tumorais (SHI *et al.*, 2012; JOBIN *et al.*, 2013; LIM *et al.*, 2013).

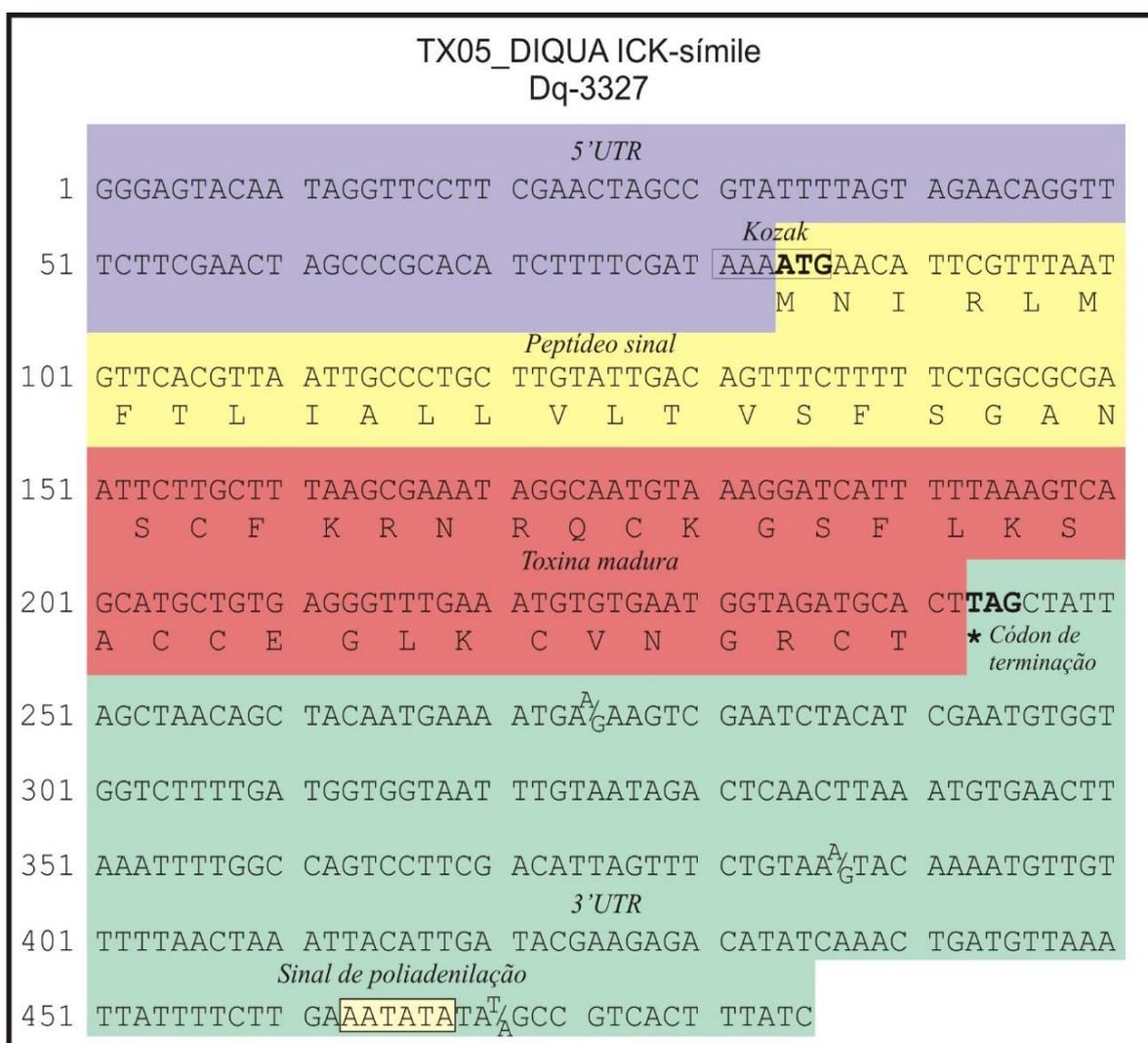
#### **4.4.3 Toxina ICK-símile**

No transcriptoma de *Dinoponera quadriceps* foram encontrados dois transcritos denominados contigs TX05\_DIQUA e NGS\_1144 que podem ser analisados conjuntamente, pois apesar de serem precursores distintos, ambos apresentam como elemento estrutural comum um domínio “Knottina”, o que permite uma discussão conjunta em termos gerais. Por serem proteínas pequenas e de baixa similaridade na sequência primária, e a pouca informação existente para sequências de formigas para as quais não existem dados estruturais, é necessária cuidadosa

interpretação de forma a detectar um padrão de interconectividade de cisteínas que permita detectar a classe a que pertencem estas sequências.

A sequência da proteína que é codificada pelo contig TX05\_DIQUA apresenta uma região 5'UTR de 83pb, seguida por uma região traduzida que se estende da posição 84-242pb. A região 3'UTR encontra-se entre as posições 243-485pb. A tradução deste contig gera uma proteína com um peptídeo sinal de 23 aminoácidos clivados entre as posições 23-24 (Aminoácidos GAN-SC), seguido de uma toxina madura de 30 aminoácidos, incluindo seis resíduos de cisteína. O ponto isoelétrico e massa preditos são 9.38 e 3327Da (Figura 23).

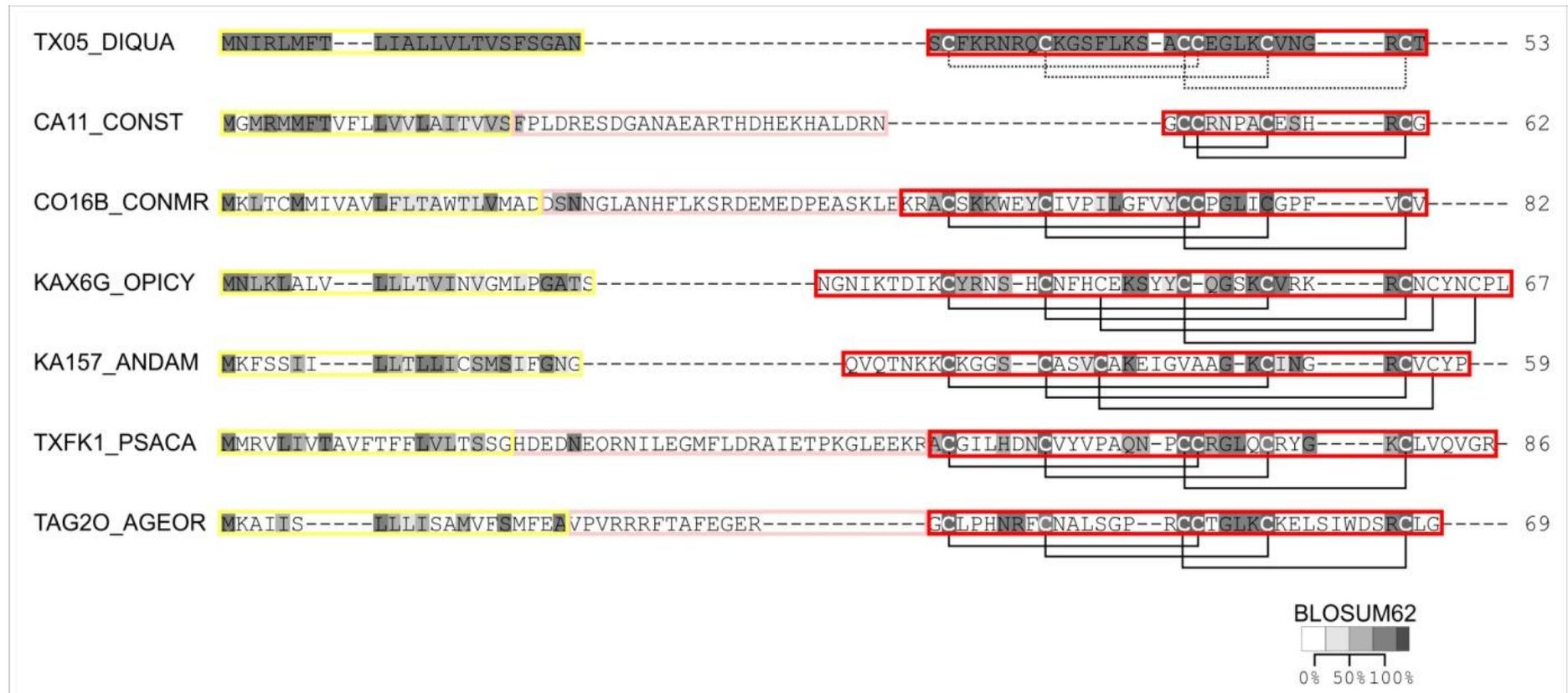
**Figura 23:** Organização do cDNA do contig 05.



5'UTR (lilás), Peptídeo sinal (amarelo), Proteína madura (vermelho), 3'UTR (verde), sinal de poliadenilação (caixa amarela). O peptídeo sinal foi predito utilizando a ferramenta SignalP.

A sequência nucleotídica do contig TX05\_DIQUA não encontra sequências nucleotídicas similares significantes quando submetida à ferramenta BlastN. Entretanto, apresenta um ORF que codifica para um precursor de 53 aminoácidos que produz alinhamentos de similaridade significativa na região do peptídeo sinal, após utilizar a ferramenta BlastP, enquanto que a parte madura diverge. O alinhamento das quatro sequências mais significativas, com maior cobertura do alinhamento está mostrado na figura 24.

**Figura 24:** Alinhamento do precursor codificado pelo contig TX05\_DIQUA com outras toxinas ICK-símile.



**TX05\_DIQUA:** Precursor codificado pelo contig TX05\_DIQUA de *Dinoponera quadriceps*; **CA11\_CONST:**  $\alpha$ -conotoxin-like S1.1 do caramujo *Conus striatus* (Uniprot: P0C1W4); **CO16B\_CONMR:**  $\mu$ -conotoxin MrVIB do caramujo *Conus marmoreus* (Uniprot Q26443); **KAX6G\_OPICY:** Potassium channel toxin  $\alpha$ -KTx 6.16 do escorpião *Opisthacanthus cayaporum* (Uniprot C5J896); **KA157\_ANDAM:** Potassium channel toxin  $\alpha$ -KTx 15.7 do escorpião *Androctonus amoreuxi* (Uniprot Q5K0E0); **TXFK1\_PSACA:** U1-theraphotoxin-Pc1a da aranha *Psalmopoeus cambridgei* (Uniprot P0C201) e **TAG20\_AGEOR:** U2-agatoxin-Ao1o da aranha *Agelena orientalis* (Uniprot Q5Y4X1). A interconectividade das cisteínas, envolvidas na formação de pontes dissulfeto, está indicado por linhas cheias abaixo da sequência. As linhas tracejadas representam o possível padrão de pontes formadas no precursor de TX05\_DIQUA inferido por similaridade com outras sequências. As cisteínas conservadas em relação a TX05\_DIQUA estão indicadas em branco. Os peptídeos sinais estão envoltos por uma caixa amarela; os pró-peptídeos por uma caixa rosa e as toxinas maduras por uma caixa vermelha. A matriz de similaridade utilizada no alinhamento foi BLOSUM62.

Podemos observar que existem diferentes padrões de organização desses precursores, sendo a região do peptídeo sinal a mais conservada. Existem sequências com e sem pró-peptídeo, assim como diferenças no tamanho das toxinas maduras e em especial grande diversidade no arranjo do padrão das pontes dissulfeto e no número de resíduos existentes entre as cisteínas (*loops*). De maneira geral a identidade é baixa variando de 28 a 35% (similaridade de 35-53%). A toxina  $\alpha$ -conotoxin-like S1.1 (CA11\_CONST de *Conus striatus*) tem 13 aminoácidos e apenas duas pontes SS. Já a toxina *Potassium channel toxin*  $\alpha$ -KTx 6.16 (KAX6G\_OPICY de *Opisthacanthus cayaporum*) apresenta 43 resíduos e quatro pontes SS. Portanto, é de se esperar grande diversidade estrutural nas sequências alinhadas dada a importância das pontes dissulfeto na estrutura tridimensional. Além disso, as sequências alinhadas pertencem a diferentes famílias de toxinas e exercem diferentes funções. Assim, CA11\_CONST pertence à superfamília conotoxinas A e bloqueia receptores nicotínicos de acetilcolina. CO16B\_CONMR de *Conus marmoreus*, pertence à superfamília conotoxinas O1 e bloqueia preferencialmente canais de sódio dependentes de voltagem (Nav1.2 e Nav1.4). KAX6G\_OPICY inibe canais de potássio dependentes de voltagem, assim como a outra toxina de escorpião KA157\_ANDAM (*Androctonus amoreuxi*).

Se por um lado as toxinas de escorpião não apresentam pró-peptídeos, tal como observado no precursor de TX05\_DIQUA, o padrão de pontes dissulfeto da toxina madura é muito diferente. No alinhamento da figura 18 acrescentamos duas sequências de precursores de toxinas de aranha, ainda que a similaridade total com o precursor de TX05\_DIQUA seja mais baixa, o alinhamento das cisteínas conservadas estabelece idêntico padrão. São elas a U1-theraphotoxin-Pc1a (TXFK1\_PSACA de *Psalmopoeus cambridgei*) e a U2-agatoxin-Ao1o (TAG2O\_AGEOR de *Agelena orientalis*). Estas toxinas pertencem à subfamília TOXIN\_7 do clã das  $\omega$ -toxinas-símile (PFAM CL0083). A subfamília que empresta nome ao clã, as  $\omega$ -toxinas de aranhas, bloqueiam canais de cálcio dependentes de voltagem e purinoreceptores incluindo inúmeras toxinas com atividade inseticida. Já o clã  $\omega$ -toxinas-símile compreende inúmeras outras toxinas de diversas famílias de diferentes origens biológicas e com diferentes alvos farmacológicos conhecidos ou não. Para uma melhor classificação funcional da toxina gerada pelo precursor codificado pelo contig TX05\_DIQUA é necessário uma inspeção mais acurada de sua

parte madura, mas que, entretanto, é pequena e rica em pontes dissulfeto, (30 resíduos, 3 pontes dissulfeto e massa predita 3327 Da).

Proteínas com domínios ricos em pontes dissulfeto (*Small Disulfid-rich protein domains*), caracterizam pequenos domínios protéicos (em geral, < 100 resíduos) nos quais o padrão geral de enovelamento (*fold*) é estabilizado primariamente pela formação de pontes dissulfeto e em menor grau pela estrutura secundária e pelas interações hidrofóbicas dessas proteínas pequenas.

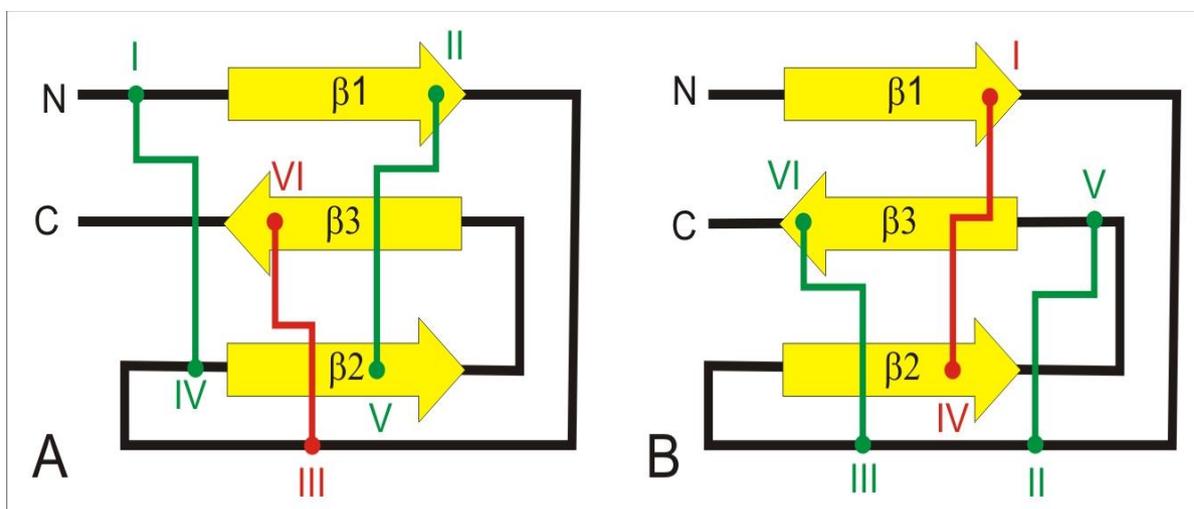
Esses domínios ricos em dissulfeto exercem uma larga série de atividades biológicas muito distintas e de grande interesse científico e médico (LAVERGNE *et al.*, 2012). Tais como, fatores de crescimento, toxinas inibidoras de canais iônicos (*Inhibitor cystein knot - ICKs*), inibidores de enzimas, domínios de ligação e dimerização em proteínas multi-domínio, hormônios, feromônios, alérgenos, defensinas, lectinas, CRISPs (*Cystein Rich Secretd Proteins*), etc. Este tipo de domínio é principalmente encontrado nas proteínas, ora de forma independente (domínio único), ora como domínios em grandes polipeptídios (multi-domínio) e, algumas vezes com repetições *in tandem*. Mais de 3000 tipos de pequenos domínios protéicos ricos em dissulfeto são encontrados em bancos de dados de classificação estrutural como o SCOP (*Structural Classification of Proteins*) (PETHICA *et al.*, 2012) com estrutura tridimensional determinada. Estes domínios podem ser agrupados em diversos subgrupos de cerca de mais de 40 subtipos estruturais, incluindo a superfamílias de proteínas “*knottin-like*”. Esses grupos de homólogos estruturais podem incluir casos de evolução convergente e/ou divergente em relação à função exercida por essas proteínas estabilizadas por pontes dissulfeto.

Vizinhos estruturais ou evolutivos de proteínas globulares grandes (>100 resíduos) são facilmente identificados e classificados utilizando-se ferramentas de comparação de sequências e estruturas, tal como o BlastP, Psi-Blast e DALI (ALTSCHUL *et al.*, 1997). Entretanto, esses métodos tendem a produzir resultados menos confiáveis quando se trata de proteínas pequenas e em geral com conservação do domínio estrutural apesar da baixa similaridade na sequência primária. Assim sendo, a classificação de pequenos domínios ricos em dissulfeto não é uma tarefa simples e, em geral, a análise não pode ser automatizada, exigindo considerável esforço de anotação manual (TOWSE; DAGGETT, 2012). Apesar disso, existem algumas ferramentas automáticas que tentam preencher esta lacuna comparando a

geometria e a topologia do arranjo das cisteínas envolvidas na formação de pontes dissulfeto (MARQUES *et al.*, 2010). Como por exemplo, o programas KNOT-MATCH (MAS *et al.*, 2001). Existem também bancos de dados especializados o Knottin-Website (GELLY *et al.*, 2004; KNOTTIN Database) com foco em subgrupos específicos como toxinas de moluscos marinhos do gênero *Conus* (PUILLANDRE *et al.*, 2012), escorpiões (GAO *et al.*, 2013), e aranhas (ESCOUBAS, 2006). Este tipo de abordagem pretende identificar um padrão ou assinatura no arranjo das cisteínas conservadas, procurando identificar sua conectividade e o tamanho dos laços ou “loops” formados pelo número de resíduos que existem entre dois resíduos de cisteína ligados por uma ponte dissulfeto (GUPTA *et al.*, 2004). Ainda que a relação evolutiva destes grupos de proteínas contendo pequenos domínios ricos em dissulfeto, estabelecidos por esta abordagem, deva ser cuidadosamente interpretado, este método permite estabelecer indicadores de homologia valiosos ou fatores biologicamente relevantes, tais como similaridade da sequência, função biológica, enovelamento topológico, entre outras características estruturais que vão além do simples padrão de interconectividade de cisteínas conservadas.

O clã de proteínas contendo domínios *Knottin-like* respondem por quase 40% da totalidade de proteínas contendo pequenos domínios ricos em dissulfeto, (vide SCOP #5715). Este grupo se caracteriza por conter um tipo de topologia de interconectividade das pontes dissulfeto muito particular denominado dissulfeto através de nó dissulfeto (*disulfide through disulfide knot*). Este dito nó é formado quando uma ponte dissulfeto cruza o anel macrocíclico formado por outras duas pontes dissulfeto e o esqueleto da estrutura proteica interconectada. Conseqüentemente, as Knottinas devem ter pelo menos seis cisteínas conservadas formando 3 pontes dissulfeto (Figura 25).

**Figura 25:** Comparação entre domínio *Knottin*. A) Representação esquemática de uma *Knottina*. B) Representação esquemática da conectividade um nó de cisteínas presente em fator de crescimento.



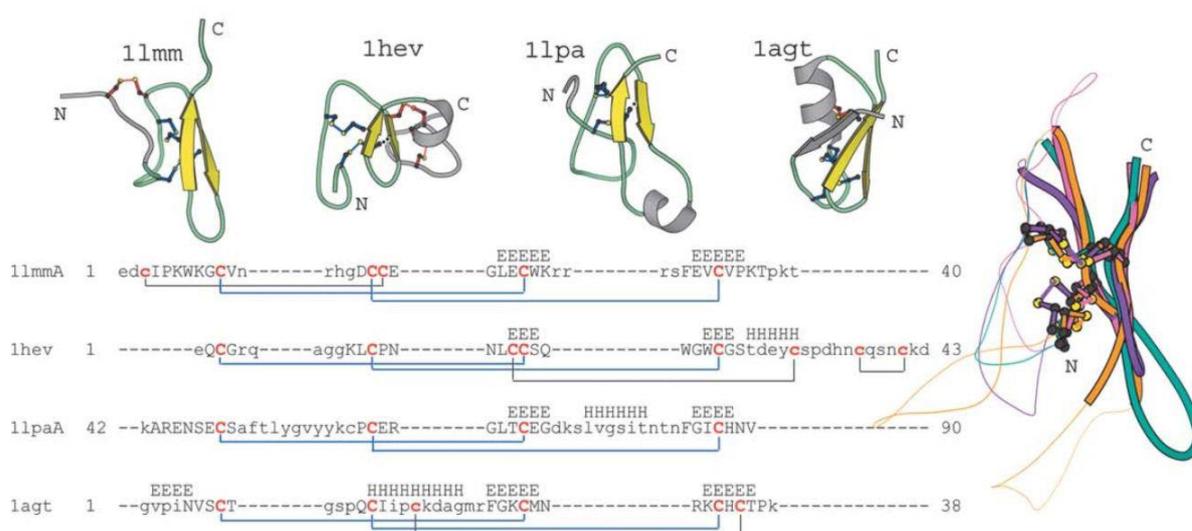
O esqueleto da estrutura primária é representado por linhas pretas, sendo indicado o N-terminal (N) e o C-terminal (C). Triplas folhas β-pregueadas anti-paralelas são indicadas por setas amarelas. O padrão de interconectividade das cisteínas via ponte dissulfeto é mostrado por linhas verdes e vermelhas numeradas em algarismos romanos, sendo que a ponte formada por III-VI(A) ou I-IV (B) (vermelha) passa por dentro das pontes I-IV e II-V (A) ou II-V e III-VI (Verdes).

Os domínios knottina de fatores de crescimento também apresentam um nó dissulfeto. Entretanto, o padrão de interconectividade é diferente (Figura 25B) e não pode ser sobreposto ao padrão clássico mostrado na figura 25A. Por outro lado, por se tratar de uma representação esquemática, outros elementos estruturais podem estar presentes, como ausência da folha β1 ou presença de α-helices “*interloops*”. Entretanto, o grampo-β (*β-hairpin*) formado pelas duas últimas folhas beta também é muito conservado, vide Figura 26. Algumas proteínas podem, ainda que similares a uma knottina, não apresentarem o nó dissulfeto representado pela figura 25A/B, e portanto não podem ser consideradas verdadeiras Knottinas.

Dentro da superfamília estrutural das Knottinas existem proteínas não relacionadas evolutivamente que apresentam o domínio knottina como resultado de evolução convergente. Por outro lado, knottinas homólogas (mesma origem gênica) podem apresentar funções muito diferentes, como resultado de evolução divergente. Assim sendo, existem knottinas de origem muito diferentes, tais como: Inibidores de serinoproteases de plantas, defensinas de plantas e insetos, toxinas de serpentes, de artrópodes (anêmonas do mar, aranhas, escorpiões, carrapatos e límulos), peptídeos

antimicrobianos, peptídeos inibidores de receptores de sabor doce do tipo Gurmarin-símile e proteínas de sinalização (*Human Agouti-related protein*), inibidores de serino proteases de sanguessugas e insetos hematófagos, grupos de proteínas de diferentes origens contendo multi-domínios EGF-símile, IGFBPs (*insulin-like growth factor-binding proteins*), Elaffin-símile, etc. A figura 26 apresenta exemplos de algumas dessas Knottinas de domínio único de diferentes origens alinhadas.

**Figura 26:** Alinhamento de knottinas de diferentes origens mostrando núcleo de conservação estrutural.



Nas figuras tridimensionais das Knottinas, o núcleo estrutural comum é mostrado em cor, sendo amarelas as folhas- $\beta$  antiparalelas e verde os segmentos *coil*. As  $\alpha$ -hélices e outros elementos que não fazem parte do núcleo estrutural são mostrados em cinza. As pontes dissulfeto estão representadas nas cores vermelho e azul no formato de bola-bastão. Um alinhamento estrutural é mostrado abaixo das figuras. Os resíduos em caixa alta representam resíduos estruturalmente alinhados enquanto que os resíduos em caixa baixa, não se alinham nas estruturas tridimensionais. As cisteínas altamente conservadas na família e que formam pontes SS são mostradas em vermelho no alinhamento conectados por linhas. Os números de acesso no RCSB-PDB são usados para identificar cada proteína mostrada na figura: Aranha 1LMM - *Psalmopoeus cambridgei psalmotoxin 1*; Planta 1HEV - *Hevea brasiliensis hevein*; Suíno 1LPA - *Sus scrofa colipase C-terminal domain* e Escorpião 1AGT - *Leiurus quinquestriatus hebraeus agitoxin*. As pontes dissulfeto em azul formam as pontes SS cruzadas características. Os elementos da estrutura secundária são mostrados acima do alinhamento, (H:  $\alpha$ -hélice e E: Folha- $\beta$ ). Os números flanqueando nas sequências indicam o primeiro e o último resíduos que compõem a sequência do domínio. A sobreposição mostra o padrão estrutural chave das Knottinas ( $\beta$ -hairpin e nó de cisteínas) das estruturas de 1lmm (verde), 1hev (roxo), 1lpa (laranja), 1agt, (rosa).

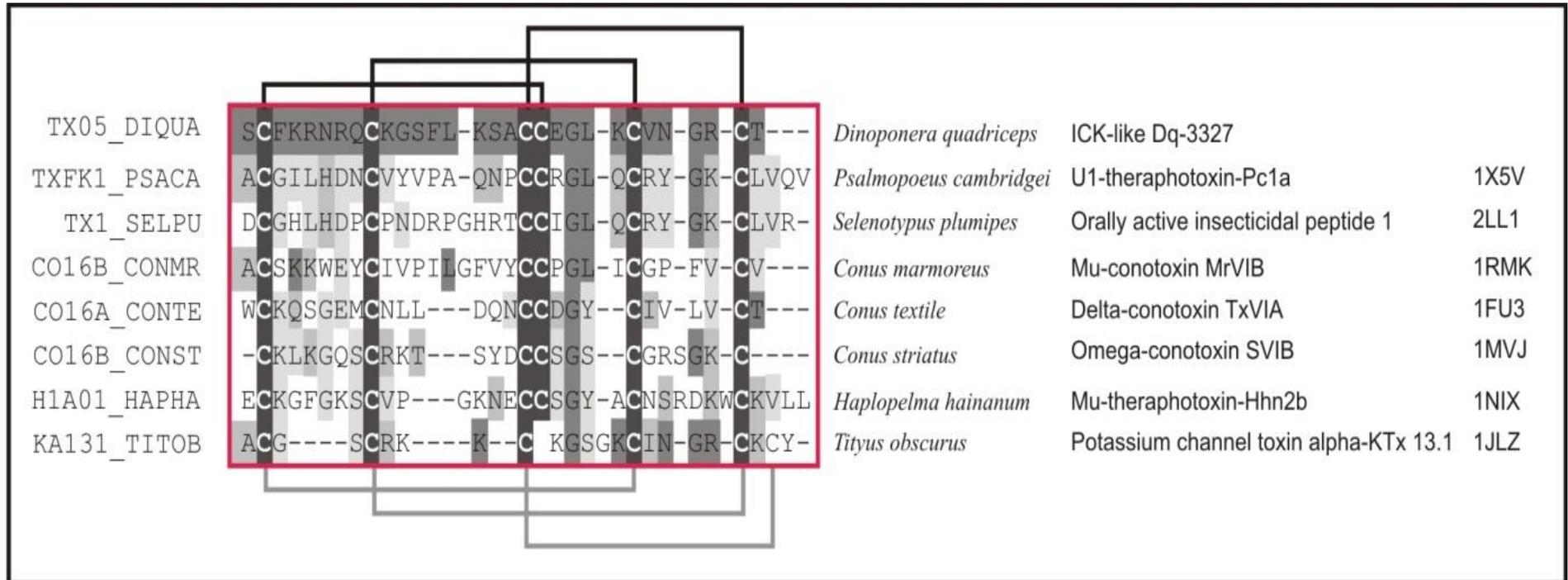
Muitas proteínas e peptídeos, contendo em geral, um único domínio independente de knottina, são denominados como “*Inhibitors Cystein Knots - ICKs*”. Entretanto, para fins de classificação, este nome deveria ser reservado para uma

subclasse especial de knottinas que compreende em geral, toxinas secretadas nas peçonhas animais, capazes de inibirem diversos subtipos de canais iônicos.

O clã de proteínas contendo um único domínio independente knottina-símile é denominado de superfamília  $\omega$ -toxina-símile por similaridade com toxinas de aranha que atuam em canais de cálcio dependentes de voltagem (Cav), tal como a Agitoxina. Entretanto, as  $\omega$ -toxina-símile podem atuar em outros canais e comumente são promíscuas, ou seja, atuam em diferentes canais iônicos. Além disso, grande parte das toxinas deste grupo não estão caracterizadas biologicamente. Por ser um grupo muito extenso de peptídeos de diferentes origens, a superfamília se subdivide em 19 outras superfamílias (Protein Family Database/PFAM CL083). A superfamília  $\omega$ -toxina é composta principalmente por  $\omega$ -conotoxinas e por toxinas de aranhas (tipo  $\omega$ -agatoxin) que atuam em canais de cálcio dependentes de voltagem denominada família  $\omega$ -toxina propriamente dita. Mas inclui também famílias de toxinas de escorpião;  $\mu$ -conotoxinas; neuropeptídeos agouti-símile; *Tachistatin-B*; *Albumin-I hormone peptide*; além de uma série de toxinas que atuam em diversos canais iônicos, em especial de sódio e potássio voltagem dependentes, tais como atracotoxinas; conotoxinas da superfamília-O. Proteínas de aranhas com atividade inseticida; lectinas de aranha; *Magi-peptide toxins*; *Hexatelidae-toxins*; Conotoxinas TVIIA/GS; Conotoxinas da superfamília-A; Conotoxinas da superfamília-P; Toxinas de insetos assassinos (Ptu1 de *Peirates turpis*); toxinas de aranha e mais recentemente a PhcrTx1 da anêmona do mar *Phymanthus crucifer* (RODRÍGUEZ *et al.*, 2013).

Utilizamos o banco de dados *Knottin Database* para melhor elucidar a sequência da potencial toxina gerada pelo precursor codificado pelo transcrito TX05\_DIQUA. Utilizando tanto o precursor completo de 53 resíduos e/ou apenas a parte madura de 30 resíduos, cuja massa predita corresponde a 3327 Da. O *Knottin Database* reconhece o padrão de seis cisteínas envolvido na formação das três pontes dissulfeto que caracteriza uma knottina e que a toxina potencial codificada por TX05\_DIQUA tem maior similaridade principalmente com toxinas de aranhas do clã  $\omega$ -toxina-símile. Sendo que o número de resíduos entre os *loops* inter-cisteína, também são encontradas em outras toxinas de aranha ( $\omega$ -toxina-símile) e conotoxinas. Para confirmar este achado resolvemos alinhar apenas a parte madura da toxina codificada pelo contig TX05\_DIQUA com outras toxinas de farmacologia bem conhecida e que possuem estrutura tridimensional determinada (Figura 27).

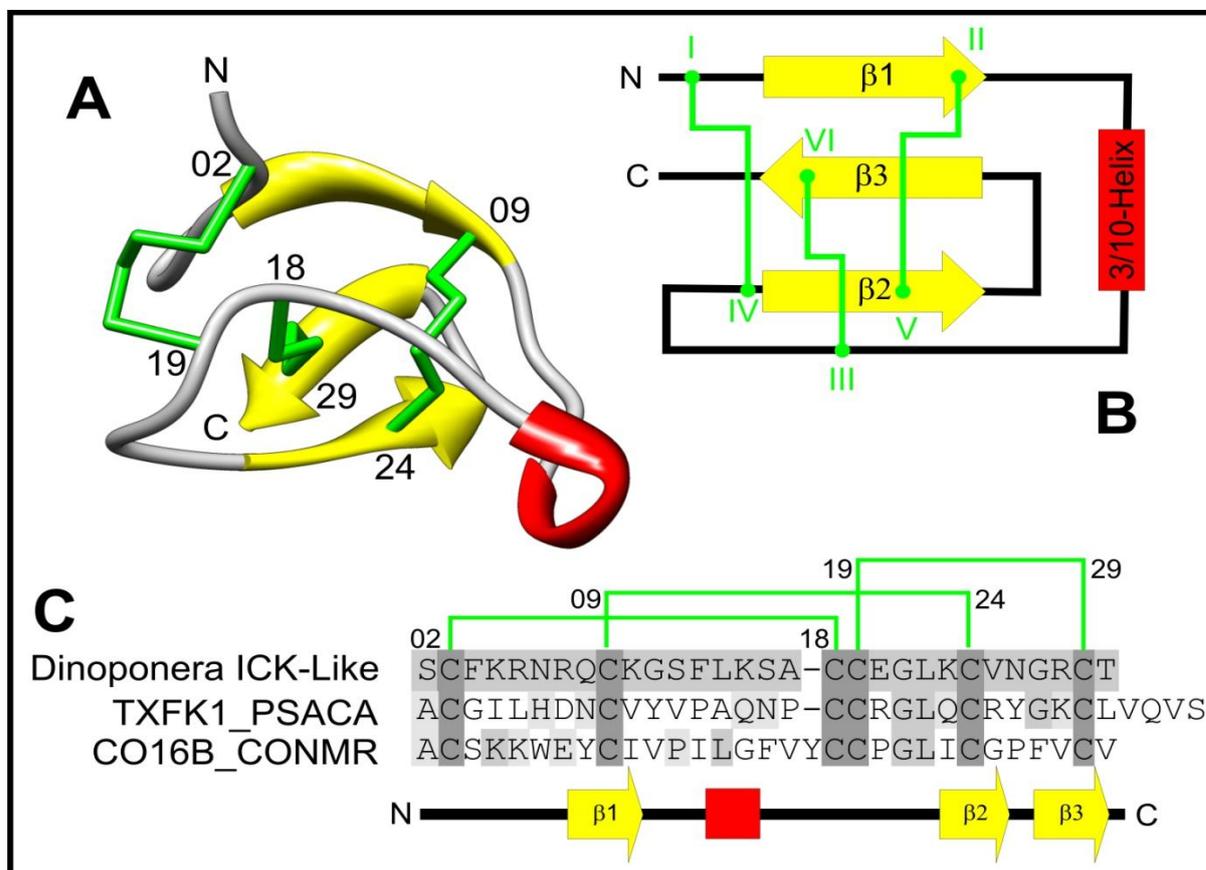
**Figura 27:** Alinhamento de toxinas maduras, com arranjo *Konttin* e que apresentam estrutura tridimensional.



**TX05\_DIQUA:** Parte madura do precursor codificado pelo contig TX05\_DIQUA de *Dinoponera quadriceps*. **TXFK1\_PSACA:** U1-theraphotoxin-Pc1a da aranha *Psalmopoeus cambridgei* (Uniprot P0C201). **TX1\_SELPU:** Orally active insecticidal peptide 1 da aranha *Selenotypus plumipes* (Uniprot K7N5K9). **CO16B\_CONMR:**  $\mu$ -conotoxin MrVIB do molusco *Conus marmoreus* (Uniprot Q26443). **CO16A\_CONTE:**  $\delta$ -conotoxin TxVIA do molusco *Conus textile* (Uniprot P18511). **CO16B\_CONST:**  $\omega$ -conotoxin SVIB do molusco *Conus striatus* (Uniprot P28881). **H1A01\_HAPHA:**  $\mu$ -theraphotoxin-Hhn2b da aranha *Haplopelma hainanum* (Uniprot D2Y1X6) e **KA131\_TITOB:** Potassium channel toxin alpha-KTx 13.1 do escorpião *Tityus obscurus* (Uniprot P83243). A caixa vermelha indica a parte madura da toxina. O padrão de pontes dissulfeto entre as cisteínas conservadas está indicado por linhas de conexão, (notar o padrão de interconectividade diferente da toxina de KA131\_TITOB indicado em cinza claro). O número de acesso da estrutura tridimensional no banco de dados RCSB PDB está indicado à direita do nome da toxina. A similaridade está indicada pela escala da matriz BLOSUM62.

Podemos observar, na figura 27, grande conservação do padrão de pontes dissulfeto nas sequências escolhidas e o mesmo número de resíduos inter-cisteínas (*loops*), exceto pela toxina indicada por KA131\_TITOB (Potassium channel TX $\alpha$ 13.1 de *Tityus obscurus*) que usamos como um controle, uma vez que esta toxina apresenta um padrão de interconectividade de cisteínas diferente, ainda que seja uma Knottin-símile. Com base neste alinhamento foi realizado um modelo tridimensional da toxina madura codificada pelo contig TX05\_DIQUA por homologia com essas estruturas bem conhecidas. O resultado do modelo de homologia tridimensional é mostrado na figura 28 A, B e C.

**Figura 28:** Modelo proposto para a toxina ICK-símile de *D. quadriceps*. A) Modelo tridimensional da toxina ICK-símile codificada por TX05\_DIQUA. B) Representação esquemática de uma knottina. C) O alinhamento das sequências das duas estruturas que geraram o melhor modelo de homologia tridimensional da toxina de TX05\_DIQUA com a estrutura secundária indicada abaixo.



As pontes dissulfeto são mostradas em verde, sendo as cisteínas numeradas de acordo com sua posição na cadeia madura. As folhas  $\beta$ -pregueadas antiparalelas em amarelo e em vermelho a  $\alpha$ -hélice. O N-terminal e C-terminal da cadeia estão indicados respectivamente por (N) e (C). **TXFK1\_PSACA**: *U1-theraphotoxin-Pc1a* da aranha *Psalmopoeus cambridgei* (Uniprot P0C201). **CO16B\_CONST**:  *$\omega$ -conotoxin SVIB* do caramujo *Conus striatus* (Uniprot P28881). A matriz de similaridade utilizada foi BLOSUM62.

Observa-se claramente no modelo gerado o padrão típico da topologia de interconexividade das pontes dissulfeto denominado: dissulfeto através de nó dissulfeto (*disulfide through disulfide knot*) que caracteriza uma knottina. Podemos inferir por similaridade que o contig TX05\_DIQUA realmente codifica para uma possível toxina madura pertencente a essa classe  $\omega$ -toxina-símile e possivelmente inibindo algum tipo de canal iônico especialmente de insetos. Por essa razão, podemos denominar este peptídeo como uma toxina ICK-símile. No entanto, sua atividade em canais iônicos assim como sua toxicidade em insetos ou outras presas só pode ser esclarecido experimentalmente.

Recentemente, foram descritas novas toxinas de aranha contendo dois domínios knottina-símile repetidos. A toxina majoritária CpTx1 da aranha *Cheiracanthium punctorium* de saco amarelo é formado por dois domínios diferentes, porém homólogos de Knottinas (VASSILEVSKI *et al.*, 2010). Caso similar também foi relatado em uma “toxina bivalente” da peçonha de tarântula, capaz de inibir receptores de capsaicina TRPV1 (BOHLEN *et al.*, 2010). Os dois domínios knottina-símile *in tandem*, atuam sinergicamente. Diversificando ainda mais a classificação de knottinas da superfamília  $\omega$ -toxins-símile, foi descrito recentemente uma toxina inseticida de aranha com um domínio knottina-símile, mas que não é uma knottina, uma vez que a molécula só tem duas pontes dissulfeto. Portanto, o “nó de pontes dissulfeto” característico não pode ser formado. Entretanto a homologia dessa toxina de aranha com outras toxinas da família Omega é claramente estabelecido (MORITA *et al.*, 2011).

O domínio Knottina destas toxinas, devido ao seu tamanho compacto e altamente estabilizada por pontes dissulfeto é tolerante a grandes alterações na composição da sequência primária, servindo de base para obtenção de novos fármacos (GRACY; CHICE, 2011). Podemos citar como exemplo o PRIALT® (Azur Pharma Inc -USA), droga aprovada pelo FDA em 2004 para o tratamento de dor severa crônica por administração intratecal derivada da Omega-conotoxina MVIIA (WALLACE, 2006). A diversidade e plasticidade destas moléculas é muito grande. Assim, knottinas são comumente encontradas em novas fontes de origem animal, tal como a Asteropina A isolada em esponjas marinhas (TAKADA *et al.*, 2006). Outras toxinas como por exemplo a Tx2-6 de *Phoneutria nigriventer* atrai atenção por um possível papel no tratamento de disfunção peniana (NUNES *et al.*, 2008). Recentemente knottinas de escorpião (Imperotoxina A e Maurocalcina) têm mostrado capacidade de atravessar membranas celulares e interferir na liberação de  $Ca^{+2}$  *in vivo* (GURROLA *et al.*, 2010; ESTÈVE *et al.*, 2005). Sendo capazes de carrear outras moléculas farmacologicamente ativas para dentro das células (CASCALES *et al.*, 2011) com evidentes aplicações em diagnóstico por imagem e aplicações como quimioterápicos.

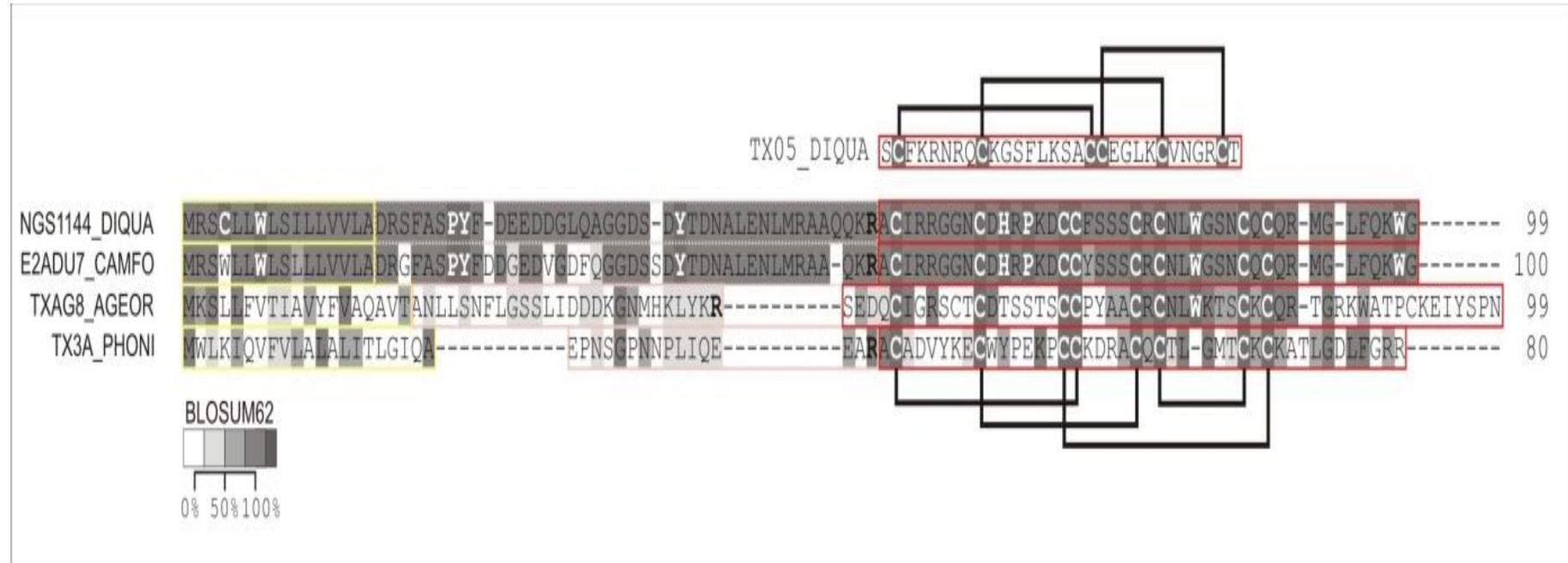
No sequenciamento por NGS identificamos outro precursor ICK-símile codificado pelo contig NGS\_1144. No Blast os principais alinhamentos com similaridade significativa com a sequência codificada pelo contig NGS\_1144, foram

sequências preditas derivadas da anotação automática de projetos de sequenciamento de genomas de Himenópteros. Por exemplo, a sequências com alta cobertura *U8-agatoxin-Ao1a-like* isoforma 1 da mamangaba (Bumblebee) *Bombus impatiens* (XP\_003485909.1) apresentando 84% de identidade e 89% de similaridade ou a sequêcia predita de *U8-agatoxin-Ao1a-like isoform 1* da abelha *Apis mellifera* (XP\_003249808.1) que apresenta 76% de identidade e 82% de similaridade. Estas sequências apresentam alta similaridade e cobertura completa entre as sequências alinhadas. Com menor grau de cobertura mas ainda com alta similaridade o precursor codificado pelo transcrito NGS\_1144 apresenta alinhamentos de alta significância com sequências também oriundas de genomas de formigas, tais como:

- Proteína hipotética EAG\_10432 de *Camponotus floridanus* (E2ADU7\_CAMFO) com 85% de identidade e 88% de similaridade;
- Proteína hipotética SINV\_06376 de *Solenopsis invicta* (E9ISS2\_SOLIN) com 68% de identidade e 77% de similaridade;
- Proteína hipotética EAI\_17246 de *Harpegnathos saltator* (E2B7Z7\_HARSA) com 91% de identidade e 97% de similaridade mas correspondendo apenas a uma parte de 44 resíduos alinhados, 81-125.

Estas sequências se caracterizam por conterem um domínio knottina, associado a outros elementos estruturais, em uma cadeia precursora de 99 a 100 aminoácidos (Figura 29).

**Figura 29:** Alinhamento do precursor codificado pelo contig NGS\_1144 de *D. quadriceps*, com homólogos que apresentam domínio Knottina.



**NGS1144\_DIQUA:** precursor codificado pelo contig NGS\_1144 de *Dinoponera quadriceps*. **E2ADU7\_CAMFO:** Proteína hipotética EAG\_10432 da formiga *Camponotus floridanus* (Uniprot E2ADU7). **TXAG8\_AGEOR:** *U8-agatoxin-Ao1a* da aranha *Agelena orientalis* (UNIPROT Q5Y4U4). **TX3A\_PHONI:** *U6-ctenitoxin-Pn1a* da aranha *Phoneutria nigriventer* (Uniprot P81793). **TX05\_DIQUA:** indica a toxina madura codificada pelo contig TX05\_DIQUA de *Dinoponera quadriceps*. Os peptídeos sinais são indicados por caixas de contorno amarelo. Os pró-peptídeos estão indicados por caixas de contorno rosa, sendo que os contornos pontilhados indicam inferência por similaridade. Os resíduos de arginina (**R**) indicados em negrito, indicam o ponto de clivagem do pró-peptídeo. As cadeias das toxinas maduras estão delimitadas por contornos vermelhos. A interconectividade das pontes dissulfeto é indicada por linhas pretas conectando as cisteínas. A matriz de similaridade utilizada no alinhamento foi BLOSUM62.

Para determinar o peptídeo sinal do precursor foi utilizada a ferramenta SignalP, que prediz a existência de um sinal de secreção N-terminal de 16 resíduos. Por similaridade, podemos inferir a existência de um pró-peptídeo que é clivado da cadeia após um resíduo de arginina conservado nas sequências alinhadas. A existência desse pró-peptídeo distingue o precursor de NGS\_1144 do precursor de TX05\_DIQUA.

Outra diferença entre as cadeias é que a parte madura da proteína codificada pelo contig NGS\_1144 é formada por 43 resíduos, com massa molecular predita de 4910 Da e contendo 8 cisteínas. Portanto, o NGS-1144 deve formar quatro pontes dissulfeto tal como indicado na figura 29. Já a TX05\_DIQUA apresenta 30 resíduos, uma massa molecular predita de 3327 Da e as três pontes de cisteínas conservadas do domínio Knottin. Assim pudemos evidenciar dois transcritos distintos com domínio Knottin (ICK-like) na biblioteca de cDNA da glândula de veneno de *D. quadriceps*.

Sequências similares são encontradas no banco de dados Knottin Database, indicando que a sequência NGS\_1144 também pertence à superfamília das  $\omega$ -toxina-símile. Em especial a sequência *U8-agatoxin-Ao1a* da aranha *Agelena orientalis* (TXAG8\_AGEOR) indicada na anotação automatizada das sequências hipotéticas dos genomas de Himenopteros. Como o padrão de pontes dissulfeto não foi experimentalmente determinado na toxina de aranha *Agelena orientalis U8-agatoxin-Ao1a* (TXAG8\_AGEOR), acrescentamos um similar de estrutura bem conhecida e arranjo de pontes SS determinado que é a toxina de aranha *Phoneutria nigriventer U6-ctenitoxin-Pn1a* (TX3A\_PHONI). Para termos de comparação note que a toxina madura codificada por TX05\_DIQUA apresenta apenas 30 resíduos e três pontes dissulfeto que formam o “nó de dissulfetos”. O número de resíduos nos intervalos inter-cisteínas (*loops*) também é diferente entre as duas cadeias, indicando que pertencem a subclasses diferentes de  $\omega$ -toxina-símile que podem ser reconhecidas pelas arquiteturas da interconectividade das cisteínas (Ver figura 29).

#### 4.4.4. Toxinas com domínio TIL

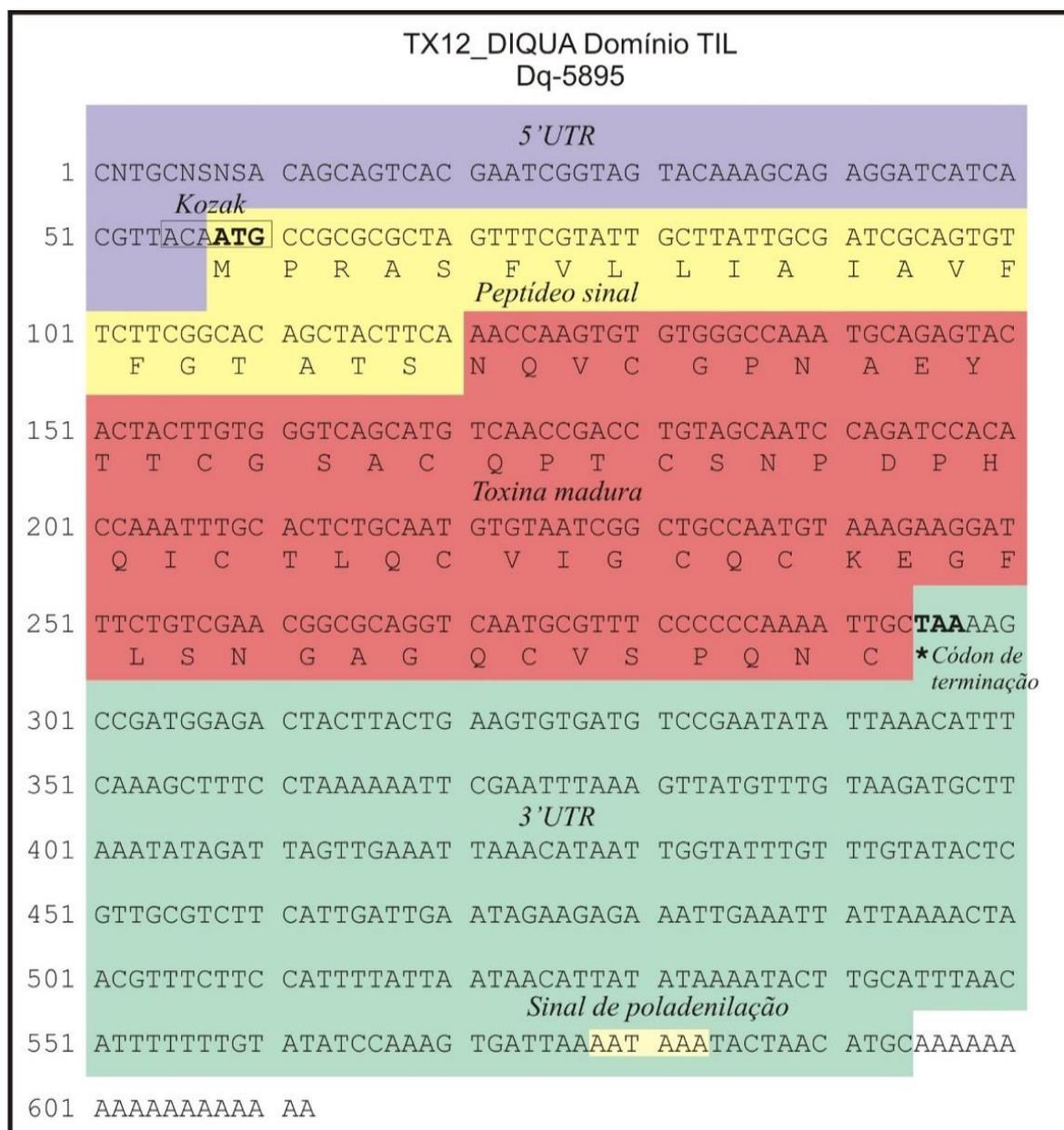
As toxinas com domínio TIL estão representadas pelo contig TX12\_DIQUA do agrupamento de sequências obtidas por Sanger. A organização do cDNA, representada na figura 30, apresenta uma região 5'UTR de 57pb com o códon de inicialização na posição 58. O peptídeo sinal tem início na posição 58 e se estende até a posição 120; é seguido pela região codificante da toxina madura que compreende 184pb e termina na posição 294. A região 3'UTR se estende entre os resíduos 295 e 594. A tradução do consenso do contig TX12\_DIQUA evidencia uma proteína precursora de 79 aminoácidos e peptídeo sinal com sítio de clivagem entre as posições 21 e 22 (Aminoácidos ATS-NQ), sendo a toxina madura representada por um segmento de 57 aminoácidos, ponto isoelétrico de 4.65 e massa de 5895Da, preditos. De fato, toxinas contendo domínio TIL geralmente consistem de 56-84 resíduos de aminoácidos (ZENG *et al.*, 2013). A TX12\_DIQUA apresenta homologia, inclusive a nível de sítio ativo, com inibidores de quimiotripsina presentes no veneno de outros himenópteros como as formigas *Acromyrmex echinator* (F4WD51) e *Nasonia vitripennis* (K7IM93), a abelha *Apis mellifera* (P56682) e a vespa *Pimpla hypochondriaca* (Q8T0W0) (Figura 31). Os resíduos de aminoácidos entre as cisteínas C5 e C6 são críticos para a função inibidora de protease e, portanto, interferem na ação e especificidade desses peptídeos (CHEN *et al.*, 2013; MICHEL *et al.*, 2012). Este precursor também produz alinhamentos significantes com proteínas não caracterizadas de *Solenopsis invicta* (E9IZF3) e *Harpegnathos saltator* (E2BE66).

TX12\_DIQUA possui domínio TIL (*Trypsin Inhibitor like cysteine rich*), portanto pertence à superfamília de peptídeos inibidores de tripsina (Pfam01826), e desta forma, se apresenta como inibidor de proteases. Este domínio, tipicamente, apresenta 10 resíduos de cisteínas que formam 5 pontes SS (1-7, 2-6, 3-5, 4-10 e 8-9) conservadas (BANIA *et al.*, 1999). Proteínas deste grupo foram primeiramente descritas em vermes dos gêneros *Ascaris sp.* e *Schistosoma sp.*, e estão organizadas em uma arquitetura complexa de múltiplos domínios com funções distintas. Zeng *et al.* (2013) demonstraram que existem extensos polimorfismos na organização genômica dos genes que codificam estes polipeptídeos. Este peptídeo, em *D. quadriceps* apresenta um peptídeo sinal e um único domínio TIL, conferindo-lhe características de potencial toxina. Ademais, o precursor de TX12\_DIQUA pode

apresentar-se como um potencial alérgeno, quando comparado com Api m 6 (P83563) de *Apis mellifera*, a qual também apresenta domínio TIL.

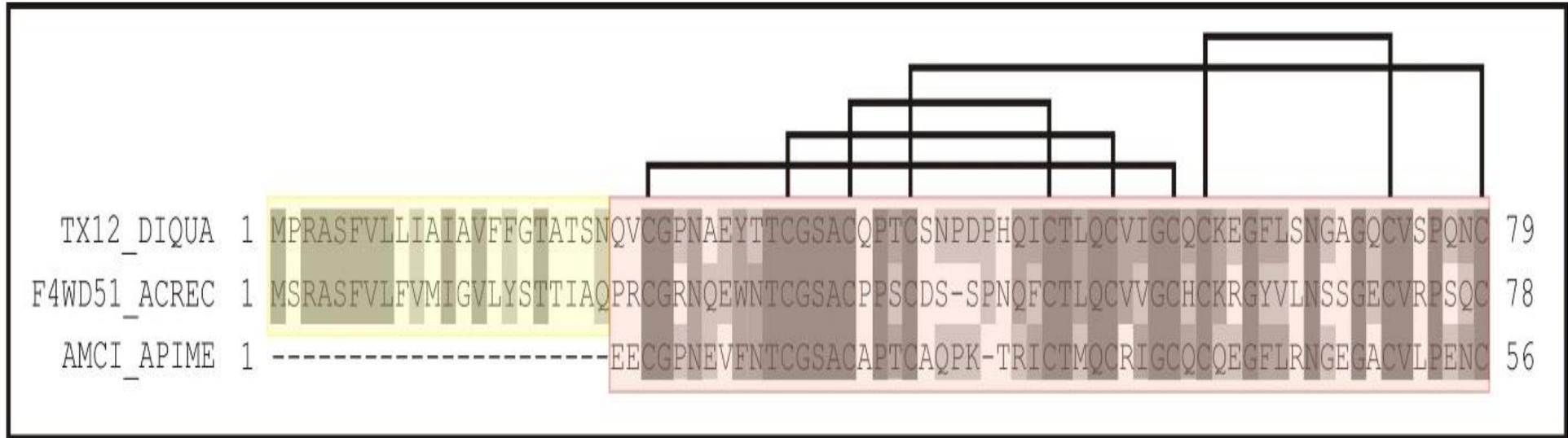
O interesse na caracterização destas proteínas reside não somente na compreensão do seu papel fisiológico mas na importância do seu envolvimento na resposta imune, na cascata de coagulação, sistema complemento, apoptose e ciclo celular, injúria tissular, progressão de tumores e bloqueio de canais iônicos (ZENG *et al.*, 2013; SHIGETOMI *et al.*, 2010; YANG *et al.*, 2013; RANASINGHE; MCMANUS, 2012; SALVADOR *et al.*, 2013; CHOO *et al.*, 2012). De fato, inibidores de proteases têm sido relatados por seu potencial terapêutico em cânceres, distúrbios cardiovasculares e inflamatórios (KENNEDY, 1998; TURK, 2006). As proteínas mais bem caracterizadas dos inibidores de serino proteases são os do tipo Kunitz, muito abundantes em diferentes tipos de organismos, incluindo anêmonas e moluscos marinhos, escorpiões, aranhas e himenópteros (MOURÃO; SHWARTZ, 2013). Estas apresentam uma ampla diversidade molecular e, além de sua função anticoagulante, têm sido propostos principalmente como inibidores de canais de potássio.

**Figura 30:** Organização do cDNA do contig TX12\_DIQUA.



5'UTR (lilás), Peptídeo sinal (amarelo), Proteína madura (vermelho), 3'UTR (verde), sinal de poliadenilação (caixa amarela). O peptídeo sinal foi predito utilizando a ferramenta SignalP.

**Figura 31:** Alinhamento do precursor codificado por TX12\_DIQUA com toxinas de venenos himenópteros com domínio TIL.



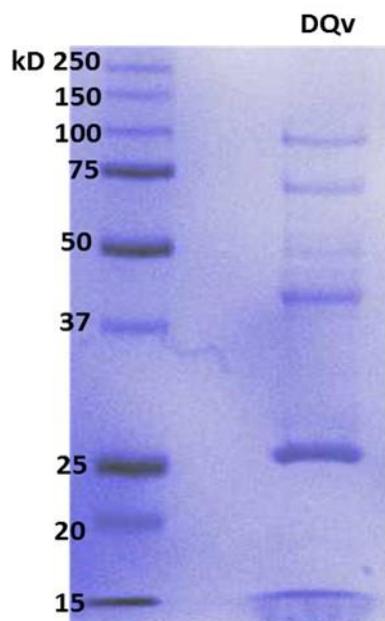
**TX12\_DIQUA:** precursor codificado pelo contig TX12\_DIQUA da biblioteca de cDNA da glândula de veneno de *D quadriceps*. **F4WD51\_ACREC:** Inibidor de quimi tripsina de *Acromyrmex echinator* (UNIPROT F4WD51). **AMCI\_APIME:** Inibidor de quimi tripsina de *Apis mellifera* (UNIPROT P56682). Peptídeo sinal está indicado em amarelo e a toxina madura em rosa. As pontes dissulfeto estão indicadas pelas linhas cheias.

## 4.5 Análise proteômica parcial do veneno de *D. quadriceps*

### 4.5.1 Extração, dosagem de proteína e perfil eletroforético SDS-PAGE da peçonha bruta de *D. quadriceps*

O veneno total de *D. quadriceps* foi extraído, entre 2010-2012, prontamente congelado, liofilizado e armazenado a -20°C. A dosagem de proteína foi realizada através do método Bradford (1976) e nos evidenciou que 64,4% do peso seco do veneno correspondem a conteúdo proteico. Utilizando eletroforese unidimensional (1-DE, SDS-PAGE), somente oito polipeptídeos maiores, com massa variando de 15-100KDa foram evidenciados, através da coloração com Coomassie Brilliant Blue (Figura 32), embora os sinais de massa demonstrem proteínas de baixo peso (112-2178 Da).

**Figura 32:** Perfil eletroforético do veneno total de *Dinoponera quadriceps* (DQv) em gel de eletroforese unidimensional SDS-PAGE.



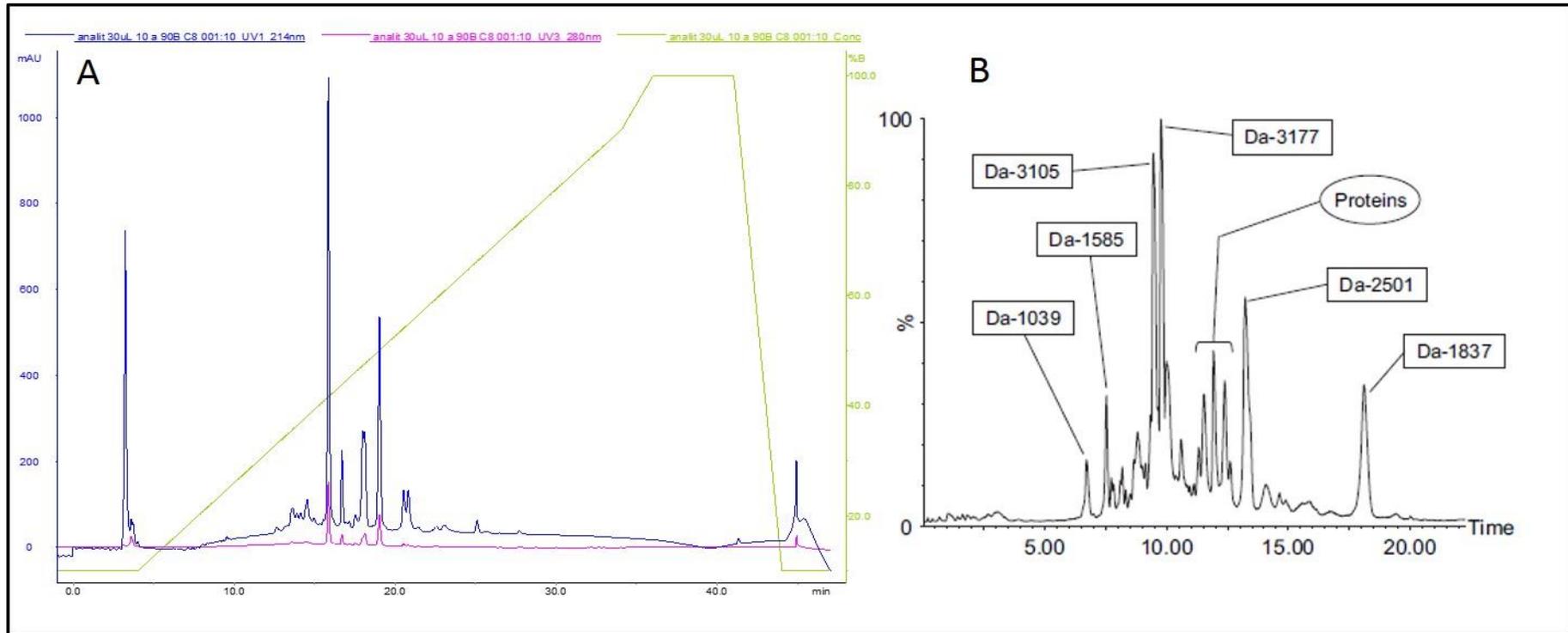
Visualizado com Coomassie Brilliant Blue R-250.

#### **4.5.2 Purificação da peçonha bruta de *D. quadriceps* por cromatografia em fase reversa (RP HPLC)**

Aplicou-se diferentes quantidades de veneno bruto (60µg, 100µg e 400µg) diluído em 10% da solução B (0,1% de ácido fórmico em acetonitrila) em coluna C-8 e C-18, com distintos gradientes e tempo, sendo o fluxo mantido a 1.0 mL/min. A avaliação do perfil de HPLC revelou a presença de compostos com diferentes perfis de hidrofobicidade e abundância (Figura 33 A). Na Figura 33 A foram empregadas as mesmas condições experimentais aplicadas para a separação da peçonha de *D. australis* (JOHNSON *et al.*, 2010) (Figura 33B), quais sejam: coluna C8, gradiente de 10-90% de solução B em 30min, com fluxo de eluição a 1 mL/min. Neste caso observamos que o veneno de *D. quadriceps* apresentou um perfil global semelhante ao de *D. australis*. No gráfico de *D. australis*, na figura 33 B representa a contagem iônica total (“*Total Ion Current*”-TIC) durante experimentos de LC-MS, que corresponde à intensidade somada de diversos sinais iônicos em um dado momento. Assim, mesmo a TIC correspondendo a sinais de íons detectados, podemos dizer que é uma representação semelhante a uma escala de absorbância em UV.

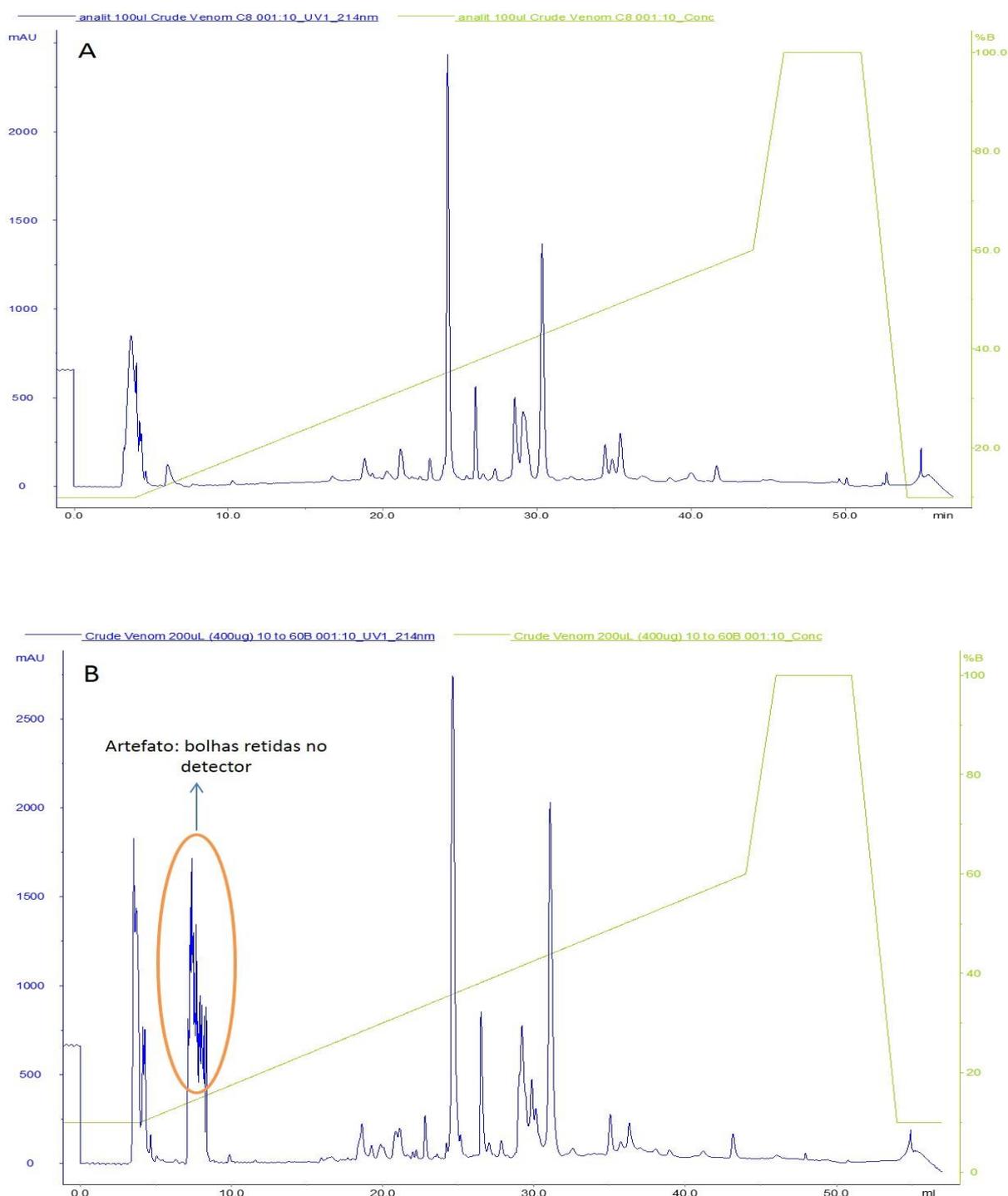
Embora o perfil obtido na purificação preliminar se mostrasse bastante semelhante ao reportado para o veneno de *D. australis*, detectou-se que a variação de 2,66% de B/min era abrupta e dificultava, operacionalmente, uma adequada separação e coleta (manual) dos componentes. Decidiu-se então, diminuir a variação do gradiente para 1,25% de B/min, utilizando uma gradiente linear de 10-60% de B em 40 min, tal como ilustrada na figura 34 A. De modo a comparar se uma coluna C-18, mais hidrofóbica, poderia melhorar as purificações, empregamos o mesmo gradiente acima descrito, gerando o perfil cromatográfico representado na figura 34 B. Conforme evidenciado, não houve mudanças significativas no perfil obtido, mostrando que a natureza dos componentes do veneno bruto de *D. quadriceps* não é influenciada por esses dois tipos de colunas cromatográficas.

**Figura 33:** Perfil de cromatográfico da peçonha bruta de *D. quadriceps* comparado com o perfil cromatográfico da peçonha bruta de *D. australis*. (A) Eluição em coluna C-8 (250x4.6mm I.D., 5µm, Vydac, EUA), através de gradiente linear 10-90% de B com fluxo de 1,0mL/min. (B) Perfil da contagem iônica total da peçonha bruta de *D. australis* (Johnson *et al.*, 2010), obtido nas mesma condições que as empregadas em A.



O eluato foi monitorado na faixa de absorvância do ultravioleta em comprimentos de onda de 214 e 280 nm (azul e rosa, respectivamente). A linha contínua em cor verde representa a variação do gradiente ao longo do tempo.

**Figura 34:** Perfil de cromatográfico da peçonha bruta de *D. quadriceps*. (A) Perfil cromatográfico de 200 µg peçonha bruta eluída em coluna C-8 (250x4.6mm I.D., 5µm, Vydac, EUA), através de gradiente linear 10-60% de B com fluxo de 1,0mL/min. (B) Perfil cromatográfico de 400 µg de peçonha bruta de eluída em coluna C-18, através de gradiente linear 10-60% de B com fluxo de 1,0 mL/min.

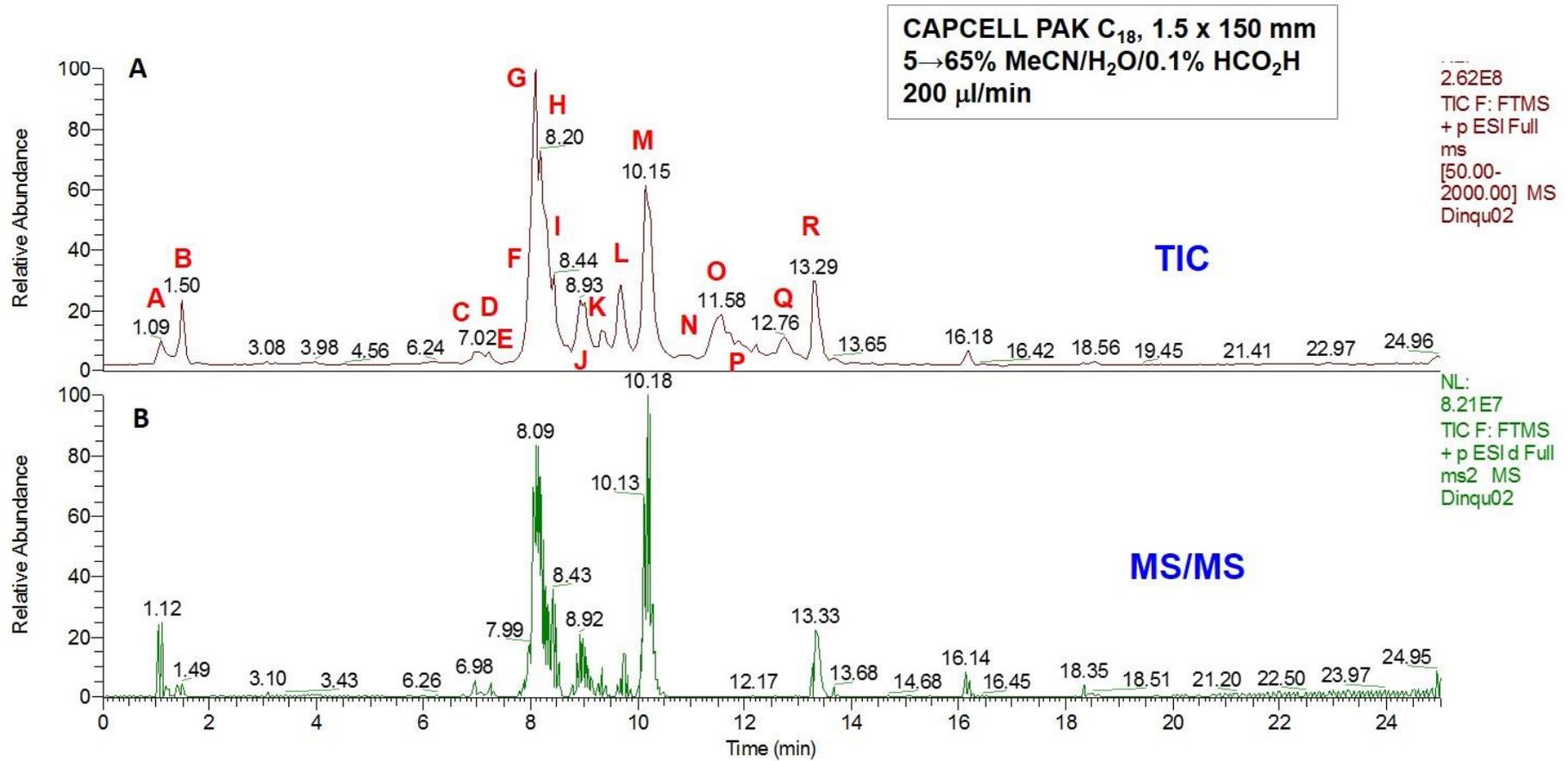


O eluato foi monitorado na faixa de absorvância do ultravioleta em comprimentos de onda de 214 nm (azul). A linha contínua em cor verde representa a variação do gradiente ao longo do tempo.

#### **4.5.3 Cromatografia líquida acoplada à espectrometria de massas (LC-MS)**

Utilizando técnica de cromatografia líquida acoplada à espectrometria de massas (LC-MS) obteve-se o perfil espectroscópico da contagem iônica total (TIC) e dos correspondentes espectros de fragmentação de íons precursores (MS/MS) dos íons detectados a partir da eluição cromatográfica, sendo os mesmos apresentados na figura 35 A e B, respectivamente. A figura 35 A representa o TIC, o qual corresponde à soma de diversos sinais iônicos em um dado momento, e evidencia os diversos compostos que diferem em tamanho, hidrofobicidade e abundância relativa na peçonha total de *D. quadriceps* (representados pelas letras de A-R). A figura 35 B representa a compilação das fragmentações dos íons precursores obtidos e representados na figura 35 A. No caso, os espectros MS/MS estão todos colapsados, sendo que a interpretação dos mesmos se deu por análise manual dos espectros de cada um dos íons obtidos e fragmentados (dados não mostrados – Dr. Katsuhiko Konno, Toyama University, Japão, comunicação pessoal).

**Figura 35:** Perfil de cromatografia líquida acoplada à espectrometria de massa da peçonha bruta de *D. quadiceps*. (A) Representação do TIC (Total Ion Current) obtido através de cromatografia de fase reversa utilizando coluna C-18, com um gradiente de 5-65% e um fluxo de 200µL/min. (B) Perfil de MS/MS observado a partir dos picos identificados em A.



O perfil de massas identificado e seus respectivos tempos de retenção, estão representados na tabela 5 onde podemos observar uma variação de massas de 112.086-3178.752 Da. Os compostos analisados até o presente e que foram sequenciados *de novo* (MS/MS) estão sinalizadas em vermelho. A figura 36 representa a estrutura de pequenas moléculas encontradas na peçonha, como histamina, prolina e dopamina, cujas estruturas foram determinadas. A presença de compostos de baixo peso molecular (aminas biogênicas, acetilcolina, carboidratos, lipídios, aminoácidos livres) em veneno de himenópteros é comumente relatada (CISZOWSKI; MIETKA-CISZOWSKA, 2007). A Tabela 6 apresenta as sequências dos peptídeos analisados, cujas estruturas foram determinadas por interpretação manual dos espectros de MS/MS. Avaliando-se estas sequências observa-se que a toxinas maduras representadas pelo contig TX01\_DIQUA e contig TX09\_DIQUA do sequenciamento por Sanger e pelo contig 01 e 02 do sequenciamento NGS, as Dinoponeratoxinas, estão representados nos resultados de espectrometria de massas como os peptídeos Dq-Ic (1320.745), Dq-Ka (2449.276), Dq-Lb (1504.857), Dq-Ma (2433.290) Dq-H (3162.756). Já os peptídeos Dq-Fa (986.603), Dq-Ia (872.542), Dq-Pa (1856.119) e Dq-Ra (1840.126), representam toxinas maduras evidenciadas através do contig TX06\_DIQUA (Pilosulina-símile) do sequenciamento por Sanger.

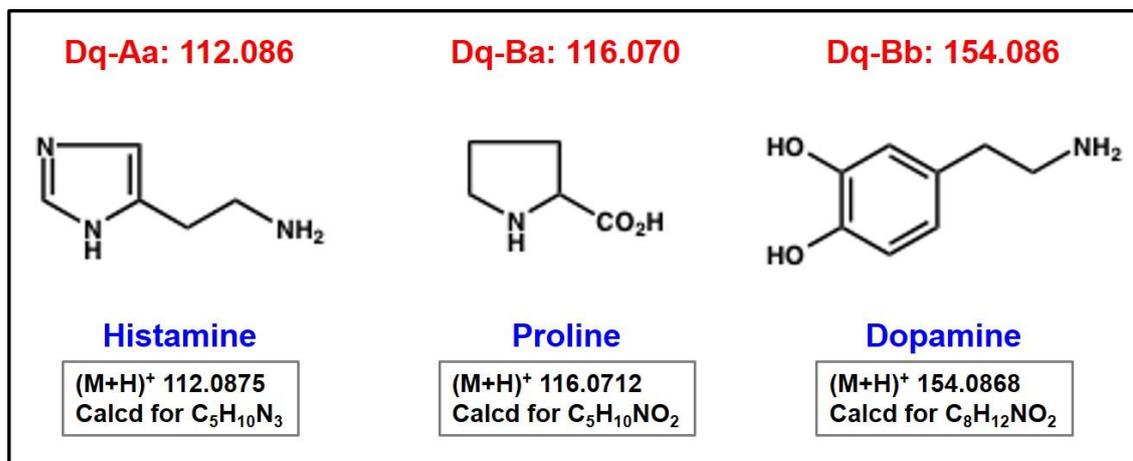
**Tabela 5:** Perfil de massas do veneno bruto de *D. quadriceps* utilizando espectrometria de massas electrospray (ESI).

Pico	TR	(M + H) <sup>+</sup>
A	1.03-1.29	112.086, 170.092
B	1.29-1.57	116.070, 154.086, 175.023, 193.033, 215.015, 257.160, 437.971
C	6.89-7.02	550.296, 665.322, 873.518, 1133.688, 3178.752
D	7.02-7.30	1002.596, 1133.688, 1187.647, 3178.752
E	7.45-7.62	343.052, 925.594, 1289.778, 1343.736
F	7.78-7.85	986.603, 1565.868, 3162.758
G	7.90-8.05	986.602, 3162.759
H	8.15-8.30	3162.756
I	8.30-8.83	872.592, 1031.547, 1320.738, 3162.757, 3197.689
J	8.83-9.26	3197.691
K	9.26-9.39	2449.276, 3197.688
L	9.53-9.79	1504.857, 2403.288, 2449.277, 2561.375
M	10.07-10.47	2433.284
N	10.53-11.09	1499.916, 1991.184, 2433.282, 2539.312, 3162.750
O	11.26-12.07	2433.272, 3162.753
P	12.14-12.36	1856.119, 1909.028, 2433.255
Q	12.60-13.08	2433.272, 3162.750
R	13.24-13.39	1840.126

Estruturas analisadas estão sinalizadas em vermelho

TR= tempo retenção

**Figura 36:** Moléculas de baixo peso molecular identificadas através de espectrometria de massas.



**Tabela 6:** Agrupamento dos peptídeos sequenciados através de espectrometria de massas.

Peptídeo	Massa	Sequência	Contig
Dq-Eb	925.594	FLPPLLRA-NH <sub>2</sub>	Não categorizado
Dq-Cd	1133.688	AHFLPPLLRA-NH <sub>2</sub>	
Dq-Ib	1031.547	GVLPPDLFR	Não categorizado
Dq-Dc	1187.647	GVLPPDLFR	
Dq-Ed	1343.736	GVGVLPPDLFR	
Dq-Ia	872.542	LVGALVSTL	TX06_DIQUA
Dq-Fa	986.603	LSLVPSLMK-NH <sub>2</sub>	
Dq-Ra	1840.126	LVGALVSTLLSLVPSLMK-NH <sub>2</sub>	
Dq-Pa	1856.119	LVGALVSTLLSLVPSLMK-NH <sub>2</sub>	
Dq-Ic	1320.745	FGWTLAKWALK	TX01_DIQUA
Dq-Lb	1504.857	FGWTLAKWALKAL	
Dq-Ma	2433.290	FGWTLAKWALKALPAAMGMKQN	
Dq-Ka	2449.276	FGWTLAKWALKALPAAMGMKQN	
Dq-Ld	2561.375	FGWTLAKWALKALPAAMGMKQNK	
Dq-H	3162.756	GLKDWNNKHKDKIVKVVKEMGKAGINAA-NH <sub>2</sub>	TX09_DIQUA

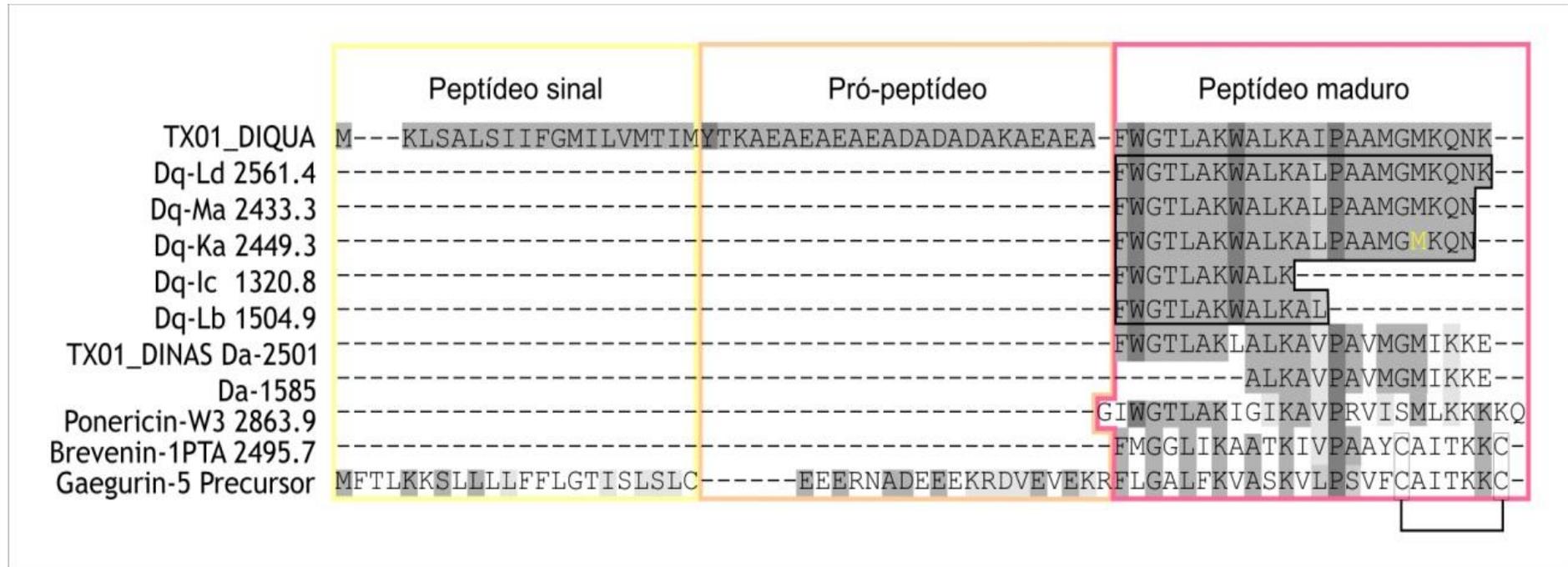
**M:** metionina oxidada; **L = L** (leucina) ou **I** (isoleucina); **NH<sub>2</sub>** denota porção C-terminal amidada

Cinco peptídeos com sequência determinada por MS/MS *de novo*, identificados na peçonha de *D. quadriceps*, derivam do precursor codificado pelo transcrito do Contig TX01\_DIQUA. A sequência madura do precursor encontra um correspondente idêntico ao peptídeo Dq-Ld com massa determinada de 2561,375 Da. O peptídeo Dq-Lb é menor com uma massa de 1504,857 Da e tem hidrofobicidade semelhante ao peptídeo Dq-Ld. O menor peptídeo Dq-Lc, com massa 1320,475 Da é o mais hidrofílico do grupo. Os peptídeos Dq-Ka e Dq-Ma diferem apenas pela oxidação da cadeia lateral do resíduo de metionina indicado em negrito (Tabela 6). Esta modificação é responsável pela diferença nas massas moleculares dos peptídeos Dq-Ka com 2449,276 Da e Dq-Ma com 2433,290 e pela diferença de hidropaticidade, sendo que o peptídeo Dq-Ma é mais hidrofóbico em comparação com o Dq-Ka.

Como os aminoácidos isoleucina e leucina são isômeros, com a mesma massa molecular, na presente análise por espectrometria de massas não é possível diferenciar entre ambos, assim, por comparação com os dados de sequenciamento por Sanger podemos afirmar que os resíduos de leucina na posição 13 dos peptídeos Dq-Ld, Dq-Ma e Dq-Ka, são, em verdade, resíduos de isoleucina, tal como determinado pelo códon correspondente à sequência nucleotídica (Figura 37).

O peptídeo Dq-Ld 2561,375 da peçonha de *D. quadriceps* tem 84% de similaridade com uma dinoponeratoxina TX01\_DINAS (UNIPROT P0CF01) encontrada no veneno de *D. australis* (JOHNSON *et al.*, 2010) (Figura 37). Esse peptídeo Da-2501 pode ser clivado em uma cadeia menor denominada Da-1585. Esta classe de peptídeos similares encontrados no veneno de *D. quadriceps* e *D. australis* são, por sua vez, similares ao peptídeo inseticida Ponericin W3, encontrado no veneno da formiga *Pachycondyla goeldii*, capaz de causar citólise e hemólise, além de apresentar atividade antimicrobiana sobre leveduras e bactérias gram-negativas (ORIVEL *et al.*, 2001). Uma vez que a Ponericin W3 foi descrita anteriormente, podemos denominar essa classe de dinoponeratoxinas como ponericina-símiles. Entretanto, o peptídeo Dq-Ld apresenta 50% de identidade com este peptídeo.

**Figura 37:** Alinhamento do precursor codificado pelo contig TX01\_DIQUA com peptídeos cuja sequência foi determinada pela análise proteômica e sequenciamento de novo MS/MS.



**TX01\_DIQUA:** Precursor de dinoponeratoxina-1; Peptídeos da peçonha de *D. quadriceps* (Dq) **Dq-Ld**; **Dq-Ma**; **Dq-Ka**; **Dq-lc**; **Dq-Lb** com a respectivas massa moleculares, estão agrupados por uma contorno negro e denominadas de acordo com o pico correspondente ao indicado no perfil TIC. **TX01\_DINAS Da 2501** corresponde a uma sequência similar identificada na peçonha de *D. australis* (UNIPROT P0CF01), a qual é clivada em uma cadeia menor correspondente ao peptídeo Da-1585. **Ponericin W3** da peçonha de *Pachycondyla Goeldi* (UNIPROT P82425). **Brevenin 1PTA** da secreção epidérmica da rã *Rana picturata* (UNIPROT P0C8T1). **Gaegurin-5** precursor da secreção epidérmica da rã *Glandirana rugosa* (UNIPROT P80399). Uma ponte dissulfeto presente nas duas últimas sequências, está indicado na figura, por uma caixa conectada que indica o padrão de ligação entre as duas cisteínas. As regiões correspondentes ao peptídeo sinal, pró-peptídeo e peptídeo maduro estão indicadas por caixas respectivamente coloridas em amarelo, laranja e vermelho. A metionina oxidada está indicada em amarelo.

O peptídeo classificado como Dq-H 3162,756 deriva do precursor codificado pelo contig TX09\_DIQUA, outro grupo de dinoponeratoxinas, ponericina-símile (Figura 38). O peptídeo Dq-H não apresenta os dois últimos resíduos de aminoácidos evidenciados no precursor e possui na região C-terminal uma amidação, o que deve ser resultante de uma modificação pós-traducional (Figura 38 e Tabela 6). O peptídeo Dq-H apresentou similaridade com a dinoponeratoxina TX77\_DINAS (UNIPROT P0CF05.1) e TX05\_DINAS (UNIPROT P0CF02.1) também identificados na peçonha de *D. australis*. Estes peptídeos Da-3177 e Da-3105, bem como Dq-H 3162, são similares às Ponericin G2 e G3 da peçonha de *Pachycondyla goeldii*. Cologna *et al.* (2013) em estudo proteômico com a peçonha de *D. quadriceps* coletadas de distintas áreas no estado da Bahia, identificou peptídeo muito semelhantes, Dq 3162 e Dq 3104. Ao compararmos com a sequência do precursor TX09\_DIQUA e com o peptídeo Ld-H, observamos que diferem no resíduo de aminoácido da posição 15 (Dq-H 3162 → K; lisina e Dq 3162 → E; Ácido glutâmico). Ademais, conforme ressaltado anteriormente, a leucina na posição 25 do peptídeo maduro Dq 3162, é, possivelmente, uma isoleucina, de acordo com o precursor TX09\_DIQUA.

**Figura 38:** Alinhamento do precursor codificado pelo contig TX09\_DIQUA com peptídeos cuja sequência foi determinada pela análise proteômica e sequenciamento de novo MS/MS.

	Peptídeo sinal	Pró-peptídeo	Peptídeo maduro	
TX09_DIQUA	MKLSAFTLAFALILMMAIMYNMAEA	AALADADADAEAI	GLKDWWNKHKDKIVKVVKEMGKAGINAAGK	69
Dq-H 3162	-----	-----	GLKDWWNKHKDKIVKVVKEMGKAGINAA--	28
Dq 3162	-----	-----	GLKDWWNKHKDKIVEVVKEMGKAGLNAA--	28
Dq 3104	-----	-----	GLKDWWNKHKDKIVEVVKDSGKAGLNAA--	28
TX77_DINAS Da 3177	-----	-----	GLKDWWNKHKDKIIDVVKEMGKAGLQAA--	28
TX05_DINAS Da 3105	-----	-----	GLKDWWNKHKDKIIAVAKEMGKAGLQAA--	28
PonericinG3 3381	-----	-----	GWKDWLNKGKEWLKKKGGPGIMKAALKAATQ	30
PonericinG2 3006	-----	-----	GWKDWLKKGKEWLKAKGGPGIVKAALQAAATQ	30

**TX09\_DIQUA:** Precursor de dinoponeratoxina; Peptídeo da peçonha de *D. quadriceps* (Dq) **Dq-H**; com a respectiva massa molecular, está destacado por um contorno negro e denominadas de acordo com o pico correspondente ao indicado no perfil TIC. **TX77\_DINAS Da 3177** corresponde a uma sequência similar identificada na peçonha de *D. australis* (UNIPROT P0CF05.1), e **TX05\_DINAS Da-3105** (UNIPROT P0CF02.1). **Dq 3162** e **Dq 3104** da peçonha de *D. quadriceps* do estado da Bahia. **Ponericin G3** e **G2** da peçonha de *Pachycondyla Goeldi* (UNIPROT P82416 e P82415, respectivamente). As regiões correspondentes ao peptídeo sinal, pró-peptídeo e peptídeo maduro estão indicadas por caixas respectivamente coloridas em amarelo, laranja e vermelho.

De certa forma, como relatado por Johnson *et al* (2010), estes peptídeos apresentam semelhanças com peptídeos bioativos da pele de anfíbios pertencente à subfamília das Brevininas (vide na figura 30 a Bevinin-1PTA e o precursor de Gaegurin-5) que caracteristicamente também apresentam atividade antimicrobiana. Estes peptídeos de anfíbio têm menor similaridade e apresentam uma ponte dissulfeto determinada para a Bravinin-1PTA (CONLON *et al.*, 2008) (Figura 36). Este peptídeo apresenta atividade antimicrobiana sobre *S. aureus* (Gram positivo) e *E. coli* (Gram negativo). Já o peptídeo maduro de Gaegurin-5 apresenta amplo espectro de atividade sobre bactérias, fungos e protozoários, sem apresentar ação hemolítica (PARK *et al.*, 1994). De maneira interessante, a sequência do precursor do peptídeo de Gaegurin-5 apresenta uma organização similar aos precursores codificados pelos contigs TX01\_DIQUA e TX09\_DIQUA, em especial quanto à presença de um peptídeo sinal de secreção, seguido por um pró-peptídeo rico em resíduos ácidos que é excisado para liberar o peptídeo maduro. Ainda que a homologia relativa seja mais baixa, a carga líquida e a presença alternada de resíduos hidrofóbicos sugere a formação de uma estrutura em  $\alpha$ -hélice anfipática catiônica similar tanto nos peptídeos maduros das formigas quanto nas brevininas de rãs, indicando uma possível similaridade na atividade biológica.

Outro grupo de peptídeos caracterizados por espectrometria de massas compõem o codificado pelo contig TX06\_DIQUA, classificados como pilosulina-símile. Observa-se que esta toxina apresenta diferentes produtos de clivagem, os quais foram identificados como Dq-Ia, Dq-Fa, Dq-Ra e Dq-Pa, todos com amidação C-terminal. O peptídeo Dq-Ia é o menor, com uma massa de 872,542. Dq-Ra 1840,126 e Dq-Pa 1856,119 diferem pela oxidação da cadeia lateral do resíduo de metionina, modificação responsável pela diferença de massas. Outro ponto interessante é que os peptídeos Dq-Fa, Dq-Ra e Dq-Pa apresentam amidação como modificação pós-traducional (vide Tabela 6, Figura 39). Cologna *et al.* (2013) descreveram peptídeo semelhante na peçonha de *D. quadriceps* coletadas de distintas áreas no estado da Bahia. Todavia esses peptídeos, classificados como dermaseptina-símile, são compostos por 17 aminoácidos e apresentam resíduo de lisina na posição 3 do peptídeo maduro, discordando dos peptídeos evidenciados neste trabalho, os quais são compostos por 18 aminoácidos e apresentam um resíduo de glicina na posição 3 do peptídeo maduro, e cujos dados podem ser confirmados pela sequência do

precursor do Contig TX06\_DIQUA, obtido por sequenciamento por Sanger. O peptídeo Dq-Ra apresenta 44% de identidade com TX37\_DINAS Da 1837 (UNIPROT P0CF03) encontrada na peçonha de *D. australis* (JOHNSON *et al.*, 2010). Este foi descrito por compartilhar semelhanças com a Poneratoxin (UNIPROT P41736) da peçonha de *Paraponera clavata* (PIEK *et al.*, 1990), funcionalmente descrito com neurotoxina e exerce sua função sobre canais de Na<sup>+</sup> (DUVAL *et al.*, 1992; SZOLAJSKA *et al.*, 2004).

Como relatado anteriormente, o precursor TX06\_DIQUA foi classificado como pilosulina-símile, tendo em vista apresentar homologia na região do peptídeo sinal e do pró-peptídeo, apresentando as toxinas maduras muitas diferenças. A avaliação dos dados proteômicos evidencia que esses peptídeos apresentam algum grau de homologia com Poneratoxina (aproximadamente 12%) e com peptídeos da subfamília Dermaseptin (*Alyteserin2c* de rãs do gênero *Alytes*, *Phylloseptin* de rãs do gênero *Phyllomedusa*), reportados como peptídeos antimicrobianos (SUBASINGHAGE *et al.*, 2011; THOMPSON *et al.*, 2007).

**Figura 39:** Alinhamento do precursor codificado pelo contig TX06\_DIQUA com peptídeos cuja sequência foi determinada pela análise proteômica e sequenciamento de novo MS/MS.

	Peptídeo sinal	Pró-peptídeo	Peptídeo maduro	
TX06_DIQUA	MKLSALS	LVFGMILMMTIMYTNAEAEATAEADADAEAEAEAEAEAL	LVGALVSTLLSLVPSLMKG	62
Dq-Ia 872			LVGALVSTL	
Dq-Fa 986			LSLVPSLMK	9
Dq-Ra 1840			LVGALVSTLLSLVPSLMK	18
Dq-Pa 1856			LVGALVSTLLSLVPSLMK	18
Dq 1840			LVK-LVSTLLSLVPSLMK	17
Dq 1856			LVK-LVSTLLSLVPSLMK	17
Dq 1897			LVK-LVSTLLSLVPSLMKG	18
TX37_DINAS Da 1837			FLGGLIGPLMSLIPGLLK	18

**TX06\_DIQUA:** Precursor Pulosulina-símile; Peptídeo da peçonha de *D. quadriceps* (Dq) **Dq-Ia; Dq-Fa; Dq-Ra; Dq-Pa** com as respectivas massas moleculares, estão destacados por um contorno negro e denominadas de acordo com o pico correspondente ao indicado no perfil TIC. Dq 1840; Dq 1856; Dq 1897 da peçonha de *D. quadriceps* do estado da Bahia. **TX37\_DINAS Da 1837** corresponde a uma sequência similar identificada na peçonha de *D. australis* (UNIPROT P0CF03). As regiões correspondentes ao peptídeo sinal, pró-peptídeo e peptídeo maduro estão indicadas por caixas respectivamente coloridas em amarelo, laranja e vermelho. Metionina oxidada está indicada em amarelo.

Como podemos observar muitos destes peptídeos foram reportados por sua ação antimicrobiana, todavia estas moléculas vêm sendo avaliadas em outros sistemas biológicos e portanto novas funções vêm sendo relatadas nos últimos anos. Por exemplo *Ponericin W3* de *Pachycondyla goeldii* (ORIVEL *et al.*, 2001) e *Poneratoxin* de *Paraponera clavata* (SZOLAJSKA *et al.*, 2004), têm sido postulados como inseticidas. Com vistas ao emprego de novas funções a Pílosulina 2 também foi relatada por sua propriedade anti-hipertensiva (DONOVAN; BALDO, 1997). Além disso, a remodelação molecular destes peptídeos também tem reduzido a sua toxicidade e mantido e/ou melhorado seu efeito farmacológico, como reportado por Inagaki *et al.* e Davie *et al.* (2004) para o caso das pilosulinas. Desta forma, está claro que esses polipeptídeos apresentam-se como um leque de novas possibilidades funcionais que necessitam ser melhor exploradas.

## 5 CONSIDERAÇÕES FINAIS

Através da combinação do sequenciamento por Sanger da biblioteca de cDNA e de sequenciamento NGS, uma descrição completa dos transcritos da glândula de veneno foi alcançada, revelando sequências que são expressas para a manutenção do equilíbrio fisiológico do tecido e produção de toxinas. Os polipeptídeos que compõem a peçonha compreendem tanto substâncias abundantes como componentes raros, envolvidos em reações locais e sistêmicas durante envenenamento. O grupo de toxinas mais abundantes é composto por alérgenos, fosfolipases e dinoponeratoxinas. Estas podem interferir na homeostase e sinalização lipídica, no transporte de vesículas e na interação proteína-proteína e proteína-ácido nucleico, interferindo no mecanismos bioquímicos do tecido da presa ou vítima. Dinoponeratoxinas representam o conjunto toxinas mais abundantes e, até o presente, exclusivamente expressos por formigas gigantes. Este grupo de pequenas toxinas, que não apresentam pontes dissulfeto encontra sequências homólogas estruturalmente na peçonha de *D. australis* e, a nível nucleotídico, compartilham similaridade com pilosulinas – peptídeos antimicrobianos e degranuladores de mastócitos encontrados em formigas do gênero *Myrmecia*. Outro grupo, menos abundante, porém não menos importante, compreende os polipeptídeos ricos em cisteínas, estruturados por pontes dissulfeto em um arranjo típico de IKCs. Estes apresentam identidade estrutural com peptídeos de escorpião, aranhas e animais marinhos como os gastrópodos de gênero *Conus*. A variedade estrutural e farmacológica de peptídeos ricos em pontes dissulfeto têm impressionado. Estas toxinas apresentam funções que variam desde bloqueadores de canais iônicos, de receptores neuronais, até ação antimicrobiana (FAN *et al.*, 2011; SAMY *et al.*, 2012). Dentre as toxinas menos abundantes também se encontram os inibidores de proteases, substâncias capazes de interferir em diferentes mecanismos patofisiológicos como a resposta imune, a cascata de coagulação e o desenvolvimento de cânceres, como descrito anteriormente.

No âmbito do desenvolvimento de novas drogas, os animais peçonhentos, desempenham um papel primordial, já que apresentam uma vasta fonte de moléculas farmacologicamente ativas. Como bem pronunciado por Escoubas e King (2009) “... o enorme potencial dos venenos como fontes de novas drogas para o futuro torna-se

mais que realidade a cada dia, como com as investigações sobre os peptídeos dos venenos que oferecem quase que infinitas possibilidades para a descoberta de novas drogas.” Aliado à isto, as últimas décadas trazem progressos significativos na caracterização de receptores, de canais iônicos e outras classes de moléculas, resultado em diferentes novos alvos para os quais se pode desenvolver drogas seletivas. Neste sentido, peptídeos de venenos são uma fonte particular de modelos, através dos quais novos agentes terapêuticos podem ser desenvolvidos (Lewis; Garcia, 2003). São pequenas moléculas, de síntese relativamente fácil, com arranjos diversos, potentes e estáveis, o que as tornam interessantes do ponto de vista farmacológico.

## 6 CONCLUSÃO

Neste estudo, nós sequenciamos e caracterizamos o transcriptoma da glândula de veneno de *D. quadriceps*, bem como procedemos a análise proteômica parcial da peçonha, tendo como foco principal as toxinas polipeptídicas que não somente podem estar envolvidas no surgimento dos sintomas na ocasião da inoculação da peçonha em humanos, mas, principalmente, por apresentarem-se como novos modelos funcionais a serem utilizados como ferramentas biotecnológicas.

Assim, com o presente trabalho, apresentamos à comunidade científica, toxinas polipeptídicas da peçonha de *D. quadriceps*, como novas ferramentas com potencial farmacológico e/ou biotecnológico. Além disso diferentes precursores com função desconhecida ou nunca descritos foram evidenciados, o que representa novas possibilidades de funções biológicas. Está claro, então, que esforços futuros devem ser empregados do sentido de melhor caracterizar funcionalmente estas moléculas.

## 7 REFERÊNCIAS

ALTSCHUL, SF; MADDEN, TL; SCHAFFER, AA; ZHANG, J; ZHANG, Z; MILLER, W; LIPMAN, DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucl Acids Res**, v25, p3389–3402, 1997.

AOKI, J; NAGAI, Y; HOSONO, H; INOUE, K; ARAI, H. Structure and function of phosphatidylserine-specific phospholipase A1. **Bioch Biophys Acta**, v1582, p26–32, 2002.

ARAÚJO, CZD; JAISSON, P. Modes de fondation des colonies chez La fourmi sans reine *Dinoponera quadricaps* Santschi (Hymenoptera, Formicidae, Ponerinae). **Actes de Colloques Insectes Sociaux**, v9, p79-88, 1994.

ATTYGALLE, AB; MORGAN, ED. Chemicals from the glands of ants. **Chemical Society Review**, v13, p245-278, 1984.

BADR, G; GARRAUD, O; DAGHESTANI, M; AL-KHALIFA, MS; RICHARD, Y. Human breast carcinoma cells are induced to apoptosis by samsum ant venom through an IGF-1-dependant pathway, PI3K/AKT and ERK signaling. **Cel Immunol**, v273, p10-16, 2012.

BANIA, J.; STACHOWIAK, D.; POLANOWSKI, A. Primary structure and properties of the cathepsin G/chymotrypsin inhibitor from the larval hemolymph of *Apis mellifera*. **Eur J Biochem**, v262 p680-687, 1999.

BATISTA, IFC; CHUDZINSKI-TAVASSI, AM; FARIA, F; SIMONS, SM; BARROS-BATESTTI, DM; LABRUNA, MB; LEÃO, LI; HO, PL; JUNQUEIRA-DE-AZEVEDO, ILM. Expressed sequence tags (ESTs) from the salivary glands of the tick *Amblyomma cajennense* (Acari: Ixodidae). **Toxicon**, v51, p823-834, 2008.

BERTHOLD, N; CZIHAL, P; FRITSCHES, SU; SCHIFFER, G; KNAPPE, D; ALBER, G; HOFFMANN, R. Novel apidaecin 1b analogs with superior serum stabilities for treatment of infections by gram-negative pathogens. **Antimicrob Agents Chemother**, v57, p402-409, 2013.

BLANK, S; SEISMANN, H; BOCKISCH, B; BRAREN, I; BREDEHORST, R; GRUNWALD, T; OLLERT M; SPILLNER, E. Identification and recombinant expression of a novel IgE-reactive 70 kDa carboxylesterase from *Apis mellifera* venom. **EMBL/GenBank/DDBJ databases**, 2008.

BLUM, M. S. **Biochemical defenses of insects**. In: ROCKSTEIN, M. Biochemistry of insects. London: Academic Press, p465-513, 1978.

BOHLEN, CJ; PRIEL, A; ZHOU, S; KING, D; SIEMENS, J; JULIUS, D. A bivalent tarantula toxin activates the capsaicin receptor, TRPV1, by targeting the outer pore domain. **Cell**, v141, p834-845, 2010.

BONASIO, R; ZHANG, G; YE, C; MUTTI, NS; FANG, X; et al. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. **Science**, v329, p1068-1071, 2010.

BORER, AS; WASSMANN, P; SCHMIDT, M; HOFFMAN, DR; ZHOU, JJ; WRIGHT, C; SCHIRMER, T; MARKOVIC-HOUSLEY, Z. Crystal structure of Sol I 2: a major allergen from fire ant venom. **J Mol Biol**, v415, p635-648, 2012.

BOUZID, W; KLOPP, C; VERDENAUD, M; DUCANCEL, F; VÉTILLARD, A. Profiling the venom gland transcriptome of *Tetramorium bicarinatum* (Hymenoptera: Formicidae): The first transcriptome analysis of an ant species. **Toxicon**, v70, p70-81, 2013.

BRADFORD, MM. A rapid and sensitive for the quatitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Anal Biochem**, v72, p248-254, 1976.

BULET, P HETRU, C; DIMARCQ, JL; HOFFMANN, D. Antimicrobial peptides in insects; structure and function. **Dev Comp Immunol**, v23, p329-344, 1999.

BUSCHINGER, A; MASCHWITZ, U. Defensive behavior and defensive mechanisms in ants. In: *Defensive mechanisms in social insects* (H.R. Hermann - Ed.), Praeger, New York; p95-150, 1984.

CAETANO, F. H.; JAFFÉ, K; ZARA, F. J. Formigas: Biologia e Anatomia. Rio Claro: FHC, 42p, 2002.

CAPPARELLI, R; DE CHIARA, F; NOCERINO, N; MONTELLA, RC; IANNACCONE, M; FULGIONE, A; ROMANELLI, A; AVITABILE, C; BLAIOTTA, G; CAPUANO, F. New perspectives for natural antimicrobial peptides: application as antinflammatory d rugs in a murine model. **BMC Immunol**, v17, p13-61, 2012.

CASCALES, L; HENRIQUES, ST; KERR, MC; HUANG, YH; SWEET, MJ; DALY, NL; CRAIK, DJ. Identification and characterization of a new family of cell-penetrating peptides: cyclic cell-penetrating peptides. **J Biol Chem**, v286, p36932-43, 2011.

CHAIM, OM; TREVISAN-SILVA, D; CHAVES-MOREIRA, D; WILLE, AC; FERRER, VP; MATSUBARA, FH; MANGILI, OC; DA SILVEIRA, RB; GREMSKI, LH; GREMSKI, W; SENFF-RIBEIRO, A; VEIGA, SS. Brown spider (*Loxosceles* genus) venom toxins: tools for biological purposes. **Toxins**, v3, p309-344, 2011.

CHEN, Z; LUO, F; FENG, J; YANG, W; ZENG, D; ZHAO, R; CAO, Z; LIU, M; LI, W; JIANG, L; WU, Y. Genomic and structural characterization of Kunitz type peptide LmKTT1a highlights diversity and evolution ofscorpion potassium chann el toxins. **PLoS One**, v8, p4, 2013.

CHERNIAK, EP. Bugs as drugs, Part 1: Insects: the new alternative medicine for the 21<sup>st</sup> century? **Altern Med Rev**, v15, p124-135, 2010.

CHIANURA, L; POZZI, F. Case Report: A 40-year-old man with ulcerated skin lesions caused by bites of safari ants. **Am J Trop Med Hyg**, v83, p9, 2010.

CHOO, YM; LEE, KS; YOON, HJ; QIU, Y; WAN, H; SOHN, MR; SOHN, HD; JIN, BR. Antifibrinolytic role of a bee venom serine protease inhibitor that acts as a plasmin inhibitor. **PLoS One**, v7, p2, 2012.

CHUDZINSKI-TAVASSI, AM; SCHATTNER, M; FRITZEN, M; POZNER, RG; REIS, CV; LOURENÇO, D; LAZZARI, MA. Effects of Lopap on human endothelial cells and platelets. **Haemostasis**, v31, p257-265, 2001.

CISZOWSKI, K; MIETKA-CISZOWSKA, A. Hymenoptera stings. **Przegl Lek**, v64, p282-289, 2007.

CLARDY, J; WALSH, C. Lessons from natural molecules. **Nature**, v432, p829-837, 2004.

COLOGNA, CT; BARBOSA, DB; SANTANA, FA; RODOVALHO, CM; OLIVEIRA, LA; BRANDEBURGO, MAM. Estudo da peçonha de *Dinoponera australis* – Roger, 1861 (Hymenoptera, Ponerinae). In: V Encontro interno de Iniciação científica, Convênio CNPq/UFU, Uberlândia, 2005.

COLOGNA, CT; CARDOSO, JS; JOURDAN, E; DEGUELDRE, M; UPERT, G; et al. Peptidomic comparison and characterization of the major components of the venom of the giant ant *Dinoponera quadriceps* collected in four different area of Brazil. **Proteomics**, v94, p413-422, 2013.

CONCEIÇÃO, K; KONNO, K; RICHARDSON, M; ANTONIAZZI, MM; JARED, C; DAFFRES, S; CAMARGO, ACM; PIMENTA, DC. Isolation and biochemical characterization of peptides presenting antimicrobial activity from skin of *Phyllomedusa hypocondrialis*. **Peptides**, v27, p3092-3099, 2006.

CONCEIÇÃO, K; KONNO, K; DE MELO, RL; ANTONIAZZI, MM; JARED, C; SCIANI, JM; CONCEIÇÃO, IM; PREZOTO, BC; CAMARGO, ACM; PIMENTA, DC. Isolation and characterization of a novel bradlinin potentiating peptide (BPP) from the skin secretion of *Phyllomedusa hypocondrialis*. **Peptides**, v28, p515-523, 2007.

CONLON, JM; KOLODZIEJEK, J; NOWOTNY, N; LEPRINCE, J; VAUDRY, H; COQUET, L; JOUENNE, T; KING, JD. Characterization of antimicrobial peptides from the skin secretions of the Malaysian frogs, *Odorrana hosii* and *Hylarana picturata* (Anura: Ranidae). **Toxicon**, v52, p465-473, 2008.

CRICK, FH; WATSON, JD. The structure of DNA. **Cold Spring Symp Quant Biol**, v18, p123-131, 1953.

CRUZ LOPEZ, L; MORGAN, ED. Explanation of bitter taste of venom of ponerine ant *Pachycondyla apicalis*. **J Chem Ecol**, v23, p705-712, 1997.

CUI, F; LIN, Z; WANG, H; LIU, S; CHANG, H; REECK, G; QIAO, C; RAYMOND, M; KANG, L. Two single mutations commonly cause qualitative change of nonspecific carboxylesterases in insects. **Insect Biochem Mol Biol**, v41, p1-8, 2011.

CZAIKOSKI, PAULA, G; MENALDO, DANILO L; MARCUSSI, SILVANA; BASEGGIO, ANNE LC; FULY, ANDRÉ L; PAULA, RAFAEL C; QUADROS, ANDREZA U; ROMÃO, PEDRO RT; BUSCHINI, MARIA LT; CUNHA, FERNANDO Q; SOARES, ANDREIMAR M; MONTEIRO, MARTA C. Anticoagulant and fibrinolytic properties of the venom of *Polybia occidentalis* social wasp. **Blood Coagul Fibrinolysis**, v21, p653-659, 2010.

DA SILVA, N.P. Noções básicas de biologia molecular. (In) *Biologia Molecular: Guia Prático*. EÇA, L.P. e col. Ed. Revinter, 2004.

DAVIES, NW; WIESE, MD; BROWN, SGA. Characterisation of major peptides in 'jack jumper' ant venom by mass spectrometry. **Toxicon**, v43, p173-183, 2004.

DETWILER, NM; HAMP, TJ; KAZIM, AL. DNA sequencing using the liquid polymer POP-7 on an ABI PRISM 3100 genetic analyzer. **Biotechniques**, v36, p932-933, 2004.

DE OLIVEIRA JUNIOR, NG; SILVA CARDOSO, MH; FRANCO, OL. Snake venoms: attractive antimicrobial proteinaceous compounds for therapeutic purposes. **Cell Mol Life Sci**, v70, p4645-4658, 2013.

DKHIL, M; ABDEL, BA; AL, Q; AL KHALIFA, M. Anti-inflammatory activities of the venom from samsumants *Pachycondyla sennaarensis*. **Afr J Pharm Pharmacol**, v4, p115-118, 2010.

DONOVAN, GR; BALDO, BA. Pulosulin 2 from ant venom, cloning and expression of a cDNA encoding it and its antihypertensive properties. PCT International Application, 27 pp. Patent No. WO 97138541997.

DOS SANTOS PINTO, JR; FOX, EG; SAIDEMBERG, DM; SANTOS, LD; DA SILVA MENEGASSO, AR; COSTA-MANSO, E; MACHADO, EA; BUENO, OC; PALMA, MS. Proteomic view of the venom from the fire ant *Solenopsis invicta* Buren. **J Proteome Res**, v11, p4643-4653, 2012.

DUVAL, A; MALÉCOT, CO; PELHATE, M; PIEK, T. Poneratoxin, a new toxin from ant venom, reveals an interconversion between two gating modes of the Na channels in frog skeletal muscle fibres. **Pflugers Arch**, v420, p239-247, 1992.

EBAID, H; AL-KHALIFA, M; ISA, AM; GADOA, S. Bioactivity of Samsum ant (*Pachycondyla sennaarensis*) venom against lipopolysaccharides through antioxidant and upregulation of Akt1 signaling in rats. **Lipids in Health Disease**, v11, p93, 2012.

ELLIS, AK; DAY, JH. Clinical reactivity to insect stings. **Current Opinion in Allergy and Clinical Immunol**, v5, p349-354, 2005.

ESCOUBAS P. Molecular diversification in spider venoms: a web of combinatorial peptide libraries. **Mol Divers**, v10, p545-54, 2006.

ESCOUBAS, P; KING, GF. Venomics as a drug discovery platform. **Expert Rev Proteomics**, v6, p221-224, 2009.

ESTÈVE, E; MABROUK, K; DUPUIS, A; SMIDA-REZGUI, S; ALTAF AJ, X; GRUNWALD, D; PLATEL, JC; ANDREOTTI, N; MARTY, I; SABATIER, JM; RONJAT, M; DE WAARD, M. Transduction of the scorpion toxin maurocalcine into cells. Evidence that the toxin crosses the plasma membrane. **J Biol Chem**, v280, n13, p12833-9, 2005.

ESSACK, M; BAJIC, VB; ARCHER, JA. Conotoxins that confer therapeutic possibilities. **Mar Drugs**, v10, p1244-65, 2012.

EWING, B; HILLIER, L; WENDL, MC; GREEN, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Res**, v8, p175-185.

EWING, B; GREEN, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. **Genome Res**, v8, p186-194, 1998.

FAN, YX; SONG, J; SHEN, HB; KONG, X. PredCSF: an integrated feature-based approach for predicting conotoxin superfamily. **Protein Pept Lett**, v18, p261-267, 2011.

FERNÁNDEZ, F; OSPINA, M. Sinopsis de las hormigas de La región Neotropical. In: Fernández F (ed.) **Introducción a las Hormigas de La región Neotropical**. Instituto de Investigación de Recursos Biológicos Alexander Von Humboldt: Colombia; 2003. p 49-64.

FOORD, OS; ROSE, EA. Long-distance PCR. **Genome Research**, v3, p149-161, 1994.

FRITZEN, M; FLORES, MPA; REIS, CV; CHUDZINSKI-TAVASSI, AM. A prothrombin activator (Lopap) modulating inflammation, coagulation and cell survival mechanisms. **Biochem Biophys Res Commun**, v333, p517-523, 2005.

GAL, R; LIBERSAT, F. On predatory wasps and zombie cockroaches: investigations of "free will" and spontaneous behavior in insects. **Commun Integr Biol**, v3, p458-461, 2010.

GANZ, T. The role of antimicrobial peptides in innate immunity. **Integr Comp Biol**, v43, p300-304, 2003.

GAO, B; HARVEY, PJ; CRAIK, DJ; RONJAT, M; DE WAARD, M; ZHU, S. Functional evolution of scorpion venom peptides with an inhibitor cystine knot fold. **Biosci Rep**, v27, p33, 2013.

GASPAR, D; VEIGA, AS; CASTANHO, MA. From antimicrobial to anticancer peptides. A review. **Front Microbiol**, v4, p294, 2013.

GELLY, JC; GRACY, J; KAAS, Q; LE-NGUYEN, D; HEITZ, A; CHICHE, L. The KNOTTIN website and database: a new information system dedicated to the knottin scaffold. **Nucleic Acids Res**, v32, p156-159, 2004.

GIBBS, GM; SCANLON, MJ; SWARBRICK, J; CURTIS, S; GALLANT, E; DULHUNTY, AF; O'BRYAN, MK. The cysteine-rich secretory protein domain of Tpx-1 is related to ion channel toxins and regulates ryanodine receptor Ca<sup>2+</sup> signaling. **J Biol Chem**, v281, p4156-4163, 2006.

GIBBS, GM; ROELANTS, K; O'BRYAN, MK. The CAP Superfamily: Cysteine-Rich Secretory Proteins, Antigen 5, and Pathogenesis-Related 1 Proteins—Roles in Reproduction, Cancer, and Immune Defense. **Endocrine Reviews**, v29, p865–897, 2008.

GRABHERR, MG; HAAS, BJ; YASSOUR, M; LEVIN, JZ; THOMPSON, DA; AMIT, I; ADICONIS, X; FAN, L; RAYCHOWDHRY, R; ZENG, Q; CHEN, Z; MAUCELI, E; HACHOEN, N; GNIRKE, A; RHIND, N; di PALMA, F; BIRREN, BW; NUSBAUM, C; LINDBLAD-TOH, K; FRIEDMAN, N; REGEV, A. Full-length transcriptome assembly from RNA-seq data without a reference genome. **Nat Biotechnol**, v29, p644-652, 2011.

GRACY, J; CHICHE, L. Structure and modeling of knottins, a promising molecular scaffold for drug discovery. **Curr Pharm Des**, v17, p4337-50, 2011.

GULLAN, PJ; CRANSTON, PS. The insects: an outline of entomology. Wiley-Blackwell: Chichester; 2010.

GULLO, VP; MCALPINE, J; LAM, KS; BAKER, D; PETERSEN, F. Drug discovery from natural products. **J Ind Microbiol Biotechnol**, v33, p523-531, 2006.

GUPTA, A; VAN VLIJMEN, HW; SINGH, J. A classification of disulfide patterns and its relationship to protein structure and function. **Protein Sci**, v13, n8, p2045-2058, 2004.

GURROLA, GB; CAPES, EM; ZAMUDIO, FZ; POSSANI, LD; VALDIVIA, HH. Imperatoxin A, a Cell-Penetrating Peptide from Scorpion Venom, as a Probe of Ca-Release Channels/Ryanodine Receptors. **Pharmaceuticals (Basel)**, v3, n4, p1093-1107, 2010.

HADDAD JUNIOR, V. Acidentes por formigas. In: Manual de diagnóstico e tratamento de acidentes por animais peçonhentos. Brasília, FNS. p65-66, 2001.

HADDAD JUNIOR, V; CARDOSO, JLC; MORAES, RHP. Description of an injury in a human caused by a false tocandira (*Dinoponera gigantea*, Perty, 1833) with a revision on folkloric, pharmacological and clinical aspects of the giant ants of the genera Paraponera and Dinoponera (sub-family ponerinae). **Rev Inst Med Trop**, v47, p235-38, 2005.

HADDAD JUNIOR, V. Identificação de enfermidades agudas causadas por animais e plantas em ambientes rurais e litorâneos: auxílio à prática dermatológica. **An Bras Dermatol**, v84, n4, p343-348, 2009.

HARVEY, AL; BRADLEY, KN; COCHRAN, SA; ROWAN, EG; PRATT, JA; QUILLFELDT, JA; JERUSALINKY, DA. What can toxins tell us for drug discovery? **Toxicol**, v36, n11, p1635-1640, 1998.

HATFIELD, MJ; POTTER, PM. Carboxylesterase inhibitors. **Expert Opin Ther**, v8, p1159-1171, 2011.

HERMANN, HR; BLUM, MS; WHEELER, JW; OVERAL, WL; SCHMIDT, JO; CHAO, J. Comparative anatomy and chemistry of the venom apparatus and mandibular glands in *Dinoponera grandis* (Guérin) and *Paraponera clavata* (F.) (Hymenoptera: Formicidae: Ponerinae). **Annals of the Entomological Society of America**, v77, n3, p272-279, 1984.

HENRIKSEN, A; KING, TP; MIRZA, O; MONSALVE, RI; MENO, K; IPSEN, H; LARSEN, JN; GAJHEDE, M; SPANGFORT, MD. Major venom allergen of yellow jackets, Ves v 5: structural characterization of a pathogenesis-related protein superfamily. **Proteins**, v45, p438-448, 2001.

HMED, B; SERRIA, HT; MOUNIR, ZK. Scorpion peptides: potential use for new drug development. **J Toxicol**, 2013.

HOFFMAN, DR. Allergens in Hymenoptera venom XXIV: the amino acid sequences of imported fire ant venom allergens Sol i II, Sol i III, and Sol i IV. **J Allergy Clin Immunol**, v91, p71-8, 1993.

HOFFMAN, DR. Ant venoms. **Curr Opin Allergy Clin Immunol**, v10, p342-346, 2010.

HOLLDOBLE, B; WILSON, EO. The ants. Belknap Press of Harvard University Press: Cambridge, 1990.

HUANG, WN; CHEN, YH; CHEN, CL; WU, W. Surface pressure-dependent interactions of secretory phospholipase A2 with zwitterionic phospholipid membranes. **Langmuir**, v27, p7034-7041, 2011.

INAGAKI, H; AKAGI, M; IMAI, HT; TAYLOR, RW; KUBO, T. Molecular cloning and biological characterization of novel antimicrobial peptides, pilosulin 3 and pilosulin 4, from a species of the Australian ant genus *Myrmecia*. **Arch Biochem Biophys**, v428, p170-178, 2004.

INAGAKI, H; AKAGI, M; IMAI, HT; TAYLOR, RW; WIESE, MD; DAVIES, NW; KUBO, T. Pilosulin 5, a novel histamine-releasing peptide of the Australian ant, *Myrmecia pilosula* (Jack Jumper Ant). **Arch Biochem Biophys**, v477, p411-416, 2008.

JOBIN, ML; BONNAFOUS, P; TEMSAMANI, H; DOLE, F; GRELARD, A; DUFOURC, EJ; ALVES, ID. The enhanced membrane interaction and perturbation of a cell

penetrating peptide in the presence of anionic lipids: toward an understanding of its selectivity of cancer cells. **Biochim Biophys Acta**, v1828, p1457-1470, 2013.

JOHNSON, SR; COPELLO, JA; EVANS, MS; SUAREZ, AV. A biochemical characterization of the major peptides from the venom of the giant neotropical hunting ant *Dinoponera australis*. **Toxicon**, v55, p702-710, 2010.

KANG, L. Two single mutations commonly cause qualitative change of nonspecific carboxylesterases in insects. **Insect Biochem Mol Biol**, v41, p1-8, 2011.

KING, TP; SPANGFORT, MD. Structure and biology of stinging insect venom allergens. **Int Arch Allergy Immunol**, v123, p99-106, 2000.

KENNEDY, AR. The Bowman-Birk inhibitor from soybeans as an anticarcinogenic agent. **Am J Clin Nutr**, v68, n6, p1406-1412, 1998.

KNAPP, O; MCARTHUR, JR; ADAMS, DJ. Conotoxins Targeting Neuronal Voltage-Gated Sodium Channel Subtypes: Potential Analgesics? **Toxins**, v4, p1236-1260, 2012.

KOU, J; NI, Y; LI, N; WANG, J; LIU, L; JIANG, ZH. Analgesic and anti-inflammatory activities of total extract and individual fractions of Chinese medicinal ants *Polyrhachis lamellidens*. **Biol Pharm Bull**, v28, p176-180, 2005.

KOYA, S; CRENSHAW, D; AGARWAL, A. Rhabdomyolysis and acute renal failure after fire ant bites. **Society of General Internal Medicine**, v22, p145-147, 2007.

KOZAK, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. **Nucleic Acids Res**, v15, n20, p8125-8148, 1987.

LAEMMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, v227, p680-685, 1970.

LAVERGNE, V; TAFT, RJ; ALEWOOD, PF. Cysteine-rich mini-proteins in human biology. **Curr Top Med Chem**, v12, n14, p1514-1533, 2012.

LEE, JH; KWON, YB; HAN, HJ; MAR, WC; LEE, HJ; YANG, IS; BEITZ, AJ; KANG, SK. Bee venom pretreatment has both an antinociceptive and anti-inflammatory effect on carrageenan-induced inflammation. **J Vet Med Sci**, v63, n3, p251-259, 2001.

LEE, EK; JEONG, KY; LYUZ, DP; LEE, YW; SOHN, JH; LIM, KJ; HONG, CS; PARK, JW. Characterization of the major allergens of *Pachycondyla chinensis* in ant sting anaphylaxis patients. **Clin Experimental Allergy**, v39, p602-607, 2009.

LELUK, K; SCHMIDT, J; JONES, D. Comparative studies on the protein composition of hymenopteran venom reservoirs. **Toxicon**, v27, n1, p105-114, 1989.

LIM, KL; SUNG, BH; SHIN, JR; LEE, YW; KIM DA, J; YANG, KS; KIM, SC. A cancer specific cell-penetrating peptide, BR2, for the eficiente delivery of na scFv into cancer cells. **PLoS One**, v8, e66084, 2013.

LIMA, PRM; BROCHETTO-BRAGA, MR. Hymenoptera venom review focusing on *Apis mellifera*. **JVAT**, v9, p149-162, 2003.

LISCOVITCH, M; CZARNY, M; FIUCCI, G; TANG, X. Phospholipase D: molecular and cell biology of a novel gene family. **Biochem J**, v3, p401-415, 2000.

LOPES, KS; RIOS, E; DANTAS, RT; LIMA, C; LINHARES, M, *et al.* (2011) Effect of *Dinoponera quadriceps* venom on chemical-induced seizures models in mice. Rencontres en Toxinologie – Meeting on Toxinology SFET Editions. Available: <http://www.sfet.asso.fr>. Accessed 20 August 2012.

LOPES, KS; RIOS, ERV; LIMA, CNC; LINHARES, MI; TORRES, AFC; HAVT, A; QUINET, YP; FONTELES, MM; MARTINS, AM. The effects of the brazilian ant *Dinoponera quadriceps* venom on chemically induced seizure models. **Neurochem Intern**, v63, p141-145, 2013.

MACKINTOSH, JA; VEAL, DA; BEATTIE, AJ; GOOLEY, AA. Isolation from an ant *Myrmecia gulosa* of two inducible O-glycosylated proline-rich antibacterial peptides. **J Biol Chem**, v13, p6139-6143, 1998.

MADANI, F; LINDBERG, S; LANGEL, ULO, FUTAKI, S; GRASLUND, A. Mechanisms of cellular uptake of cell-penetrating peptides. **J Biophys**, v2011, 2011.

MAGALHÃES, GS; CAPORRINO, MC; DELLA-CASA, MS; KIMURA, LF; PREZOTTO-NETO, JP; FUKUDA, DA; PORTES-JUNIOR, JA; NEVES-FERREIRA, AG; SANTORO, ML; BARBARO, KC. Cloning, expression and characterization of a phospholipase D from *Loxosceles gaucho* venom gland. **Biochimie**, v9, p1773-1783, 2013.

MARQUES, JR; DA FONSECA, RR; DRURY, B; MELO A. Conformational characterization of disulfide bonds: a tool for protein classification. **J Theor Biol**, v7, n267(3), p388-95, 2010.

MAS, JM; ALOY, P; MARTÍ-RENOM, MA; OLIVA, B; DE LLORENS, R; AVILÉS, FX; QUEROL, E. Classification of protein disulphide-bridge topologies. **J Comput Aided Mol Des**, v15, n5, p477-87, 2001.

MASCHWITZ, UWJ; KLOFT, W. Morphology and function of the venom apparatus of insects – bees, wasps, ants and caterpillars. In: BUCKLEY, E. E.; BUCHERL, W. Venomous animals and their venoms. New York: Academic Press, v3, p1-56, 1971.

MASEK, T; VOPALENSKY, V; SUCHOMELOVA, P; POSPISEK, M. Denaturing RNA electrophoresis in TAE agarose gels. **Anal Biochem**, v336, p46-50, 2005.

MATUSZEK, MA; HODGSON, WC; SUTHERLAND, SK; KING, RG. Pharmacological studies of jumper ant (*Myrmecia pilosula*) venom: evidence for the presence of histamine, and haemolytic and eicosanoid-releasing factors. **Toxicon**, v30, p1081-1091, 1992.

MCDEMOTT, M; WAKELAM, MJ; MORRIS, AJ. Phospholipase D. **Biochem Cell Biol**, v82, p225-253, 2004.

MESSING, J; GRONENBORN, B; MULLER-HILL, B; HANS HOPSCHEIDER, P. Filamentous Coliphage M13 as a cloning vehicle: Insertion of a HindII fragment of the lac regulatory region in M13 replicative form in vitro. **Proc Natl Acad Sci USA**, v79, p3642-3646, 1977.

MICHEL, Y; MCINTYRE, M; GINGLINGER, H; OLLERT, M; CIFUENTES, L; BLANK, S; SPILLNER, E. The putative serine protease inhibitor Api m 6 from *Apis mellifera* venom: recombinant and structural evaluation. **Immunol**, v22, n7, p476-484, 2012.

MIN, XJ; BUTLER, G; STORMS, R; TSANG, A. OrfPredictor: predicting protein-coding regions in EST-derived sequences. **Nucleic Acids Res**, v33, p677-680, 2005.

MONNIN, T; PEETERS, C. Monogyny and regulation of worker mating in the queenless ant *Dinoponera quadriceps*. **Animal Behaviour**, v55, n2, p299-306, 1998.

MORITA, S; MATSUSHITA, N; KAWACHI, T; AYABE, R; MIYASHITA, M; MIYAKAWA, T; NAKAGAWA, Y; NAGATA, K; MIYAGAWA, H; TANOKURA, M. Solution structure of a short-chain insecticidal toxin LaIT1 from the venom of scorpion *Liocheles australasiae*. **Biochem Biophys Res Commun**, v411, n4, p738-44, 2011.

MORTARI, MR; CUNHA, AO; FERREIRA, LB; DOS SANTOS, WF. Neurotoxins from invertebrates as anticonvulsants: from basic research to therapeutic application. **Pharmacol Ther**, v114, p171-183, 2007.

MOURÃO, CBF; SCHWARTZ, EF. Protease inhibitors from marine venomous animals and their counterparts in terrestrial venomous animals. **Mar Drugs**, v11, p2069-2112, 2013.

NAM, KW; JE, KH; LEE, JH; HAN, HJ; KANG, SK; MAR, W. Inhibition of COX-2 activity and proinflammatory cytokines (TNF-alpha and IL-1 beta) production by water-soluble sub-fractionated parts from bee (*Apis mellifera*) venom. **Arch Pharm Res**, v26, p383-388, 2003.

NEWMAN, DJ; CRAGG, GM. Natural Products as Sources of New Drugs over the 30 Years from 1981 to 2010. **J Nat Prod**, v75, p311-335, 2012.

NYGAARD, S; ZHANG, G; SCHIØTT, M; LI, C; WURM, Y; HU, H; ZHOU, J; JI, L; QIU, F; RASMUSSEN, M; PAN, H; HAUSER, F; KROGH, A; GRIMMELIKHUIJZEN, CJ; WANG, J; BOOMSMA, JJ. The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. **Genome Res**, v2, n8, p1339-48, 2011.

NUNES, KP; COSTA-GONÇALVES, A; LANZA, LF; CORTES, SF; CORDEIRO, MN; RICHARDSON, M; PIMENTA, AM; WEBB, RC; LEITE, R; DE LIMA, ME. Tx2-6 toxin of the *Phoneutria nigriventer* spider potentiates rat erectile function. **Toxicon**, v51, p1197-206, 2008.

ORIVEL, J; DEJEAN, A. Comparative effect of the venoms of ants of the genus *Pachycondyla* (Hymenoptera: Ponerinae). **Toxicon**, v39, p195-201, 2001.

ORIVEL, J; REDEKER, V; LE CAER, JP; KRIER, F; REVOL-JUNELLES, AM; LONGEON, A; CHAFOTTE, A; DEJEAN, A; ROSSIER, J. Ponericins, new antibacterial and insecticidal peptides from the venom of the ant *Pachycondyla goeldii*. **J Biol Chem**, v276, p17823-17829, 2001.

PADAVATTAN, S; SCHMIDT, M; HOFFMAN, DR; MARKOVIC-HOUSLEY, Z. Crystal structure of the major allergen from fire ant venom, Sol i 3. **J Mol Biol**, v383, p178-185, 2008.

PALMA, MS; YAMANE, T; CAMARGO, ACM. Biodiversidade: preservação e bioprospecção, 2001. Disponível em: <<http://www.comciencia.br/reportagens/biodiversidade/bio13.htm>>. Acesso em: Outubro 2013.

PARK, JM; JUNG, JE; LEE, BJ. Antimicrobial peptides from the skin of a Korean frog, *Rana rugosa*. **Biochem. Biophys Res Commun**, v205, p948-954, 1994.

PETHICA, RB; LEVITT, M; GOUGH, J. Evolutionarily consistent families in SCOP: sequence, structure and function. **BMC Struct Biol**, v12, p27, 2012.

PIEK, T; DUVAL, A; HUE, B; KARST, H; LAPIED, B; MANTEL, P; NAKAJIMA, T; PELHATE, M; SCHMIDT, JO. Poneratoxin, a novel peptide neurotoxin from the venom of the ant, *Paraponera clavata*. **Comp Biochem Physiol C**, v99, n3, p487-495, 1991.

POPE, JE; DEER, TR. Ziconotide: a clinical update and pharmacologic review. **Expert Opin Pharmacother**, v14, p957-966, 2013.

POWERS, JP; HANCOCK, RE. The relationship between peptide structure and antibacterial activity. **Peptides**, v24, n11, p1681-1691, 2003.

PUILLANDRE, N; KOUA, D; FAVREAU, P; OLIVERA, BM; STÖCKLIN, R. Molecular phylogeny, classification and evolution of conopeptides. **J Mol Evol**, v74, n5-6, p297-309, 2012.

RACHAIAH, NM; JAYAPPAGOWDA, LA; SIDDABYRAPPA, HB; BHARATH, VK. Unusual case of acute renal failure following multiple wasp stings. **N Am J Med Sci**, v4, n2, p104-106, 2012.

RANASINGHE, S; MCMANUS, DP. Structure and function of invertebrate Kunitz serine protease inhibitors. **Dev Comp Immunol**, v39, n3, p219-227, 2012.

RATZKA, C; FÖRSTER, F; LIANG, C; KUPPER, M; DANDEKAR, T; FELDHAAR, H; GROSS, R. Molecular characterization of antimicrobial peptide genes of the carpenter ant *Camponotus floridanus*. **PLoS One**, v7, p8, 2012.

REDDY, KV; YEDERY, RD; ARANHA, C. Antimicrobial peptides: premises and promises. **Int J Antimicrob Agents**, v24, n6, p536-547, 2004.

REIS, CV; FARSKY, SHP; FERNANDES, BL; SANTORO, ML; OLIVA, MLV; MARIANO, M; CHUDZINSKI-TAVASSI, AM. *In vivo* characterization of Lopap, prothrombin activator serine protease from the *Lonomia obliqua* caterpillar venom. **Thrombosis Res**, v. 102, p. 437-443, 2001.

RICHMOND, GS; SMITH, TK. Phospholipases A1. **Inter J Mol Sci**, v12, p588-612, 2011.

RIFFLET, A; GAVALDA, S; TÉNÉ, N; ORIVEL, J; LEPRINCE, J; GUILHAUDIS, L; GÉNIN, E; VÉTILLARD, A; TREILHOU, M. Identification and characterization of a novel antimicrobial peptide from the venom of the ant *Tetramorium bicarinatum*. **Peptides**, v38, n2, p363-370, 2012.

ROCHA, T; de SOUZA, BM; PALMA, MS; da CRUZ-HOFLING, MA; HARRIS, JB. The neurotoxicological effects of mastoparan Polybia-MPII at the murine neuromuscular junction: an ultrastructural and immunocytochemical study. **Histochem Cell Biol**, v132, p395-404, 2009.

RODRÍGUEZ, AA; SALCEDA, E; GARATEIX, AG; ZAHARENKO, AJ; PEIGNEUR, S; LÓPEZ, O; PONS, T; RICHARDSON, M; DÍAZ, M; HERNÁNDEZ, Y; STÄNDKER, L; TYTGAT, J; SOTO E. A novel sea anemone peptide that inhibits acid-sensing ion channels. **Peptides**, 2013. doi: S0196-9781(13)00220-9.

SALVADOR, LA; TAORI, K; BIGGS, JS; JAKONCIC, J; OSTROV, DA; PAUL, VJ; LUESCH, H. Potent elastase inhibitors from cyanobacteria: structural basis and mechanisms mediating cytoprotective and anti-inflammatory effects in bronchial epithelial cells. **J Med Chem**, v56, n3, p1276-1290, 2013.

SAITOU, N; IMANISHI, T. Relative efficiencies of the Fitch-Margoliash, maximum-parsimony, maximum-likelihood, minimum-evolution, and neighbor-joining methods of phylogenetic tree construction in obtaining the correct tree. **Mol Biol Evol**, v6, p514-525, 1989.

SAMY, RP; GOPALAKRISHNAKONE, P; STILES, BG; GIRISH, KS; SWAMY, SN; HEMSHEKHAR, M; TAN, KS; ROWAN, EG; SETHI, G; CHOW, VT. Snake venom phospholipases A (2): a novel tool against bacterial diseases. **Curr Med Chem**, v19, n36, p6150-6162, 2012.

SANGER, F; NICKLEN, S; COULSON, AR. DNA sequencing with chain-terminating inhibitors. **Proc Natl Acad Sci**, v74, p5463–5467, 1977.

SCHMIDT, JO; BLUM, MS. A harvester ant venom: chemistry and pharmacology. **Science**, v200, p1064-1066, 1978.

SCHMIDT, JO; BLUM, MS; OVERAL, WL. Comparative enzymology of venoms from stinging Hymenoptera. **Toxicon**, v24, n9, p907–921, 1986.

SCHOETERS, E; BILLEN, J. Morphology and ultra structure of the convoluted gland in the ant *Dinoponera australis* (Hymenoptera: Formicidae). **Int J Insect Morphol Embryol**, v24, n3, p323-332, 1995.

SEISMANN, H; BLANK, S; CIFUENTES, L; BRAREN, I; BREDEHORST, R; GRUNWALD, T; OLLERT, M; SPILLNER, E. Recombinant phospholipase A1 (Ves v 1) from yellow jacket venom for improved diagnosis of hymenoptera venom hypersensitivity. **Clin Mol Allergy**, v8, p7, 2010.

SHI, NQ; GAO, W; XIANG, B; QI, XR. Enhancing cellular uptake of activable cell-penetrating peptide-doxorubicin conjugate by enzymatic cleavage. **Int J Nanomedicine**, v7, p1613-1621, 2012.

SHIGETOMI, H; ONOGI, A; KAJIWARA, H; YOSHIDA, S; FURUKAWA, N; HARUTA, S; TANASE, Y; KANAYAMA, S; NOGUCHI, T; YAMADA, Y; OI, H; KOBAYASHI, H. Anti-inflammatory actions of serine protease inhibitors containing the Kunitz domain. **Inflamm Res**, v59, p679-687, 2010.

SIMMACO, M; MIGNOGNA, G; CANOFENI, S; MIELE, R; MANGONI, ML; BARRA, D. Temporins, antimicrobial peptides from the European red frog *Rana temporaria*. **Eur J Biochem**, v242, p788-792, 1996.

SIQUIEROLI, ACS; SANTANA, FA; RODRIGUES, RS; VIEIRA, CU; CARDOSO, R; GOULART, LR; BONETTI, AM. Phage display in venom gland in *Dinoponera australis* (Hymenoptera: Formicidae). **J Venom Anim Toxins incl Trop Disc**, IX Symposium of the Brazilian Society on Toxinology, v13, n1, p291, 2007.

SMITH, CD; ZIMIN, A; HOLT, C; ABOUHEIF, E; BENTON, R; CASH, E; CROSET, V; CURRIE, CR; ELHAIK, E; ELSIK, CG; FAVE, MJ; FERNANDES, V; GADAU, J; GIBSON, JD; GRAUR, D; GRUBBS, KJ; HAGEN, DE; HELMKAMPF, M; HOLLEY, JA; HU, H; VINIEGRA, AS; JOHNSON, BR; JOHNSON, RM; KHILA, A; KIM, JW; LAIRD, J; MATHIS, KA; MOELLER, JA; MUÑOZ-TORRES, MC; MURPHY, MC; NAKAMURA, R; NIGAM, S; OVERSON, RP; PLACEK, JE; RAJAKUMAR, R; REESE, JT; ROBERTSON, HM; SMITH, CR; SUAREZ, AV; SUEN, G; SUHR, EL; TAO, S; TORRES, CW; VAN WILGENBURG, E; VILJAKAINEN, L; WALDEN, KK; WILD, AL; YANDELL, M; YORKE, JA; TSUTSUI, ND. Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*). **Proc Natl Acad Sci U S A**, v108, n14, p5673-8, 2011a.

SMITH, CR; SMITH, CD; ROBERTSON, HM; HELMKAMPF, M; ZIMIN, A; YANDELL, M; HOLT, C; HU, H; ABOUHEIF, E; BENTON, R; CASH, E; CROSET, V; CURRIE, CR; ELHAIK, E; ELSIK, CG; FAVÉ, MJ; FERNANDES, V; GIBSON, JD; GRAUR, D; GRONENBERG, W; GRUBBS, KJ; HAGEN, DE; VINIEGRA, AS; JOHNSON, BR; JOHNSON, RM; KHILA, A; KIM, JW; MATHIS, KA; MUNOZ-TORRES, MC; MURPHY, MC; MUSTARD, JA; NAKAMURA, R; NIEHUIS, O; NIGAM, S; OVERSON, RP; PLACEK, JE; RAJAKUMAR, R; REESE, JT; SUEN, G; TAO, S; TORRES, CW; TSUTSUI, ND; VILJAKAINEN, L; WOLSCHIN, F; GADAU, J. Draft genome of the red harvester ant *Pogonomyrmex barbatus*. **Proc Natl Acad Sci U S A**, v108, n14, p5667-72, 2011b.

SONODA, H; AOKI, J; HIRAMATSU, T; ISHIDA, M; BANDO, K; NAGAI, Y; TAGUCHI, R; INOUE, K; ARAI, H. A Novel Phosphatidic acid-selective phospholipase A1 that produces lysophosphatidic acid. **J Biol Chem**, v277, p 34254 –34263, 2002.

SOUSA, PL; QUINET, Y; PONTE, EL; DO VALE, JF; TORRES, AF; PEREIRA, MG; ASSEURY, AM. Venom's antinociceptive property in the primitive ant *Dinoponera quadriceps*. **J Ethnopharmacol**, v144, n1, p213-216, 2013.

SUBASINGHAGE, AP; O'FLYNN, D; CONLON, JM; HEWAGE, CM. Conformational and membrane interaction studies of the antimicrobial peptide alyteserin-1c and its analogue [E4K]alyteserin-1c. **Biochim Biophys Acta**, v8, p1975-84, 2011.

SUEN, G; TEILING, C; LI, L; HOLT, C; ABOUHEIF, E; BORNBERG-BAUER, E; BOUFFARD, P; CALDERA, EJ; CASH, E; CAVANAUGH, A; DENAS, O; ELHAIK, E; FAVÉ, MJ; GADAU, J; GIBSON, JD; GRAUR, D; GRUBBS, KJ; HAGEN, DE; HARKINS, TT; HELMKAMPF, M; HU, H; JOHNSON, BR; KIM, J; MARSH, SE; MOELLER, JA; MUÑOZ-TORRES, MC; MURPHY, MC; NAUGHTON, MC; NIGAM, S; OVERSON, R; RAJAKUMAR, R; REESE, JT; SCOTT, JJ; SMITH, CR; TAO, S; TSUTSUI, ND; VILJAKAINEN, L; WISSLER, L; YANDELL, MD; ZIMMER, F; TAYLOR, J; SLATER, SC; CLIFTON, SW; WARREN, WC; ELSIK, CG; SMITH, CD; WEINSTOCK, GM; GERARDO, NM; CURRIE, CR. The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. **PLoS Genet**, v7, n2, p1002007, 2011.

SZOLAJSKA, E; POZNANKI, J; FERBER, ML; MICHALIK, J; GOUT, E; FENDER, P; BAILLY, I; DUBLET, B; CHROBOCZEK, J. Poneratoxin, a neurotoxin from ant venom. Structure and expression in insect cells and construction of a bio-insecticide. **Europ J Biochem**, v271, n11, p2127-2136, 2004.

TAGUCHI, S; BULET, P; HOFFMANN, JA. A novel insect defensin from the ant *Formica rufa*. **Biochimie.**, v80, n4, p343-346, 1998.

TAKADA, K; HAMADA, T; HIROTA, H; NAKAO, Y; MATSUNAGA, S; VAN SOEST, RW; FUSETANI, N. Asteropine A, a sialidase-inhibiting conotoxin-like peptide from the marine sponge *Asteropus simplex*. **Chem Biol**, v13, n6, p569-574, 2006.

TAMURA, K; PETERSON, D; PETERSON, N; STECHER, G; NEI, M; KUMAR, S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood,

Evolutionary Distance, and Maximum Parsimony Methods. **Molecular Biology and Evolution**, v28, p2731-2739, 2011.

TANKERSLEY, MS. The stinging impact of the imported fire ant. **Current Opinion in Allergy and Clinical Immunology**, v8, p354–359, 2008.

THOMPSON, AH; BJOURSON, AJ; ORR, DF; SHAW, C; MCCLEAN, S. A combined mass spectrometric and cDNA sequencing approach to the isolation and characterization of novel antimicrobial peptides from the skin secretions of *Phyllomedusa hypochondrialis* azurea. **Peptides**, v28, n7, p1331-1343, 2007.

TORRES, AFC; QUINET, YP; HAVT, A; RÁDIS-BAPTISTA, G; MARTINS, AMC (2013) Molecular pharmacology and toxinology of venom from ants. In: Rádis-Baptista G editor. Biochemistry, Genetics and Molecular Biology: An Integrated view of the molecular recognition and toxinology – From analytical procedures to biomedical applications. ISBN 978-953-51-1151-1 pp. InTech, DOI: 10.5772/53539. Available: <http://www.intechopen.com/books/an-integrated-view-of-the-molecular-recognition-and-toxinology-from-analytical-procedures-to-biomedical-applications/molecular-pharmacology-and-toxinology-of-venom-from-ants>. Accessed in 14 August 2013.

TOWSE, CL; DAGGETT, V. When a domain is not a domain, and why it is important to properly filter proteins in databases: conflicting definitions and fold classification systems for structural domains make filtering of such databases imperative. **Bioessays**, v34, n12, p1060-1069, 2012.

TURK, B. Targeting proteases: successes, failures and future prospects. **Nat Rev Drug Discov**, v5, n9, 785-799, 2006.

TSUTSUI, ND; SUAREZ, AV; SPAGNA, JC; JOHNSTON, JS. The evolution of gene size in ants. **BMC Evol Biol**, v8, p64, 2008.

VAN VAERENBERGH, M; CARDOEN, D; FORMESYN, EM; BRUNAIN, M; VAN DRIESSCHE, G; BLANK, S; SPILLNER, E; VERLEYEN, P; WENSELEERS, T; SCHOOF, L; DEVREESE, D; de GRAAF, DC. Extending the honey bee venom with the antimicrobial peptide apidaecin and a protein resembling wasp antigen 5. **Insect Mol Biol**, v22, p199-210, 2013.

VASSILEVSKI, AA; FEDOROVA, IM; MALEEVA, EE; KOROLKOVA, YV; EFIMOVA, SS; SAMSONOVA, OV; SCHAGINA, LV; FEOFANOV, AV; MAGAZANIK, LG; GRISHIN, EV. Novel class of spider toxin: active principle from the yellow sac spider *Heiracanthium puncturium* venom is a unique two-domain polypeptide. **J Biol Chem**, v285, n42, p32293-302, 2010.

VERANO-BRAGA, T; ROCHA-RESENDE, C; SILVA, DM; LANZER, D; MARTIN-EAUCLAIRE, MF; BOUGIS, PE; DE LIMA, ME; SANTOS, RAS; PIMENTA, AMC. *Tityus serrulatus* hypotensins: a new family of peptides from scorpion venom. **Biochem Biophys Res Com**, v371, p515-520, 2008.

VIEGAS JUNIOR C; BOLZANI, VS; BARREIRO, EJ. Os produtos naturais e a química medicinal moderna. **Quim Nova**, v29, p326-337, 2006.

VILJAKAINEN, L; PAMILO, P. Identification and molecular characterization of defensin gene from the ant *Formica aquilonia*. **Insect Mol Biol**, v14, n4, p335-338, 2005.

VINHOTE, JFC; TORRES, AFC; DANTAS, RT; PRACIANO, TP; MENEZES, RPPB; SOUSA, D; BRITO, TS; LIMA, FJB; TOYAMA, MH; MAGALHÃES, PJ; MONTEIRO, HAS; MARTINS-NUNES, AMC. Renal- and calcium-dependent vascular effects of *Polybia paulista* wasp venom. **JVAT**, v17, p199-208, 2011.

VOGT, RG; PRESTWICH, GD; LERNER, MR. Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects. **J Neurobiol**, v22, n1, p74-84, 1991.

WAGSTAFF, SC; SANZ, L; JUÁREZ, P; HARRISON, RA; CALVETTE, JJ. Combined snake venomomics and venom gland transcriptomic of the ocellated carpet viper *Echis ocellatus*. **J of proteomics**, v17, p609-523, 2009.

WALLACE, MS. Ziconotide: a new nonopioid intrathecal analgesic for the treatment of chronic pain. **Expert Rev Neurother**, v6, n10, p1423-1428, 2006.

WANG, CP; WU, YL. Study on mechanism underlying the treatment of rheumatoid arthritis by Keshiling. **Zhongguo Zong Yao Za Zhi**, v31, p155-158, 2006.

WARD, PS. Ants. **Curr Biol**, v7, n16, p152-155, 2006.

WERREN, JH; RICHARDS, S; DESJARDINS, CA; NIEHUIS, O; GADAU, J; et al. Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. **Science**, v327, p343-348, 2010.

WIESE, MD; CHATAWAY, TK; DAVIES, NW; MILNE, RW; BROWN, SGA; GAI, WP; HEDDLE, RJ. Proteomic analysis of *Myrmecia pilosula* (jack jumper) ant venom. **Toxicon**; v47, n2, p208-217, 2006.

WIESNER, J; VILCINSKAS, A. Antimicrobial peptides: the ancient arm of the human immune system. **Virulence**, v1, p440-464, 2010.

WILSON, EO. The insect societies. Belknap Press of Harvard University Press: Cambridge, 1971.

WINNEBECK, EC; MILLAR, CD; WARMAN, GR. Why does insects RNA look degraded? **Journal of Insect Science**, v10, p159, 2010.

WURM, Y; WANG, J; RIBA-GROGNUM, O; CORONA, M; NYGAARD, S; HUNT, BG; INGRAM, KK; FALQUET, L; NIPITWATTANAPHON, M; GOTZEK, D; DIJKSTRA, MB; OETTLER, J; COMTESSE, F; SHIH, CJ; WU, WJ; YANG, CC; THOMAS, J; BEAUDOING, E; PRADERVAND, S; FLEGEL, V; COOK, ED; FABBRETTI, R; STOCKINGER, H; LONG, L; FARMERIE, WG; OAKLEY, J; BOOMSMA, JJ; PAMILO,

P; YI, SV; HEINZE, J; GOODISMAN, MA; FARINELLI, L; HARSHMAN, K; HULO, N; CERUTTI, L; XENARIOS, I; SHOEMAKER, D; KELLER, L. The genome of the fire ant *Solenopsis invicta*. **Proc Natl Acad Sci U S A**, v108, n14, p5679-84, 2011.

YANDEK, LE; POKORNY, A; ALMEIDA, PF. Wasp mastoparans follow the same mechanism as the cell-penetrating peptide transportan 10. **Biochemistry**, v48, p7342-7351, 2009.

YANG, W; FENG, J; WANG, B; CAO, Z; LI, W; WU, Y; CHEN, Z. BF9, the first functionally characterized snake toxin peptide with Kunitz-Type protease and potassium channel inhibiting properties. **J Biochem Mol Toxicol**, 2013.

ZALAT, S; NABIL, Z; HUSSEIN, A; RAKHA, M. Biochemical and haematological studies of some solitary and social bee venoms. **Egyptian J of Biology**, v1, p55-71, 1999.

ZALAT, S; SCHMIDT, J; MOAWAD, TI. Lipase and phospholipase activities of Hymenoptera venoms (wasps and ants). **Egyptian J Biol**, v5, p138-147, 2003.

ZELEZETSKY, I; PAG, U; ANTCHEVA, N; SAHL, HG; TOSSI A. Identification and optimization of an antimicrobial peptide from the ant venom toxin pilosulin. **Biochem Biophys**, v434, p358-364, 2005.

ZENG, XC; LIU, Y; SHI, W; ZHANG, L; LUO, X; NIE, Y; YANG, Y. Genome-wide search and comparative genomic analysis of the trypsin inhibitor-like cysteine-rich domain-containing peptides. **Peptides**, v13, p283-290, 2013.

ZHANG, Z; ZHU, S. Comparative genomics analysis of five families of antimicrobial peptide-like genes in seven ant species. **Immunol**, v38, n2, p262-274, 2012.