

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CAMPUS SOROCABA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E MONITORAMENTO
AMBIENTAL

Análise de peptídeos antimicrobianos na formiga *Odontomachus chelifer* (Latreille, 1802) (Formicidae: Ponerinae) a partir de dados transcriptômicos

Josilene de Jesus Menk

Sorocaba - SP
2024

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CAMPUS SOROCABA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E MONITORAMENTO
AMBIENTAL

Josilene de Jesus Menk

Análise de peptídeos antimicrobianos na formiga *Odontomachus chelifer* (Latreille, 1802) (Formicidae: Ponerinae) a partir de dados transcriptômicos

Tese apresentada à Banca Examinadora do Programa de Pós-Graduação em Biotecnologia e Monitoramento Ambiental da Universidade Federal de São Carlos, atendendo ao requisito para a obtenção do título de Doutor em Biotecnologia e Monitoramento Ambiental.

Orientação: Prof. Dr. Flavio Henrique da Silva
Coorientação: Profa. Dra. Renata Santos Rodrigues

Sorocaba - SP
2024



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências e Tecnologias Para a Sustentabilidade
Programa de Pós-Graduação em Biotecnologia e Monitoramento Ambiental

Folha de Aprovação

Defesa de Tese de Doutorado da candidata Josilene de Jesus Menk, realizada em 03/05/2024.

Comissão Julgadora:

Prof. Dr. Flavio Henrique da Silva (UFSCar)



Documento assinado digitalmente

FLAVIO HENRIQUE DA SILVA

Data: 07/05/2024 10:41:22-0300

Verifique em <https://validar.iti.gov.br>

Profa. Dra. Milene Ferro (UNESP)

Prof. Dr. Malson Neilson de Lucena (UFMS)

Prof. Dr. Marcelo de Bello Cioffi (UFSCar)

Prof. Dr. Ricardo Carneiro Borra (UFSCar)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Biotecnologia e Monitoramento Ambiental.

Agradecimentos

Agradeço a Deus por ter me dado a vida. Agradeço a toda a minha família, que me ensinaram a sempre lembrar de onde eu vim. As minhas amigas e amigos que acompanharam o meu processo durante toda essa jornada a distância. Sou grata por todos que acompanharam essa fase da minha vida e fizeram parte direta ou indiretamente.

Ao meu orientador prof. Dr. Flavio Henrique, minha total gratidão por sempre estar presente, me orientando e ensinando, compartilhando sua sabedoria e conhecimento, contribuindo para o meu aprendizado e desenvolvimento e sempre tentando extrair o meu melhor. À profa. Dra. Renata, que foi minha coorientadora, pela contribuição e colaboração nesta pesquisa magnífica.

Ao grupo do artigo (Mestre Yan, graduando Henrique, prof. Dr. Flavio Henrique, profa. Dra. Milene, profa. Dra. Renata, prof. Dr. Celio), obrigada pelo engajamento, colaboração, conhecimento, sabedoria, empenho, comprometimento e dedicação que cada um desempenhou para que esse artigo fosse construído e publicado; esse artigo é fruto dessas mentes e mãos que trabalharam arduamente desempenhando um trabalho em equipe efetivo e produtivo.

A toda a equipe do LBM/UFSCAR-São Carlos (Laboratório de Biologia Molecular) "LBMers": Amanda, Bianca, Camila, Eduardo, Felipe, Gabrielli, Henrique, Iara, Karen, Letícia, Thais e Yan e principalmente a todos que passaram pelo LBM; onde a cultura de transferir conhecimento, ajudar e cooperar, compartilhar conhecimento e apoiar foi sempre mútuo. Obrigada a todos por terem participado dessa fase riquíssima da minha vida profissional e pessoal, espero poder ter contribuído mutuamente com algo positivo na vida de cada um.

À todos os meus professores da graduação na UNISO (Universidade de Sorocaba) sem eles eu jamais teria conseguido chegar até aqui e meus sinceros agradecimentos à profa. Dra. Denise Grotto, que me "mostrou" a ciência no último semestre da graduação, que despertou esse fascínio em aprender e pesquisar. Grata por todos os professores da UFSCAR/Sorocaba e UFSCAR/São Carlos com quem pude ter a satisfação e o privilégio em ter aula e aprender um pouco mais de um universo totalmente novo. Em especial aos professores Iolanda e Tiago Palladino UFSCAR/Sorocaba, que compartilharam e ensinaram em suas aulas e onde ouvi pela primeira vez sobre o método Illumina; me encantei por algo totalmente fora dos meus conhecimentos técnicos mas que de primeira me pareceu encantador apesar de ser totalmente complexo para o meu intelecto. Minha gratidão eterna à profa. Dra. Andrea Soares (DGE/UFSCAR-São Carlos), onde me ajudou prontamente sempre que precisei.

À banca da minha qualificação (profa. Dra. Andrea (UFSCAR/São Carlos), profa. Dra. Milene (UNESP), profa. Dra. Tassia (UFU) e prof. Dr. Vladimir (UNIOESTE), que foram fundamentais para o desenvolvimento da tese.

À Universidade Federal de São Carlos (UFSCAR - campus São Carlos e Sorocaba) e ao Programa de Biotecnologia e Monitoramento Ambiental (PPGBMA).

À CAPES pela concessão da bolsa de demanda social de doutorado.

“O que conhecemos é uma gota, o que ignoramos é um oceano”
(Isaac Newton).

Resumo

As peçonhas dos himenópteros contêm uma mistura complexa de potentes substâncias bioativas, as quais possuem potenciais aplicações biotecnológicas. As peçonhas de formigas são compostas por ácido fórmico, alcalóides, amins biogênicas, enzimas, hidrocarbonetos, peptídeos, polipeptídeos e proteínas. Esses componentes apresentam propriedades com atividade antifúngica, antiparasitária, antimicrobiana, antiviral, fosfolipásica, citolítica e neurotóxica. A peçonha da espécie *Odontomachus chelifer*, uma formiga carnívora, foi investigada neste estudo de forma inédita. *O. chelifer* é uma formiga carnívora, com hábitos noturnos, grandes que possuem um corpo anatômico característico da sua espécie, apresentando uma mandíbula robusta. Sua picada causa dor aguda e intensa, coceira prolongada e queimação. Neste estudo, investigamos, usando técnicas *in silico*, os peptídeos que compõem um suposto arsenal antimicrobiano da glândula de peçonha da formiga neotropical *Odontomachus chelifer*. Com base nos transcritos do corpo e da glândula de peçonha desse inseto, foi possível determinar o secretoma da glândula, que continha cerca de 1.022 peptídeos com supostos peptídeos sinalizadores. A maioria desses peptídeos (75,5%) era desconhecida, não correspondendo a nenhum banco de dados de referência, motivando-nos a extrair informações funcionais por meio de técnicas baseadas em análise de dados automatizados. Utilizamos metodologias complementares como dados transcriptômicos, extração e sequenciamento de RNA, sequenciamento e montagem *De Novo*; para investigarmos a existência de peptídeos antimicrobianos (PAMs) na glândula de peçonha de *O. chelifer*, encontrando 112 candidatos não redundantes. Os PAMs foram submetidos as ferramentas online de bioinformática (AI4AMP, AMPIfy, AMP Scanner v2, amPEPpy, Macrel v.1.2.0, UNIPROT versão 2022_05, CPAMR3, CAMPSign, Blastp) e o banco de dados NCBI. Os PAMs candidatos são mais globulares e hemolíticos do que os peptídeos restantes no secretoma. Há evidências de transcrição para 97% dos candidatos a PAM no mesmo gênero de formigas, com um deles também verificado como traduzido, apoiando assim nossos achados. A maioria dessas sequências com potencial antimicrobiano (94,8%) correspondeu a transcrições do corpo da formiga, indicando seu papel não apenas como toxinas da peçonha.

Palavras-chave: Himenópteros, peçonha, peptídeos antimicrobianos; peptídeos de formiga; *Odontomachus chelifer*.

Abstract

Hymenopteran venoms contain a complex mixture of potent bioactive substances, which have potential biotechnological applications. Ant venoms are composed of formic acid, alkaloids, biogenic amines, enzymes, hydrocarbons, peptides, polypeptides and proteins. These components have antifungal, antiparasitic, antimicrobial, antiviral, phospholipase, cytolytic and neurotoxic properties. The venom of the species *Odontomachus chelifer*, a carnivorous ant, was investigated in this study in a novel way. *O. chelifer* is a carnivorous, nocturnal, large ant with an anatomical body characteristic of its species, featuring a robust mandible. Its sting causes acute and intense pain, prolonged itching and burning. In this study, we investigated, using *in silico* techniques, the peptides that make up a supposed antimicrobial arsenal from the venom gland of the neotropical ant *Odontomachus chelifer*. Based on transcripts from the body and venom gland of this insect, it was possible to determine the secretome of the gland, which contained around 1,022 peptides with supposed signaling peptides. Most of these peptides (75.5%) were unknown and did not correspond to any reference database, motivating us to extract functional information using techniques based on automated data analysis. We used several complementary methodologies such as transcriptomic data to investigate the existence of antimicrobial peptides (AMPs) in the venom gland of *O. chelifer*, finding 112 non-redundant candidates. The PAMs were submitted to online bioinformatics tools (AI4AMP, AMPlify, AMP Scanner v2, amPEPpy, Macrel v.1.2.0, UNIPROT version 2022_05, CPAMR3, CAMPSign, Blastp) and the NCBI database. The candidate AMPs were predicted to be more globular and hemolytic than the remaining peptides in the secretome. There is evidence of transcription for 97% of the candidate AMPs in the same genus of ants, with one of them also verified as translated, thus supporting our findings. Most of these sequences with antimicrobial potential (94.8%) corresponded to transcripts from the ant's body, indicating their role not only as venom toxins.

Keywords: hymenoptera; antimicrobial peptides; ant peptides; *Odontomachus chelifer*.

LISTA DE FIGURAS

- Figura 1. *Odontomachus chelifer*. Cabeça (A), vista superior (B) e vista lateral (C).
Fonte: Damorin et al., 2019. 12
- Figura 2. Morfologia da formiga *Odontomachus monticola*. *Odontomachus monticola* (A), Glândula da peçonha da formiga *Odontomachus monticola* (B). Adaptado de Torres KAZUMA et al., 2017. 13
- Figura 3. Representação esquemática de alguns mecanismos de ação de PAMs. Modelo barril (A), Modelo de carpete (B) e Modelo de poro toroidal (C). A cor azul representa as porções hidrofóbicas dos PAMs, enquanto a cor vermelha representa as partes hidrofílicas dos PAMs. Fonte: Lima, 2021. 21
- Figura 4. Peptídeos analisados. Análise gráfica do número de peptídeos encontrados no transcriptoma total de *O. chelifer*. Os gráficos comparam a abundância de peptídeos previstos no transcriptoma total com o número de PAMs previstos no mesmo transcriptoma. Fonte: Menk et al., 2023. 32
- Figura 5. Análise gráfica do número de peptídeos encontrados no transcriptoma total de *O. chelifer*. Os gráficos comparam a abundância de peptídeos previstos no transcriptoma total com o número de PAMs previstos no mesmo transcriptoma. Fonte: Menk et al., 2023. 40
- Figura 6. Árvore filogenética dos peptídeos antimicrobianos (PAMs) previstos. Os PAMs previstos de *O. chelifer* e cinco PAMs de referência de outras espécies de formigas. Os peptídeos foram agrupados com base nas cinco sequências PAM de referência. Sequências dos demais grupos foram deixadas sem cor. Também estão disponíveis modelos de estrutura tridimensional para sequências pertencentes a diferentes grupos de peptídeos (A) semelhantes a crustina, (B) semelhantes a defensina, (C) semelhantes a himenoptaecina, (D) tipo ICK e (E) semelhantes a abaecina, seguindo o mesmo padrão de cores: roxo para α -hélice, amarelo para conformação β . As sequências usadas para previsão de estrutura estão listadas abaixo de cada modelo. Alinhamentos detalhados de todas as sequências de cada grupo da árvore, incluindo tecido de origem e seus peptídeos sinais. Fonte: Menk et al., 2023. 42

Figura 7. Alinhamento entre crustina e sequência de referência. Alinhamento global entre a crustina de referência e as sequências PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do PAM de referência. Fonte: Menk *et al.*, 2023. 43

Figura 8. Alinhamento entre defensina e sequência de referência. Alinhamento global entre a defensina de referência e as sequências de PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do PAM de referência. Fonte: Menk *et al.*, 2023. 44

Figura 9. Alinhamento entre himenoptaecina e sequência de referência. Alinhamento global entre a himenoptaecina de referência e as sequências PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do PAM de referência. Fonte: Menk *et al.*, 2023. 46

Figura 10. Alinhamento entre PAM do tipo ICK e sequência de referência. Alinhamento global entre o PAM do tipo ICK de referência e as sequências PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do PAM de referência. Fonte: Menk *et al.*, 2023. 47

Figura 11. Previsão da estrutura tridimensional do homólogo da toxina tipo ICK da poneratoxina putativa de *O. chelifer* (TRINITY_DN24839_c0_g1_i1.p1) Amarelo: conformação β e roxo: α -hélices. Fonte: Menk *et al.*, 2023. 48

Figura 12. Alinhamento de sequência primária entre duas sequências PAM previstas de *O. chelifer* e a abaecina A0A348G6C5 de *O. monticola*. Resíduos semelhantes estão em caixas. Resíduos idênticos estão em negrito. A linha vermelha indica os motivos PXP/PXXP. Fonte: Menk *et al.*, 2023. 49

Figura 13. Alinhamento entre abaecina e sequência de referência. Alinhamento global entre a abaecina de referência e as sequências PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do peptídeo de referência. Fonte: Menk *et al.*, 2023. 49

LISTA DE TABELAS

Tabela 1 - Peptídeos encontrados em formigas	14
Tabela 2 - Estudos de peçonha de formigas	16
Tabela 3 - Transcriptomas realizados em formigas	17
Tabela 4 - Medicamentos originados a partir de toxinas	23

Sumário

1. Introdução.....	10
2. Revisão de Literatura.....	13
2.1 Transcriptoma de formigas.....	13
2.2 Tipos de peptídeos em formigas.....	15
2.3 Peptídeos encontrados em peçonha.....	17
2.4 Peptídeos antimicrobianos (PAMs).....	19
2.5 Aplicações biotecnológicas da peçonha.....	22
3. Objetivos.....	24
3.1 Objetivo geral.....	24
3.2 Objetivos específicos.....	24
4. Metodologia.....	25
4.1 Transcriptoma da glândula da peçonha de <i>O. chelifer</i>	25
4.2 Extração e sequenciamento de RNA do corpo de <i>O. chelifer</i>	25
4.3 Controle de qualidade, pré-processamento e montagem de novo de <i>O. chelifer</i>	26
4.4 Obtenção de regiões protéicas codificadoras do transcriptoma de <i>O. chelifer</i>	27
4.5 Predição e anotação dos peptídeos que compõem o secretoma da glândula da peçonha de <i>O. chelifer</i>	27
4.6 Previsão de peptídeos antimicrobianos potencialmente secretados.....	28
4.7 Mapeamento cruzado.....	29
4.8 Teste estatístico.....	29
4.9 Agrupamento e filogenia de PAMs previstos.....	29
5. Resultados e discussão.....	31
5.1 Pré-processamento e montagem das sequências de <i>O. chelifer</i>	31
5.2 Peptídeos potencialmente secretados encontrados em <i>O. chelifer</i>	31
5.3 Atividade antimicrobiana prevista de peptídeos que podem atuar como toxinas na glândula da peçonha de <i>O. chelifer</i>	33
5.4 Filogenia de candidatos a PAM.....	35
5.5 PAMs podem compor o sistema imunológico de <i>O. chelifer</i>	43
5.6 Conservação da transcrição de PAMs candidatos no gênero <i>Odontomachus</i>	44
6. Conclusão.....	46
7. Perspectivas futuras.....	46
8. Referências.....	47
9 . Apêndices.....	63

1. Introdução

A maioria dos insetos Himenópteros (formigas, abelhas e vespas) desenvolveram aparelhos urticantes que são usados para injetar peçonha no seu alvo (PALMA, 2006). As formigas são animais peçonhentos e numerosos, apresentando uma distribuição ubíqua em ambientes terrestres (AILI et al., 2014; TOUCHARD et al., 2016). Sua diversidade se reflete em sua peçonha, que é composta por uma mistura de polipeptídeos, enzimas, alcalóides, aminas biogênicas, ácido fórmico e hidrocarbonetos, todos produzidos pela glândula de peçonha e armazenados no abdômen do animal no reservatório de peçonha (TOUCHARD et al., 2016). As peçonhas das espécies de formigas das subfamílias Ponerinae, Myrmicinae, Pseudomyrmecinae e Ecitoninae são geralmente ricas em proteínas e peptídeos (PALMA, 2006).

As peçonhas de himenópteros são coquetéis versáteis com estruturas bioquímicas que foram otimizadas pela evolução para selecionar alvos proteicos para defesa contra predadores ou patógenos, captura de presas e parasitismo (CASEWELL et al., 2013). No que diz respeito à sua natureza química, peçonhas são um *pool* extraordinário de toxinas e compostos não tóxicos, como sais, açúcares, ácido fórmico, aminas biogênicas, alcalóides, aminoácidos livres, hidrocarbonetos, peptídeos e proteínas. Evidências científicas sugerem que as toxinas participariam de uma série de processos biológicos, incluindo dor, paralisia, citólise, hemólise, reações alérgicas, resposta pró-inflamatória, atividades inseticidas e antimicrobianas (GUIDO-PATIÑO e PLISSON, 2022).

As peçonhas em insetos (aranhas, vespas, formigas, abelhas, centopéias e escorpiões) são promissores como uma das fontes mais ricas em PAMs (peptídeos antimicrobianos) (HAEBERLI et al., 2000; VASSILEVSKI et al., 2009; SANTOS et al., 2010; KUHN-NENTWIG et al., 2013; ZHANG, 2015; BORGES et al., 2016; GOPALAKRISHNAKONE et al., 2016; REIS et al., 2018). As peçonhas das formigas são formadas por substâncias orgânicas ricas em moléculas não proteicas, peptídeos lineares, diméricos e com ligações dissulfeto, proteínas bioativas, enzimas, alcalóides e hidrocarbonetos (LIMA, 2014; TOUCHARD et al., 2016).

Esses componentes orgânicos têm apresentado propriedades relevantes. Dentre eles ação contra células tumorais (VÉTILLARD e BOUZID, 2016); atividade antimicrobiana (ORIVEL, JÉRÔME et al., 2001; INAGAKI et al., 2004; ZELEZETSKY et al., 2005; WANANDY et al., 2015),

As formigas *Odontomachus* (BOLTON, 1995) Ponerinae compõem 77 espécies, 3 das quais são fósseis. Essas formigas vivem em partes subtropicais e tropicais do mundo e são fáceis de identificar devido a forma e ação cinética da abertura de mandíbula de 180°, que pode produzir cliques quando fechada (RAIMUNDO, FREITAS, OLIVEIRA, 2009). *Odontomachus chelifer* (Latreille, 1802) (figura 1), é uma formiga carnívora, com hábitos noturnos (LEMOS et al., 2020) e que pode ser encontrada no cerrado brasileiro e nas regiões neotropicais (KEMPF, 1972). Morfologicamente a *O. chelifer* possui uma cabeça que determina as subfamílias da espécie (BROWN, 1978; FERNANDES, 2017; GUIMARÃES, 2012). *O. chelifer* são consideradas formigas grandes (12 a 15mm) (GUIMARÃES et al., 2023), e sua picada causa uma dor aguda, queimação e coceira prolongada (SILVA, 2015; GUIMARÃES et al., 2023). O aparato da peçonha é composto por um ferrão, seguido da glândula Dufour, tubos secretórios e reservatório de peçonha. A peçonha é armazenada no reservatório que se encontra na gaster, localizado no último segmento do corpo da formiga e é associado a outras duas estruturas chamadas de glândula de Dufour e a glândula de peçonha (figura 2) (TORRES, 2013; LOPES, 2014).

Seus ninhos são formados sob camadas de material vegetal ou em galerias abandonadas por outras espécies de animais. Os hábitos alimentares de *O. chelifer* consistem em uma dieta oportunista de artrópodes mortos encontrados na região que habitam e de cupins mortos por ação de suas poderosas mandíbulas e peçonha (RAIMUNDO, FREITAS, OLIVEIRA, 2009).

Utilizando ferramentas de bioinformática, exploramos o conjunto de peptídeos encontrados no transcriptoma da formiga *O. chelifer*, verificando também evidências de tradução e transcrição em uma espécie congênica (*O. monticola*).



Figura 1. *Odontomachus chelifer*. Cabeça (A), vista superior (B) e vista lateral (C). Fonte: Damorin et al., 2019.

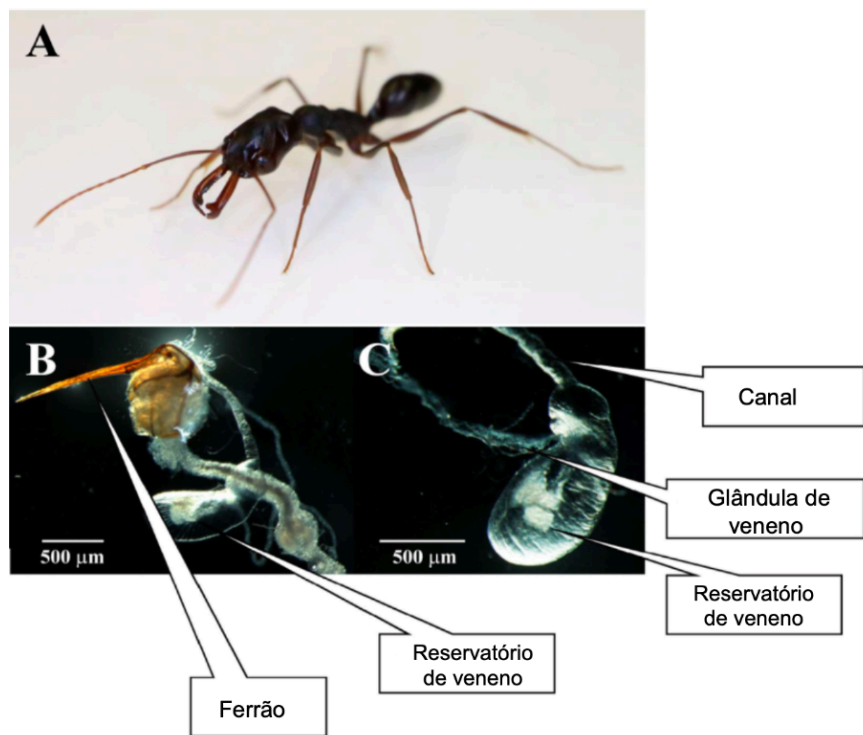


Figura 2. Morfologia da formiga *Odontomachus monticola*. *Odontomachus monticola* (A), Glândula da peçonha da formiga *Odontomachus monticola* (B). Adaptado de Torres KAZUMA et al., 2017.

2. Revisão de Literatura

2.1 Transcriptoma de formigas

Estudos genômicos e transcriptômicos são trabalhados por meio da bioinformática, área em que compreende uma análise sistemática de dados que engloba a biologia, estatística e a computação (CARVALHO, 2021). Estudos transcriptômicos identificam e caracterizam uma coleção de dados biológicos como sequências de RNA presentes em um tecido em um determinado momento (PACHECO, 2020; GODOY, 2021).

O transcriptoma é uma abordagem que pode fornecer informações valiosas para analisar a anotação de genoma, construir árvores filogenéticas, a evolução e a expressão gênica (HOLDING et al., 2018). Abordagens transcriptômicas demonstraram (tabela 1) ser eficazes para analisar a complexidade dos componentes das peçonha de formigas (GUIMARÃES et al., 2023) fornecendo informações a respeito dos genes e proteínas expressas.

Tabela 1 - Transcriptomas realizados em formigas

NOME CIENTÍFICO	TRANSCRIPTOMA	REFERÊNCIA
<i>Atta laevigata</i>	corpo	RODOVALHO et al., 2011
<i>Tetramorium bicarinatum</i>	glândula da peçonha	BOUZID et al., 2013
<i>Dinoponera quadriceps</i>	glândula da peçonha	TORRES et al., 2014
<i>Odontomachus monticola</i>	glândula da peçonha	KAZUMA et al., 2017
<i>Formica exsecta</i>	corpo	DHAYGUDE et al., 2017
<i>Prenolepis imparis</i>	corpo	TONIONE et al., 2020
<i>Paraponera clavata</i>	glândula da peçonha	AILI et al., 2020
<i>Cataglyphis nodus</i>	cérebro	HABENSTEIN et al., 2021
<i>Crematogaster rogenhoferi</i>	antenas	ZHOU et al., 2021
<i>Myrmica rubra</i> e <i>Myrmica ruginodis</i>	glândula da peçonha	HURKA et al., 2022
<i>Odontomachus chelifer</i>	glândula da peçonha	GUIMARÃES et al., 2023
<i>Lasius flavus</i>	glândula da peçonha	WANG et al., 2023

Wang e colaboradores (2023), identificaram no transcriptoma da glândula de veneno da formiga *Lasius flavus* novas toxinas peptídicas. Foram identificadas 17 proteínas, sendo sete proteínas que foram anotadas como fosfolipase B, proteínas

receptoras de odorante (Odorant binding protein, OBPs), apolipoproteína D, proteína *Takeout*, anidrase carbônica, *nose resistant to fluoxetine protein* e tioredoxina redutase (TrxR).

Recentemente, o transcriptoma da glândula da peçonha da *Odontomachus chelifer* possibilitou a investigação da atividade enzimática e funcional da peçonha e forneceu uma visão global dos transcritos expressos na glândula da peçonha. A peçonha bruta demonstrou atividade antiparasitária e de fosfolipase A2 (GUIMARÃES et al., 2023).

Hurka e colaboradores (2022) identificaram a composição das peçonhas de duas espécies de formigas (*Myrmica rubra* e *Myrmica ruginodis*) e identificaram novas toxinas da família semelhante ao fator de crescimento epidérmico (EGF) utilizando o transcriptoma. Descobriram 44 proteínas pertencentes a três famílias de proteínas, protease S1 na sua maioria (~84%), fosfatase (~9%) e a superfamília CAP (proteínas secretoras ricas em cisteína, antígeno 5 e proteínas relacionadas à patogênese).

O transcriptoma da formiga *Formica exsecta* (DHAYGUDE et al., 2017) possibilitou caracterizar os genes expressos utilizando indivíduos de diferentes estágios de vida assim podendo identificar informações mais abrangentes em relação ao número de genes expressos e o número de isoformas. Neste transcriptoma foi revelado DNA metiltransferases do tipo DNMT1 e DNMT3, proteínas enzimáticas modificadoras de DNA (KAUSAR et al., 2021).

O transcriptoma da glândula da peçonha e a composição da peçonha da formiga *Odontomachus monticola* também foi analisado. Foram identificados 92 peptídeos, proteínas, peptídeos semelhantes a pilosulina, proteína semelhante a quitinase, peptídeo semelhante a icarapina, proteína semelhante a hialuronidase, dipeptidil peptidase 4, quatro peptídeos semelhantes a inibidores ICK e três proteínas semelhantes a fosfolipase A (KAZUMA et al., 2017).

Outra formiga predadora popularmente conhecida como tocandira (*Dinoponera quadriceps*) teve o transcriptoma da glândula da peçonha analisado identificando novas sequências de toxinas que compreendem proteínas alergênicas, esterases (fosfolipases e carboxilesterase), toxinas letais e toxinas peptídicas com motivos ICK (inhibitory cystine knot). A sequências codificaram dois conjuntos, um de polipeptídeos mais abundantes (substâncias alergênicas, proteínas letais e

esterases) e outro conjunto de polipeptídeos menos abundantes (toxinas ricas em cisteínas conservadas) (TORRES et al., 2014).

Os genes transcritos da glândula da peçonha da espécie de formiga *Tetramorium bicarinatum* foi analisado por meio da montagem do transcriptoma e esse estudo permitiu analisar a expressão dos genes da peçonha. Foram identificados alérgenos da peçonha (alérgeno 3, Sol II e Sol IV), peptídeos do tipo pilosulina, fosfolipase A1 e A2, hialuronidase, uma toxina nomeada de secapina e outros tipos de enzima foram identificados como carboxipeptidases e inibidores de metaloproteinase (BOUZID et al., 2013).

Rodvalho e colaboradores (2011) analisaram os transcritos da formiga cortadeira *Atta laevigata*. Eles encontraram vinte e cinco enzimas do tipo citocromo P450, doze enzimas relacionadas a desintoxicação de xenobióticos incluindo glutatona S-transferase, glutamato-cisteína ligase e aldeído oxidase.

Existem muitos exemplos de famílias de peptídeos encontrados na peçonha de formigas, como poneratoxina, ponericina, ectatamina e pilosulina, todos com algum grau de atividade citolítica ou neurotóxica. Peptídeos da peçonha de formiga com propriedades perturbadoras de membrana, por exemplo, ponericanas, são geralmente associados a propriedades hemolíticas, antimicrobianas e/ou inseticidas. Sua ação como agentes disruptores de membrana pode permeabilizar as barreiras celulares, facilitando que os peptídeos neurotóxicos atinjam seus alvos. Os peptídeos de poneratoxina e ectatamina são tóxicos para células de vertebrados e invertebrados, principalmente por causa de suas habilidades neurotóxicas, que estão relacionadas ao bloqueio da transmissão sináptica pela modulação dos canais de sódio e cálcio dependentes de voltagem (AILI et al., 2014; TOUCHARD et al., 2016).

2.2 Tipos de peptídeos em formigas

Peçonha são ricas em peptídeos biologicamente ativos, comuns em himenópteros (PEREIRA, 2012; DAMORIN, 2019). Os peptídeos são classificados de acordo com o tamanho de sua cadeia proteica (cadeia curta ou longa), especificidade (toxinas de crustáceos, insetos e mamíferos), mecanismo de ação (citotoxinas ou neurotoxinas) e ainda pela presença ou ausência de ligações de dissulfeto (MAGALHÃES, 2017).

A maioria das toxinas presentes na peçonha dos himenópteros são geralmente semelhantes a peptídeos policationicos curtos e lineares, abundantes em α -hélices, responsáveis por ações antimicrobianas, liberação de histamina, lise celular e hemólise (PALMA, 2006; AILI et al., 2014; BOUZID et al., 2014)

Em peçonha de abelhas foram encontrados peptídeos com ação antimicrobiana e anticancerígena (melitina) (PALMA, 2006; KUNITZ, 2015) e neurotóxica (apamina) (PALMA, 2006; GU et al., 2020). Na peçonha de vespas foram encontrados peptídeos com atividade antimicrobiana, quimiotática e proteolítica (SAIDEMBERG, 2011). Há duas classes de peptídeos antimicrobianos que são mais encontrados em vespas, os mastoparanos (HIRAI et al., 1980; BAPTISTA-SAIDEMBERG, 2007; DE AZEVEDO, 2015; KIM et al., 2016; CHEN et al., 2018) que possuem atividade antitumoral contra células de melanoma (ABD EL-WAHED et al., 2021) e os quimiotáticos que promovem degranulação em mastócito (AGUIAR, 2006; BAPTISTA-SAIDEMBERG, 2007).

A diversidade de peptídeos encontrados na peçonha de formigas (tabela 2) tem sido investigada nas diversas subfamílias (Ectatomminae, Myrmeciinae, Pseudomyrmecinae, Paraponerinae e Ponerinae) (FONTANA et al., 2015). Peptídeos como pilosulinas, que são peptídeos antimicrobianos e os principais alérgenos foram encontrados na formiga *Myrmecia pilosula* (WU et al., 1998; INAGAKI et al., 2004; ZELEZETSKY et al., 2005; WIESE et al., 2007; WANANDY et al., 2015). Nas espécies *Paraponera clavata* e *Anochetus Emarginatus* foi identificado o neuropeptídeo poneratoxina (PIEK et al., 1991; PALMA, 2013; TOUCHARD et al., 2016), na espécie *Pachycondyla goeldii* foi caracterizado peptídeos conhecidos como ponerinas com propriedades antibacterianas e inseticidas (ORIVEL et al., 2001; PALMA, 2006). Fontana e colaboradores (2015), descreveram na espécie *Ectatomma tuberculatum* a neurotoxina ectatomina.

Himenópteros, especialmente as formigas, possuem moléculas biologicamente ativas com atividades farmacológicas atrativas para o desenvolvimentos de novos produtos farmacêuticos devido as suas atividades biológicas e estruturas moleculares (SOUZA, 2006; FONTANA et al, 2015).

Além dos peptídeos, proteínas, neurotoxinas com baixo peso molecular, como as poliaminas compõem a peçonha de himenópteros (PALMA, 2006; CORREA-VIANA, 2018).

Guimarães e colaboradores (2023) investigaram parcialmente as atividades enzimáticas e funcionais da glândula da peçonha da formiga *Odontomachus chelifer*.

Tabela 2 - Peptídeos encontrados em formigas

NOME CIENTÍFICO	PEPTÍDEOS	ATIVIDADE	REFERÊNCIA
<i>Paraponera clavata</i>	poneratoxina	neurotoxina	PIEK et al., 1991
<i>Myrmecia pilosula</i>	pilosulinas	antimicrobianos	WU et al., 1998
<i>Pachycondyla goeldii</i>	ponericinas	antibacterianas e inseticidas	ORIVEL et al., 2001
<i>Ectatomma tuberculatum</i>	ectatomina	neurotoxina	FONTANA et al., 2015.
<i>Dinoponera quadriceps</i>	dinoponeratoxinas	antimicrobiana	MONTEIRO, 2022.

2.3 Peptídeos encontrados em peçonha

Diversos grupos de animais possuem peçonha que podem ser utilizadas como fontes promissoras pois são compostos por misturas ricas em proteínas, peptídeos e moléculas orgânicas em função disso podem ser aplicados na medicina como agente terapêutico e na agricultura como bioinseticida por exemplo. (TOUCHARD et al., 2016; HURKA et al., 2022).

Os himenópteros são classificados em grupos sociais solitários. O grupo social possui peçonha, que são ferramentas de defesa para proteger a colônia de predadores (PALMA, 2006). A massa molar dos componentes peptídicos das peçonhas dos himenópteros variam de 1400 a 7000 Da, em sua grande maioria, são peptídeos anfipáticos policatiônicos lineares com alto teor de α -hélices em suas estruturas secundárias. Com mais de 154.000 espécies existentes, os himenópteros representam uma fonte inexplorada de secreções da peçonha e possivelmente uma diversidade química única com propriedades biológicas formidáveis (HUBER, 2017; GUIDO-PATIÑO e PLISSON, 2022).

Os peptídeos encontrados na peçonha de formiga geralmente seguem um padrão estrutural também encontrado nos peptídeos da peçonha de outros insetos himenópteros, sendo curtos, lineares e policatiônicos (PALMA, 2006). Esses peptídeos possuem alto teor de α -hélices e são responsáveis pela lise celular, hemólise, liberação de histamina, atividade neurotóxica e antimicrobiana (PALMA, 2006; TOUCHARD et al., 2016). Além dos peptídeos lineares, na peçonha de formigas também podem conter peptídeos estruturalmente diversos, como peptídeos do tipo ICK (*inhibitor cystine knot*), formado por ligações dissulfeto, que formam nós de cistina (ALLI et al., 2014).

A diversidade de peptídeos encontrados na peçonha de formigas mostra que eles representam "armas poderosas" que, quando injetadas em presas ou predadores, causam uma grande variedade de efeitos, garantindo a sobrevivência de formigas na maioria dos ecossistemas do mundo (PALMA, 2006). Esses peptídeos, quando isolados, podem ser utilizados em diversas aplicações biotecnológicas, como no tratamento de doenças (por exemplo, artrite e infecções), juntamente com a conservação de alimentos e controle de pragas na agricultura (AGARWAL et al., 2022).

Há décadas, estudos sobre as peçonhas das formigas vêm sendo realizados (tabela 3) utilizando as formigas *Paraponera clavata* (PIEK et al., 1991), *Myrmecia pilosula* (MATUSZEK et al., 1992), *Pachycondyla goeldii* (ORIVEL et al., 2001), *Neoponera goeldii* (SENETRA; NECELIS; CAPUTO, 2020), *Solenopsis saevissima* (FOX et al., 2010); *Solenopsis invicta* (CHEN, Li e FADAMIRO, Henry Y., 2009; HOFFMAN D. R, 2010), *Harpegnathos saltator* (HAIGHT, 2012), *Anochetus emarginatus* (TOUCHARD et al., 2016); *Tetramorium bicarinatum* (GUZMAN et al., 2017), *Manica rubida* (TOUCHARD et al., 2020), *Dinoponera quadriceps* (MONTEIRO et al., 2022).

Tabela 3 - Estudos de peçonha de formigas

NOME CIENTÍFICO	NOME POPULAR	ALVO DE ESTUDO	REFERÊNCIA
-----------------	--------------	----------------	------------

<i>Paraponera clavata</i>	Formiga cabo verde	Evidência de neurotoxina peptídica	PIEK et al., 1991
<i>Myrmecia pilosula</i>	'jack jumper'	Evidência de histamina	MATUSZEK et al., 1992
<i>Pachycondyla goeldii</i>		Peptídeos antibacteriano e inseticida	ORIVEL et al., 2001
<i>Neoponera goeldii</i> *		Atividade antimicrobiana	SENETRA; NECELI; CAPUTO; 2020
<i>Solenopsis saevissima</i>	lava pés	Estudo do mecanismo da peçonha	FOX et al., 2010
<i>Solenopsis invicta</i>	lava pés	Identificação de novos alcalóides	CHEN, Li; FADAMIRO, Henry Y., 2009
<i>Harpegnathos saltator</i>	formiga saltadora de jerdon	Produção da peçonha	HAIGHT, 2012
<i>Dinoponera quadriceps</i>	Tocandira	Caracterização de polipeptídeos	TORRES et al., 2014
<i>Dinoponera quadriceps</i>	Tocandira	Efeito antiparasitário	SOUSA et al., 2016
<i>Anochetus emarginatus</i>		Caracterização de uma poneratoxina	TOUCHARD et al., 2016
<i>Tetramorium bicarinatum</i>		Identificação e caracterização de peptídeo antimicrobiano	GUZMAN et al., 2017
<i>Manica rubida</i>		Caracterização de peptídeo	TOUCHARD et al., 2020
<i>Dinoponera quadriceps</i>	Tocandira	Atividade antifúngica	DODOU LIMA et al., 2020
<i>Dinoponera quadriceps</i>	Tocandira	Peptídeos antimicrobiano	MONTEIRO et al., 2022

2.4 Peptídeos antimicrobianos (PAMs)

Os PAMs são peptídeos que compõem o sistema imune dos animais, plantas e seres humanos e possuem características anfipática e catiônica. Em sua maioria

são moléculas pequenas e possuem a capacidade de combater bactérias, protozoários e vírus (COLE; LEHRER, 2003; TORRENT et al., 2013; ZASLOFF, 2002). Eles possuem atividade antibacteriana, antifúngica, antiviral e antiparasitária (STĄCZEK, CYTRYŃSKA, ZDYBICKA-BARABAS, 2023).

Os PAMs podem interagir com alvos intracelulares inibindo a síntese de RNA mensageiro; podem bloquear a síntese proteica e também dificultar o enovelamento conformacional das proteínas nativas (JENSSEN; HAMILL; HANCOCK, 2006), sendo capazes de agir como membranolíticos, imunomoduladores ou por interação com alvos intracelulares, incluindo ácido nucleico e maquinaria de síntese proteica e apoptose. Esses peptídeos podem contribuir para desenvolver fármacos, pois possuem amplo espectro de atividade antimicrobiana que age contra bactérias, fungos e vírus por apresentar um complexo mecanismo de ação que pode reduzir a chance de resistência (CAVALCANTE et al., 2016; GREBER, DAWGUL, 2017; LI et al., 2012; SIERRA et al., 2017; SORRUMA et al., 2020).

Peptídeos antimicrobianos (PAMs) derivados da peçonha de formigas podem ser usados como um antibiótico alternativo para tratar infecções causadas por microrganismos (TONK; VILCINSKAS, 2016). Isso se deve principalmente à sua atividade de amplo espectro, visando bactérias Gram-positivas e Gram-negativas, e ademais, serem utilizados para o tratamento de doenças causadas por fungos, vírus e parasitas, permitindo seu uso em uma variedade de tratamentos de doenças (LI et al., 2012). O principal modo de ação dos PAMs é romper a membrana do alvo logo após a ligação a ela (figura 3). Além disso, em alguns casos, eles são capazes de criar poros na membrana, causando lise celular ou vazamento. Alguns PAMs podem ter alvos componentes intracelulares (ácidos nucleicos e proteínas) (AMORIM-CARMO et al., 2022).

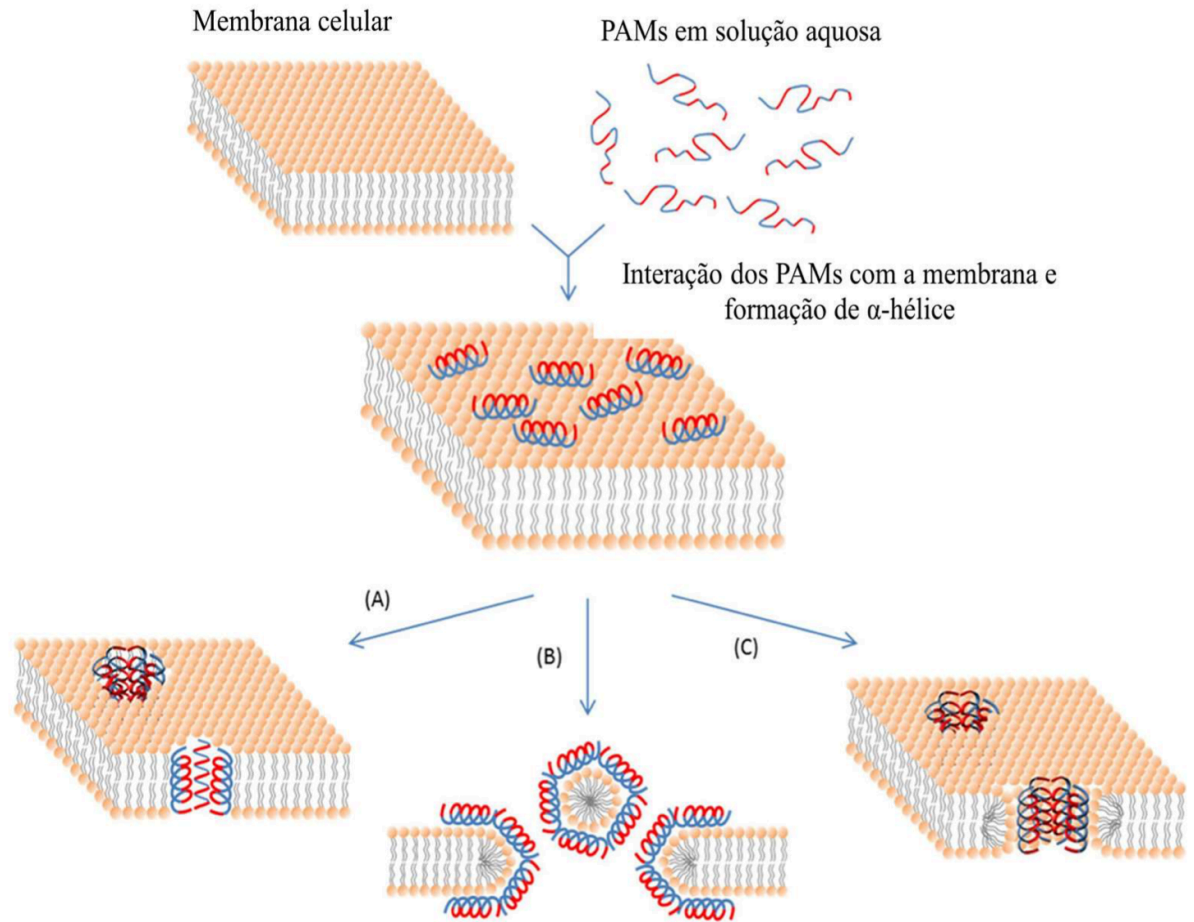


Figura 3. Representação esquemática de alguns mecanismos de ação de PAMs. Modelo barril (A), Modelo de tapete (B) e Modelo de poro toroidal (C). A cor azul representa as porções hidrofóbicas dos PAMs, enquanto a cor vermelha representa as partes hidrofílicas dos PAMs. Fonte: Lima, 2021.

Em toxinas de formiga normalmente são encontrados PAMs como: bicarinalinas (*Tetramorium bicarinatum*), dinoponeratoxinas (*Dinoponera quadriceps*), pilosulinas (*Myrmecia pilosula*), ponerinas (*Neoponera goeldii*) e pilosulinas (*Myrmecia pilosula*) (DODOU LIMA, DE PAULA CAVALCANTE, RÁDIS-BAPTISTA, 2020).

A bicarinalina, é um peptídeo encontrado em formigas *Tetramorium bicarinatum*, que demonstrou atividade semelhante aos antibióticos já utilizados para o tratamento de infecções por *Helicobacter pylori* (GUZMAN, et al., 2017). As dinoponeratoxinas são peptídeos antimicrobianos identificados na glândula da peçonha da formiga *Dinoponera quadriceps* (TORRES et al, 2014).

Outro exemplo de PAMs polivalentes são as pilosulinas, marcadas como um dos principais peptídeos alergênicos biologicamente ativos em formigas (INAGAKI, et al., 2004; WIESE et al., 2007; WANANDY, 2015). Esta classe de peptídeos exibe uma ampla gama de atividades, como atividade antimicrobiana e hemólise (DONOVAN; BALDO, 2001; KING et al., 1998; WU et al., 1998; ZELEZETSKY et al., 2005).

A análise peptidômica da peçonha bruta da formiga *Dinoponera quadriceps* foi investigado (TORRES, 2013) e seus efeitos biológicos foram testados abordando o potencial antimicrobiano e antiparasitário (LIMA, 2014), atividade repelente (DOS SANTOS ARAUJO, 2019) e ação antifúngica de peptídeos antimicrobianos (LIMA, 2021).

Recentemente, foi investigada a atividade enzimática e funcional da peçonha da glândula da *O. chelifera* (GUIMARÃES et al., 2023), cuja composição e propriedades, até então, eram desconhecidas. A peçonha bruta de *O. chelifera* apresentou atividade antiparasitária e de fosfolipase A2. Neste estudo, também foi realizada o sequenciamento de RNA (Illumina), em que foram identificados os 30 genes mais expressos da glândula da peçonha e transcritos importantes envolvidos na função da peçonha como a U-poneritoxina (01), Om1-like (pilosulina), quitinase 2, alérgeno 3, quimotripsina (1 e 2) e glutathione S-transferase.

2.5 Aplicações biotecnológicas da peçonha

A diversidade bioquímica das toxinas presentes nas peçonhas de formigas revelam ser uma fonte promissora e muito potente para aplicações biotecnológicas. A caracterização da composição bioquímica e funcional das toxinas da peçonha bruta permite a identificação de moléculas potenciais para o tratamento de doenças.

As moléculas presentes nas peçonhas possuem propriedades terapêuticas e biotecnológicas. Registros utilizando anticorpos terapêuticos estão mudando a indústria com o reconhecimento das moléculas biológicas, particularmente os peptídeos que estão sendo validados como inovações terapêuticas (ESCOUBAS et al., 2009).

Ao longo das décadas foram investigados substâncias presentes na peçonha (tabela 4) e medicamentos foram desenvolvidos. Um exemplo de sucesso foi o captopril, molécula sintética antihipertensiva que simula os peptídeos da bradicinina, descobertos no veneno da cobra brasileira *Bothrops jararaca*.

Tabela 4 - Medicamentos originados a partir de toxinas

Produto	Composto	Origem	Referência
Captopril	Bradicinina	<i>Bothrops jararaca</i>	FERREIRA, 1965
Cola de fibrina	Serino proteínase	<i>Crotalus durissus</i>	BARROS, 2009
Ziconotida (Prialt)	ω - conopéptido	<i>Conus magus</i>	MILJANICH, 2004
Ozempic	Semaglutida	<i>Heloderma suspectum</i>	GOLDENBERG, 2019
Botox	Toxina botulínica	<i>Clostridium botulinum</i>	DE MELLO SPOSITO, 2004

Novas moléculas farmacologicamente ativas podem desempenhar um papel essencial para a descobertas de medicamentos, mas a maioria dos animais peçonhentos são de pequeno porte, assim dificultando a quantidade de venenos disponíveis para aplicação de técnicas para uma investigação completa porém as abordagens transcriptômicas e proteômicas estão se revelando eficazes para desvendar a complexidade dos componentes das peçonhas (COLOGNA et al., 2013).

A demanda por novas moléculas farmacologicamente ativas leva ao desenvolvimento de medicamentos baseados em outros materiais, como por exemplo os peptídeos. Os peptídeos da peçonha tem especificidades que podem desenvolver produtos biofarmacêuticos, pois suas atividades biológicas e farmacológicas têm se mostrado eficientes e o mercado global está expandindo (DOS SANTOS et al., 2020; BIN et al., 2022).

Dadas as possíveis aplicações de peptídeos antimicrobianos, e o fato das formigas possuírem moléculas promissoras, há ainda a necessidade de enfatizar as investigações das atividades funcionais e identificar quais são as moléculas presentes no arsenal de toxinas das formigas.

Neste estudo, realizamos análises transcriptômicas de novos peptídeos e proteínas de *Odontomachus chelifer*, uma espécie até então não analisada quanto a este aspecto.

3. Objetivos

3.1 Objetivo geral

Investigação *in silico* dos possíveis peptídeos antimicrobianos presentes na formiga *Odontomachus chelifer*.

3.2 Objetivos específicos

Utilizando-se ferramentas de bioinformática para analisar o transcriptoma da formiga *Odontomachus chelifer*, pretende-se:

- Obter regiões proteínicas codificadoras do transcriptoma de *O. chelifer*;
- Prever e anotar peptídeos que compõem o secretoma da glândula da peçonha de *O. chelifer*
- Prever peptídeos antimicrobianos potencialmente secretados;
- Mapear os transcritos da glândula da peçonha da *O. chelifer* contra os transcritos montados do corpo.

4. Metodologia

4.1 Transcriptoma da glândula da peçonha de *O. chelifera*

Os dados transcriptômicos da glândula da peçonha de *O. chelifera* foram previamente gerados por Guimarães e colaboradores (2023) e depositados no banco de dados GenBank SRA. A análise do transcriptoma de *O. chelifera* (extração de RNA, sequenciamento de Illumina, controle de qualidade, pré-processamento e montagem de novo) foi descrita por Guimarães e colaboradores (2023).

4.2 Extração e sequenciamento de RNA do corpo de *O. chelifera*

O RNA total do corpo de *O. chelifera* foi isolado usando o Reagente TRIzol (Invitrogen). A amostra foi macerada por aproximadamente 2 a 3 minutos usando um pistilo e a extração foi realizada de acordo com as instruções do fabricante para o reagente TRIzol. O RNA foi analisado no espectrofotômetro Nanodrop ND-1000 (Thermo Scientific) e sua integridade verificada em gel de agarose 1%. O material (~3 µg de RNA) foi tratado com DNase I livre de RNase (Invitrogen) em uma reação de volume final de 20 µL. As bibliotecas foram geradas utilizando os kits de preparação de pequenas bibliotecas de RNA TruSeq da Illumina. Os mRNAs foram isolados usando esferas de oligo (dT) e por meio de incubação com um tampão de fragmentação 1x fornecido com o kit de preparação para amostras pequenas de RNA Illumina. Os fragmentos foram utilizados como moldes em reações com transcriptase reversa SuperScript II (Invitrogen) e primers de hexâmeros randômicos para obter o cDNA de primeira fita, seguido de uma reação com DNA polimerase I e RNase H (Illumina) para gerar o cDNA de fita dupla. Os fragmentos foram reparados nas extremidades, as extremidades 3' foram adeniladas e os adaptadores foram ligados. Fragmentos apropriados (200 ± 30 pb) foram isolados por eletroforese em gel de agarose, purificados e enriquecidos por reação em cadeia da

polimerase (PCR) usando primers adaptadores específicos. A qualidade das bibliotecas foi validada usando um Agilent 2100 Bioanalyzer (Agilent Technologies). Por fim, as bibliotecas foram sequenciadas usando HiSeq1000 (Illumina) em leituras pareadas (2 x 100 pb), de acordo com as instruções do fabricante.

4.3 Controle de qualidade, pré-processamento e montagem *de novo* de *O. chelifer*

O programa FASTQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) foi utilizado para realizar o controle de qualidade do transcriptoma do corpo de *O. chelifer*. Um fenômeno comum na plataforma Illumina é conhecido como um desvio no conteúdo do GC (CONESA, et al., 2016; HANSEN, BRENNER, DUDOIT, 2010). Para resolver este problema, os primeiros 15 nucleotídeos de todas as leituras foram cortados usando a ferramenta `fastx_trimmer` disponível no FATXToolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). Os últimos 4 pb também foram removidos usando o mesmo programa. SeqyClean (ZEBANNIKOV, ARBEEV, YASHIN, 2017) foi usado para filtrar bases de baixa qualidade (Phred Q score <30). A montagem de novo foi realizada com Trinity v2.14.0 (GRABHERR, et al., 2011) usando parâmetros padrão. Com o objetivo de fornecer uma visão quantitativa e abrangente do nível de completude de nossa montagem, usamos o programa Benchmarking Universal Single-Copy Orthologs (BUSCO) v5 (WATERHOUSE, et al., 2018).

A montagem foi comparada com um conjunto predefinido de bancos de dados Insecta e as sequências foram categorizadas como “completa, cópia única”, “completa, cópia duplicada”, “fragmentada” ou “ausente”, de acordo com a sequência alinhada.

4.4 Obtenção de regiões protéicas codificadoras do transcriptoma de *O. chelifer*

O total de 50.220 transcritos da glândula da peçonha de *O. chelifer* foi usado para obter as regiões codificantes de proteínas usando TransDecoder v5.5.0 (HAAS, et al., 2013) na plataforma OmicsBox v3.0.25 (BioBam, Valência, Espanha) (GOTZ, et al., 2008). As proteínas foram filtradas por tamanho até um máximo de 200 aminoácidos.

4.5 Predição e anotação dos peptídeos que compõem o secretoma da glândula da peçonha de *O. chelifer*

Foram preditas 125.123 proteínas a partir dos transcritos da glândula da peçonha de *O. chelifer*. Os peptídeos sinal foram previstos usando Razor v.1 (BHANDARI, GARDNER, LIM, 2021), deixando apenas peptídeos com locais de clivagem previstos e caindo na faixa de 8 a 98 aminoácidos. Esse limite de comprimento preciso foi definido para permitir operações downstream de aprendizado de máquina que às vezes requerem peptídeos nesses intervalos. Os 1.022 peptídeos restantes, editados em suas sequências maduras, foram pesquisados no banco de dados AntiFAM v.7 (EBERHARDT, et al., 2012) com HMMER v.3.3.2 (MISTRY, et al., 2013) para eliminar sequências espúrias.

A atividade hemolítica dos peptídeos secretados foi predita usando Macrel v.1.2.0 (SANTOS -JÚNIOR, et al., 2020) e a desordem junto com os domínios globulares foram computados usando IUpred3.0 (ERDOS, PAJKOS, DOSZTÁNYI, 2021), este último não usando procedimentos estáveis. O secretoma previsto foi anotado usando eggNOG-mapper v.2.1.10 (CANTALAPIEDRA et al., 2021) para avaliar as classes de COG e traçar um perfil de metabolismo.

4.6 Previsão de peptídeos antimicrobianos potencialmente secretados

Os peptídeos previstos do secretoma da glândula da peçonha de *O. chelifer* foram submetidos a diferentes sistemas de previsão de peptídeos antimicrobianos (PAMs): (i) aqueles baseados em redes neurais: AI4PAM (LIN et al., 2021), AMPLify (LI et al., 2022) e AMP Scanner v2 (VELTRI, KAMATH, SHEHU, 2018) e (ii) aqueles baseados em florestas aleatórias: amPEPpy (LAWRENCE et al., 2021), usando o modelo para sequências maduras (FINGERHUT et al., 2020) e Macrel v.1.2.0 (SANTOS-JÚNIOR et al., 2020). O resultado final para cada peptídeo foi obtido através do sistema de votação, onde os peptídeos designados como PAMs candidatos por três ou mais sistemas foram mantidos e os demais descartados. Para verificar se os PAMs candidatos teriam maior probabilidade de causar hemólise - uma justificativa plausível para sua presença na peçonha - foi aplicado o teste Exato de Fisher nas contagens dos PAMs candidatos e os demais peptídeos do secretoma foram organizados de acordo com sua atividade hemolítica. A razão de chances foi então calculada.

Os PAMs candidatos foram anotados usando ferramentas online do UNIPROT versão 2022_05 e CPAMR3 (WAGHU et al., 2016), o servidor web CAMPSign (<http://www.campsign.bicnirrh.res.in/>) com seu banco de dados Hidden Markov Models (HMM) de famílias PAM e o modo Blast completo contra vários bancos de dados PAM diferentes usando um valor E máximo de 10^{-5} . A anotação também prosseguiu localmente usando Blastp (CAMACHO et al., 2009) contra a versão 254 (SAYERS et al., 2023) do banco de dados NCBI e analisando os resultados para aqueles com uma identidade $\geq 90\%$, sem lacunas abertas, alinhamentos desde o início do peptídeo e novamente um E-valor máximo de 10^{-5} , mantendo apenas o melhor acerto para cada sequência.

4.7 Mapeamento cruzado

Inferimos a presença desses PAMs previstos no corpo de *O. chelififer* (APÊNDICE B) através do mapeamento de transcritos da glândula da peçonha de *O. chelififer* contra os transcritos montados do corpo usando BlastN (ATSCHUL et al., 2023) na plataforma OmicsBox. A evidência da tradução dos PAMs previstos foi reunida pesquisando as sequências peptídicas contra um banco de dados de peptídeos *O. monticola* detectados por LC-ESI-MS (ATSCHUL et al., 2023). Usando o PALADIN (WESTBROOK et al., 2017), mapeamos os peptídeos potencialmente secretados contra o transcriptoma da glândula da peçonha de *O. chelififer* em termos de leitura de aminoácidos, para verificar a abundância de peptídeos potencialmente antimicrobianos versus o secretoma residual da glândula da peçonha de *O. chelififer*. Também usando PALADIN (WESTBROOK et al., 2017), mapeamos as leituras de aminoácidos do transcriptoma da glândula de *O. monticola* contra os peptídeos antimicrobianos previstos de *O. chelififer*, inferindo não apenas sua presença em diferentes espécies, mas quantificando sua abundância entre espécies.

4.8 Teste estatístico

As diferenças entre proporções foram verificadas por meio do teste Exato de Fisher e as diferenças entre medianas foram avaliadas por meio do teste U de Mann-Whitney, ambos implementados no pacote Scipy 1.9.1 (VIRTANEN et al., 2020) do Python 3.9. O limite de significância foi estabelecido como $P < 0,05$. As visualizações gráficas das distribuições foram usando os pacotes Seaborn (WASKOM, 2021) e (HUNTER, 2007).

4.9 Agrupamento e filogenia de PAMs previstos

Com base no trabalho de Zhang e Zhu (2012), incluímos diferentes sequências de PAM de cinco famílias amplamente presentes em táxons de formigas

nos PAMs previstos para comparar suas características e possivelmente inferir suas atividades/mecanismos. Estes incluíram uma defensina 2 (A0A348G6A9), himenoptaecina (A0A348G6C2) e abaecina (A0A348G6C5) de *Odontomachus monticola*; uma poneritoxina semelhante a ICK (P0DSL6) de *Dinoponera quadriceps*; e uma crustina 4 (A0A0J7L9H5) de *Lasius niger*. As sequências adicionais tiveram seus peptídeos sinal removidos antes das operações a jusante. Todos os peptídeos foram então agrupados usando CDHIT (FU et al., 2012) com 100% de identidade e 90% de sobreposição da sequência mais curta (parâmetros: -c 1 -aS .9 -G 0 -l 5"). Em seguida, as sequências foram alinhadas usando MAFFT (KATO et al., 2002) no modo 'auto' com o alinhamento posterior aparado usando ClipKIT (STEENWYK et al., 2020). A árvore filogenética foi produzida com FastTree2 (PRICE, DEHAL, ARKIN, 2010), usando pseudocontas para corrigir a esparsidade, estimando as taxas de transição usando a distribuição gama e estimando a verossimilhança máxima do modelo Le Gascuel - LG (LE, LARTILLOT, GASCUEL, 2008) implementado. Para estimar a distribuição de recursos e a proximidade entre as sequências potencialmente secretadas e antimicrobianas, os mesmos 22 características bioquímicas (por exemplo, pl, carga e índice de Boman) usados por Macrel (SANTOS-JÚNIOR et al., 2020) para a previsão de PAM foram calculados e comparados como funções de densidade. Após os alinhamentos, e mantendo a maioria dos domínios de referência, as sequências agrupadas para referências na árvore filogenética seguiram procedimentos downstream. ESPript (<https://esript.ibcp.fr>) (ROBERT, GOUET, 2014) foi usado para editar e visualizar alguns alinhamentos. As previsões de estruturas secundárias foram geradas usando PSIPRED 4.0 (BUCHAN, JONES, 2019). As previsões da estrutura tridimensional

do peptídeo foram feitas com ColabFold v1.5.2: AlphaFold2 usando MMseqs2 (MIRDITA et al., 2022).

5. Resultados e discussão

5.1 Pré-processamento e montagem das sequências de *O. chelififer*

O sequenciamento Illumina do corpo de *O. chelififer* gerou 63.012.700 *reads*. Após a análise de pré-processamento, descartamos 15,7% dos *reads* iniciais e obtivemos 52.469 contigs em uma montagem *de novo* com o programa *Trinity*. A montagem de 93,42% foi concluída, de acordo com nossa análise BUSCO contra *Insecta_db10*, enquanto faltavam 4,02% dos grupos BUSCO e 2,56% foram apenas parcialmente recuperados. Kazuma e colaboradores (2017) geraram utilizando o sequenciamento Illumina da glândula de veneno da *O. monticola* 44.087.058 *reads*. Hurka e colaboradores (2022) investigaram os componentes do veneno das formigas *Myrmica rubra* e *Myrmica ruginodis* e geraram 69.645.529 e 69.318.655, *reads* respectivamente. Wang e colaboradores (2023) construíram um banco de dados transcriptômicos da formiga amarela *Lasius flavus* que gerou 41.335.556 *reads*.

5.2 Peptídeos potencialmente secretados encontrados em *O. chelififer*

Após a montagem do transcriptoma da glândula da peçonha de *O. chelififer*, obtivemos 125.123 proteínas dele deduzidas, das quais pudemos prever que 1.819 possuem peptídeos de sinal. Depois de filtrar as sequências maduras para aquelas com comprimentos entre 8 e 98 aminoácidos, as sequências compunham 1.022 peptídeos (Figura 4 e APÊNDICE A). Apenas 1 peptídeo atingiu o AntiFAM v.7 (EBERHARDT et al., 2012) com um domínio incompleto e foi descartado dos processos *downstream*. As sequências tiveram uma desordem média de $0,32 \pm 0,27$,

com 366 dos 1.022 peptídeos avaliados sendo previstos para conter domínios globulares. Desarranjo acima de 0,45 indicam possíveis proteínas espúrias ou estruturas de proteínas desordenadas que precisam principalmente de ajuda no dobramento (NECCI, PIOVESAN, TOSATTO, 2016; ERDOS, PAJKOS, DOSZTÁNYI, 2021). Curiosamente, uma minoria das sequências previstas (282/1.022 peptídeos) foi atribuída a esse peptídeo sinal.

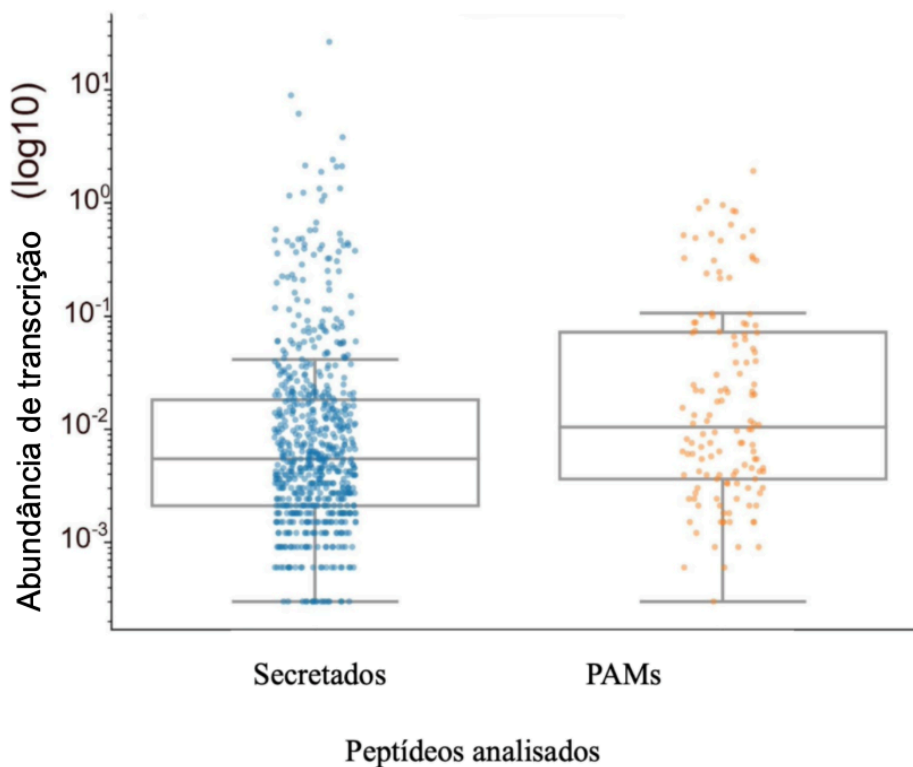


Figura 4. Peptídeos analisados. Análise gráfica do número de peptídeos encontrados no transcriptoma total de *O. chelifera*. Os gráficos comparam a abundância de peptídeos previstos no transcriptoma total com o número de PAMs previstos no mesmo transcriptoma. Fonte: Menk et al., 2023.

Kazuma e colaboradores (2017) após a montagem do transcriptoma da glândula da peçonha de *O. monticola*, obtiveram 49.639 sequências, do qual 92 codificam peptídeos e proteínas semelhantes à toxina. Nas formigas *Myrmica rubra* e *Myrmica ruginodis* (Hurka et al., 2022) leituras foram agrupadas 512.855 (*M. rubra*) e 466.158 (*M. ruginodis*) e anotaram 255 de 288 (*M. rubra*) e 348 de 381 (*M.*

ruginodis) de candidatos á peptídeos utilizando o Blast. Wang e colaboradores (2023) tiveram 42.047 predições, sendo anotados 219 transcritos.

Apenas uma pequena fração (250 de 1.022 peptídeos) das sequências maduras potencialmente secretadas da glândula da peçonha de *O. chelifera* foi anotada com o banco de dados Egnog v.5 (HUERTA-CEPAS et al., 2019), possivelmente devido à baixa homologia entre os peptídeos de formiga e o comprimento completo no banco de dados. Apesar da pequena representatividade dessas proteínas anotadas, foi possível observar a categoria mais representada, que é a classe metabólica desconhecida (58/250 peptídeos). De fato, a categoria de COG mais abundante com função definida entre o transcriptoma previsto foi a categoria de mecanismos de transdução de sinal (39/250 peptídeos), seguida pela modificação pós-traducional e categoria chaperonas (28/250 peptídeos), a categoria de tradução (14/250 peptídeos) e o íon inorgânico categoria de transporte e metabolismo (11/250 peptídeos).

5.3 Atividade antimicrobiana prevista de peptídeos que podem atuar como toxinas na glândula da peçonha de *O. chelifera*

Usando um sistema de votação com seis classificadores de PAM diferentes, previmos um total de 136 PAMs candidatos dos transcritos da glândula da peçonha de *O. chelifera*, representando 112 peptídeos não redundantes. Não houve diferença significativa ($P_{\text{Mann-Whitney's U}} = 0.92$) entre a desordem dos PAMs candidatos e o transcriptoma restante (o transcriptoma previsto exceto para os PAMs previstos). No entanto, os PAMs candidatos tendem a ser mais globulares do que os peptídeos restantes (razão = 3.5, $P_{\text{Fisher's Exact}} = 4.6 \times 10^{-11}$) o que sugere que nossos PAMs candidatos podem ser mais solúveis.

Os PAMs previstos foram mais longos do que o secretoma médio, com medianas de 60,5 e 28 ($P_{\text{Mann-WhitneyU}} = 4 \times 10^{-19}$, Figura 5A), respectivamente. Também descobrimos que o comprimento médio dos PAMs (Q50 = 19) de bancos de dados públicos (por exemplo, DRPAM, APD3, dbPAM) é ainda menor que o secretoma médio, indicando que os PAMs candidatos de *O. chelifera* podem ser naturalmente mais longos. Curiosamente, PAMs mais longos geralmente causam a formação de poros na membrana bacteriana, enquanto os mais curtos promovem a micelação da membrana (LI et al., 2021).

PAMs candidatos de *O. chelifer* também foram carregados mais positivamente ($pI_{Q50} = 10$) do que o secretoma total ($pI_{Q50} = 7.6$, $P_{MannWhitneyU} = 2.2 \times 10^{-10}$) (Figura 5B e 5C). Cargas positivas são essenciais para a capacidade dos PAMs de se ligarem à membrana carregada de bactérias (LI et al., 2021), o que implica uma modulação de sua atividade antimicrobiana com base no pH (LI et al., 2021) e pode indicar o melhor ambiente para sua ação, como a própria peçonha ou outro local do corpo. Ao analisar a propensão dos PAMs previstos para se ligarem a outras proteínas via índice de Boman (HANSEN et al., 2020), não observamos nenhuma diferença ($P_{MannWhitneyU} = 0.39$) entre os PAMs previstos e o secretoma restante (Figura 5D). Isso sugere que a maioria dos peptídeos do secretoma, incluindo os PAMs, realizam interações proteína-proteína, dada a probabilidade de densidade ter picos em valores altos ($> 2,1$).

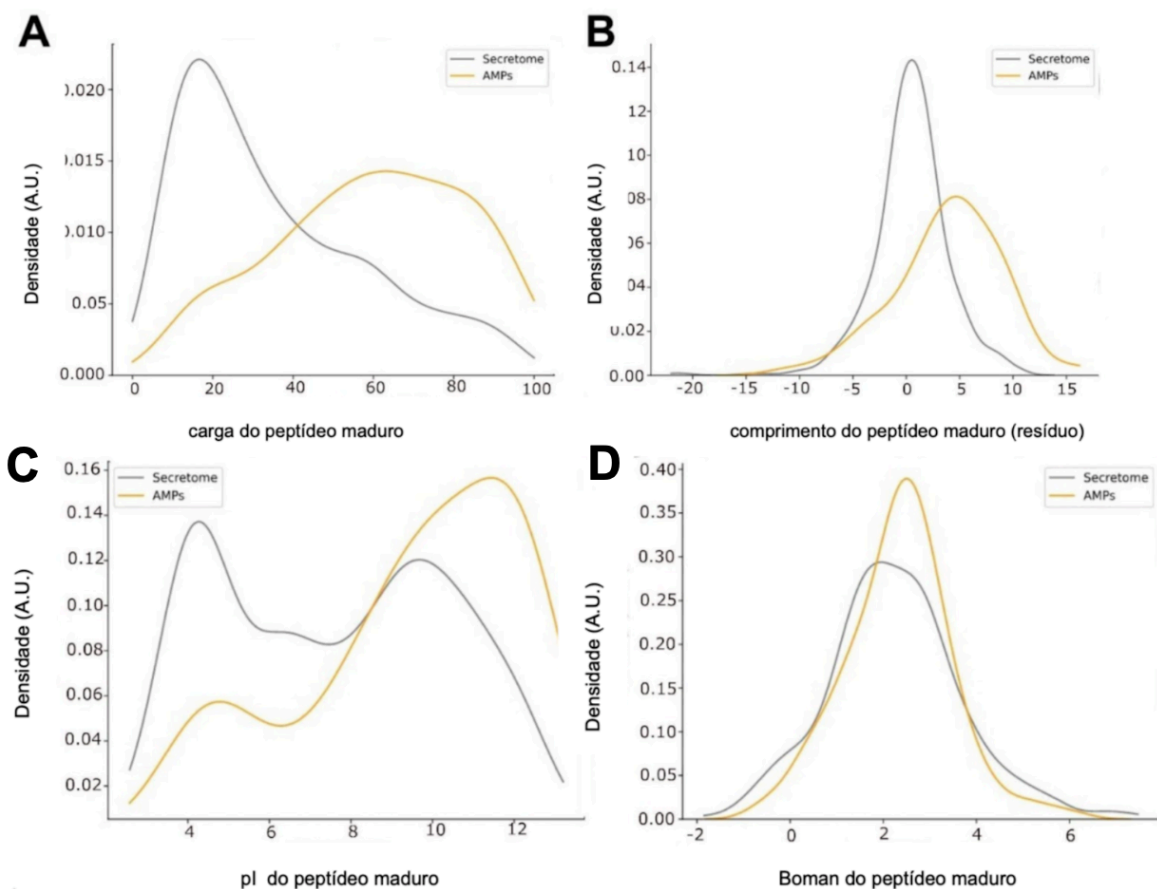


Figura 5. Análise gráfica do número de peptídeos encontrados no transcriptoma total de *O. chelififer*. Os gráficos comparam a abundância de peptídeos previstos no transcriptoma total com o número de PAMs previstos no mesmo transcriptoma. Fonte: Menk et al., 2023.

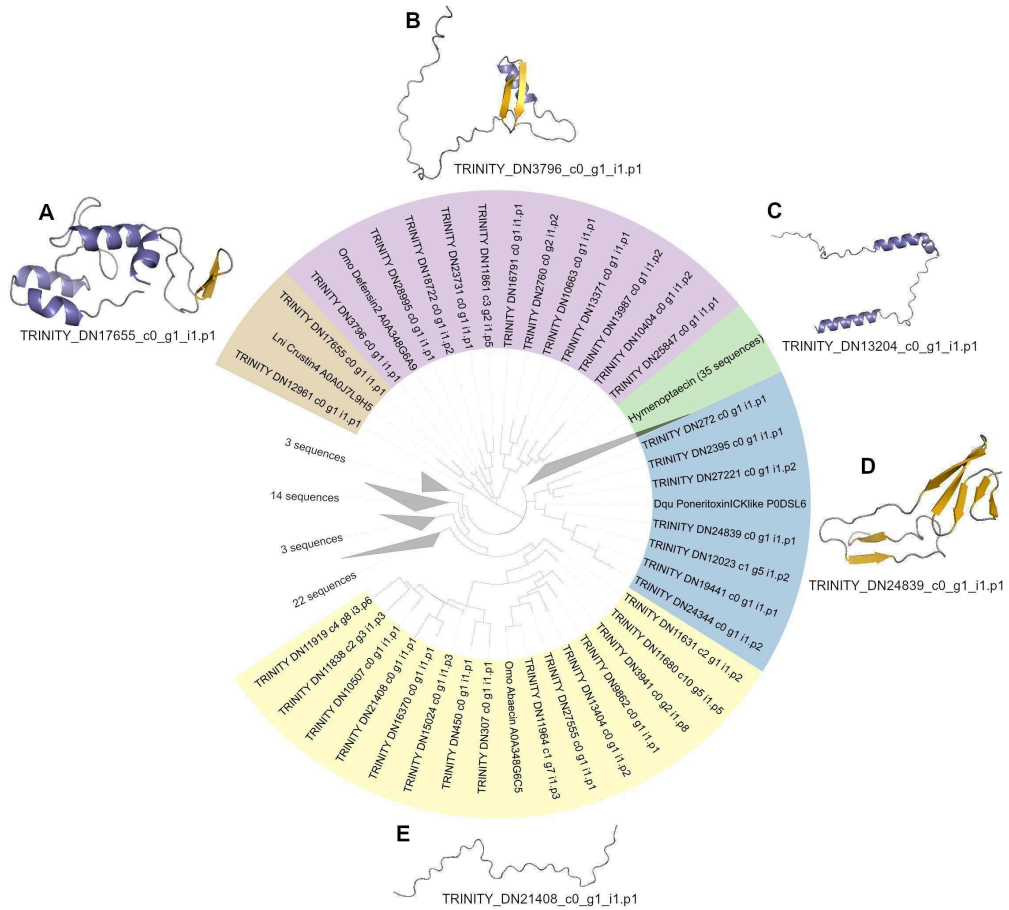
A peçonha é um coquetel complexo de compostos e geralmente apresentam diversidade funcional. Isso pode ser devido à multifuncionalidade da toxina, na qual a mesma toxina atinge múltiplas funções, por exemplo, estar presente no alvo. As membranas celulares são provavelmente os alvos potenciais mais onipresentes para uma toxina, muitas vezes mostrando propriedades inseticidas, hemolíticas e/ou antimicrobianas associadas (TOUCHARD et al., 2016; SCHENDEL et al., 2019). Baseado em resultados utilizando as ferramentas de bioinformática (MACREL v.1.2.0), foi predito que os PAMs candidatos eram mais propensos a causar hemólise (razão = 4.45, $P_{\text{Fisher's Exact}} = 2.3 \times 10^{-9}$) do que o secretoma restante, o que pode ser uma justificativa plausível para sua presença na peçonha. Um possível papel das hemolisinas na peçonha de insetos himenópteros é a indução de dor em vertebrados como defesa contra predadores. Essas estratégias também são descritas em abelhas e formigas que utilizam sua peçonha apenas para defesa (AILI, TOUCHARD, 2014).

Obtivemos cruzamentos insuficientes (identidade <75%) com sequências conhecidas ao anotar os PAMs candidatos usando vários bancos de dados (por exemplo, Uniprot, NCBI, CPAMR3). A única exceção foi o candidato TRINITY_DN3796_c0_g1_i1.p1 que foi atribuído à família de defensinas por meio de Hidden Markov Models com a ferramenta CPAMsign (<http://www.cPAMsign.bicnirrh.res.in/>). Ao usar outras ferramentas de anotação, como Blastp, observamos alguns acertos (29/136) (identidade > 90% e valor E < $1,0 \times 10^{-5}$) contra proteínas principalmente de comprimento total de *O. brunneus* (22/29), indicando uma montagem incompleta dessas entidades ou um truncamento em *O. chelififer*.

5.4 Filogenia de candidatos a PAM

Observamos a relação filogenética entre os PAMs previstos e cinco peptídeos antimicrobianos de referência de outras espécies (figura 6). Os PAMs de referência foram escolhidos com base no trabalho de Zhang e Zhu (2012), no qual uma análise comparativa do genoma de sete espécies diferentes de formigas encontrou 69

genes semelhantes a PAM que podem cair em uma das cinco famílias semelhantes a PAM. As sequências de PAM previstas de *O. chelifer* foram agrupadas com base nos PAMs de referência, revelando agrupamentos de peptídeos semelhantes a crustina, himenoptaecina, defensina, abaecina e ICK (figura 7, 8, 9 e 10).



As defensinas de invertebrados têm seis resíduos de cisteína conservados, os quais formam três ligações dissulfeto que estabilizam um arranjo complexo de α -hélices e β -folhas, também conhecido como motivo CSa β . Embora a maioria deles mostre atividade antimicrobiana contra bactérias Gram-positivas, existem várias defensinas também direcionadas a bactérias Gram-negativas e/ou fungos (KOEHBACH, 2017). Foram agrupados doze sequências dos PAMs previstos de *O. chelifer* com a defensina-2 de *O. monticola*, incluindo nossa única sequência de PAM anotada que foi modelado (TRINITY_DN3796_c0_g1_i1.p1) (figura 6B), compartilhando 79,16% de identidade com o peptídeo de referência (figura 8). O motivo CSa β é claramente conservado neste PAM, semelhante aos da formicina e da heliomicina (KOEHBACH, 2017). O loop N-terminal é fundamental para a atividade das defensinas, e o motivo CSa β é sensível à sua modificação (LANDON et al., 2008). Em defensinas de formigas, esse loop apresenta comprimento variável, e essa diversidade pode facilitar a evolução de novos efeitos antimicrobianos para sobreviver em diferentes ambientes (ZHANG, ZHU, 2012). Curiosamente, o loop N-terminal da defensina modelada de *O. chelifer* parece ser mais longo do que os descritos anteriormente para outras espécies.

	1	10	20	30	40
Omo_Defensin2_A0A348G6A9MKLLAVFA	IFCVL	LAYVSADPIP	AVYDGP	TYELK.SIEEPESNDMSMGDAP....
TRINITY_DN13371_c0_g1_i1.p1AVSIGALLS	FLGFV	.VACICSCRRNENK	ITSANRSLP	DIKPKDSREHEDIAE...
TRINITY_DN28995_c0_g1_i1.p1	RNANISPPPLSL	LFLLI	..VPISTRRRHDC	SSYTTDGSV	RFRQRLAAFAENLSI...
TRINITY_DN13987_c0_g1_i1.p2	.MEHTGKFFVLL	FAIFV	..AVADENKEQDR	PKTFRRLIP	ADILRDFPGMCFAS...
TRINITY_DN25847_c0_g1_i1.p1	.MLSRDIIIIIL	SIFGL	C.FTASIADAVMC	NCEHCDEE	KVIGSEEEITTPAWRV
TRINITY_DN23731_c0_g1_i1.p1	.FTMDYRSART	FAFLIA	.STSCGDVPR	RSRAENY	DARKIASYINKFPFKY
TRINITY_DN1861_c3_g2_i1.p5MVFNLLLV	NILLIYLTHIF	GLR.....
TRINITY_DN10404_c0_g1_i1.p2MKRTYAVIV	LVIYICNIFLAI	SLGAKYCF	RFTWPGPSNKTVNCT
TRINITY_DN10663_c0_g1_i1.p1MLKMRWLA	MIMCLYAG...	LASARTEED	QQASNTDGIPI...
TRINITY_DN2760_c0_g2_i1.p2MRFYSVLS	VLLLLVALTCVA	EAGGSLR	ARRRSFPGDIGNTL
TRINITY_DN18722_c0_g1_i1.p2MNRNISLA	VVLVLL	VASA..CAQFY	KSGNCP	LQSTVSNCTPRCMS
TRINITY_DN16791_c0_g1_i1.p1MKAILSVA	LLMCLAMVYSY	TETNFEL	NCPIPDKENATILLP
TRINITY_DN3796_c0_g1_i1.p1MKLLAVFA	IFCVL	LAYATADPIP	AVYDGP	TDELN.PFEESASDEV
	50	60	70	80	90
Omo_Defensin2_A0A348G6A9	...IRQRRVTC	DLLSWQS..	KWFSFNHSA.	CAVKCLAQR	RRRGRCIDGVCI.....
TRINITY_DN13371_c0_g1_i1.p1SVPQDI	YETTE	TMGDHSELY	ATDQDTGQE	QIISERETQEV
TRINITY_DN28995_c0_g1_i1.p1	.ATISSCVI	ASERVG	KEQRGKKKA	QDRRC	SVGALAYHGR
TRINITY_DN13987_c0_g1_i1.p2	...KCATIEP	TKSWEL	SPFCGRST	CVFPAD	DNSGRFLFEL
TRINITY_DN25847_c0_g1_i1.p1	DDDGNR	TICAMDR	DFFNNR	TFPSICY	MSCYNHCTR
TRINITY_DN23731_c0_g1_i1.p1	...TKYTERI	AAGSYAN	EGQFFMA	VVH...RL	IDGRRVAQC
TRINITY_DN11861_c3_g2_i1.p5	...FYNSSV	QPNTK	NMWDENP.	KYSCELN	TDNVCIKYTY
TRINITY_DN10404_c0_g1_i1.p2NSPPVQC.	FYRQAE	CPNIGINS	NE.CPCK	KITLANVPE
TRINITY_DN10663_c0_g1_i1.p1DTMKTG	FESVFK	RPEDEP	ETVILQ	PFDEISLES
TRINITY_DN2760_c0_g2_i1.p2QKCCPN	KCGHR	SCAD..PS	AVSTGSG.	YKGSY...QD
TRINITY_DN18722_c0_g1_i1.p2PIYCD.SFFI	CQNGEPI	PMKCP	KGLDFNPE
TRINITY_DN16791_c0_g1_i1.p1	...IRQRRVTC	DLLSWNS..	KWFSFNHSA.	CAAKCLA	QRRRRGRCRGGICV
Omo_Defensin2_A0A348G6A9	..CR.....				
TRINITY_DN13371_c0_g1_i1.p1	..ARFAS.....				
TRINITY_DN28995_c0_g1_i1.p1				
TRINITY_DN13987_c0_g1_i1.p2	..CKLSDKTNK.....				
TRINITY_DN25847_c0_g1_i1.p1	..YKLRDGAC.....				
TRINITY_DN23731_c0_g1_i1.p1	..CVAKHPRRF.....				
TRINITY_DN1861_c3_g2_i1.p5				
TRINITY_DN10404_c0_g1_i1.p2	..IRFHSRNVYNL.....				
TRINITY_DN10663_c0_g1_i1.p1	RSCMGK	LQELFSKLS..			
TRINITY_DN2760_c0_g2_i1.p2	..CRLRQR.....				
TRINITY_DN18722_c0_g1_i1.p2	..CQFDRN	TRREKCTRA			
TRINITY_DN16791_c0_g1_i1.p1	..CKASG.....				
TRINITY_DN3796_c0_g1_i1.p1	..CR.....				

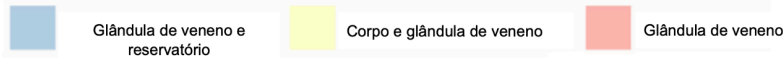
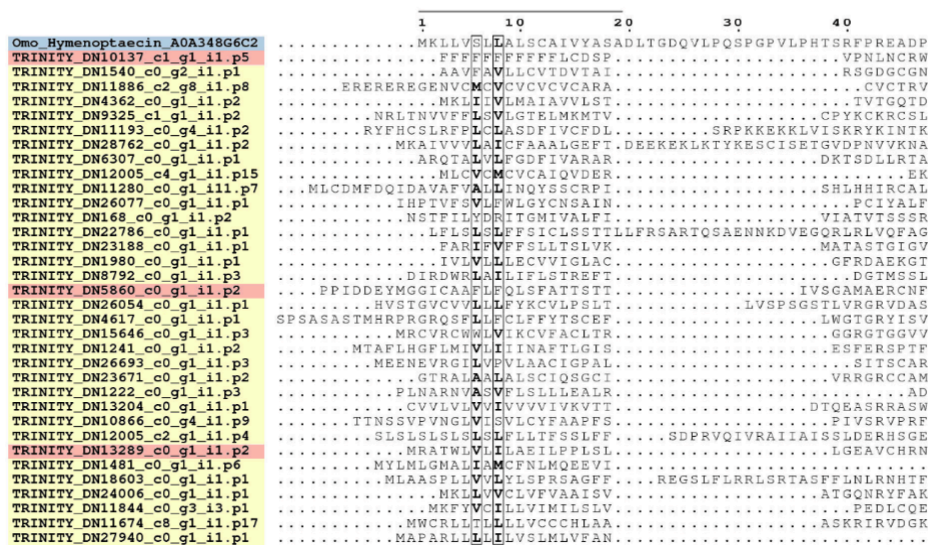


Figura 8. Alinhamento entre defensina e sequência de referência. Alinhamento global entre a defensina de referência e as sequências de PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do PAM de referência. Fonte: Menk et al., 2023.

PAMs ricos em glicina, como himenoptaécinas, são encontrados em um número diversificado de himenópteros (ZHANG, ZHU, 2012) e geralmente têm como alvo um amplo espectro de bactérias Gram-positivas e Gram-negativas (YI et al., 2014). Trinta e quatro sequências PAM preditas de *O. chelifera* agrupadas com a himenoptaécina de *O. monticola*. As himenoptaécinas, assim como outros peptídeos ricos em glicina, não adotam nenhuma estrutura tridimensional particular e são consideradas peptídeos lineares (KOEHBACH, CRAIK, 2019). Embora a himenoptaécina modelada de *O. chelifera* seja principalmente linear (Figura 6C), surpreendentemente há duas α -hélices visíveis em sua estrutura e ela mantém 18,81% de identidade com o peptídeo de referência (figura 9). Isso pode significar que essa sequência não propriamente uma himenoptaécina, mas ainda pode ser um PAM relacionado a esse grupo.



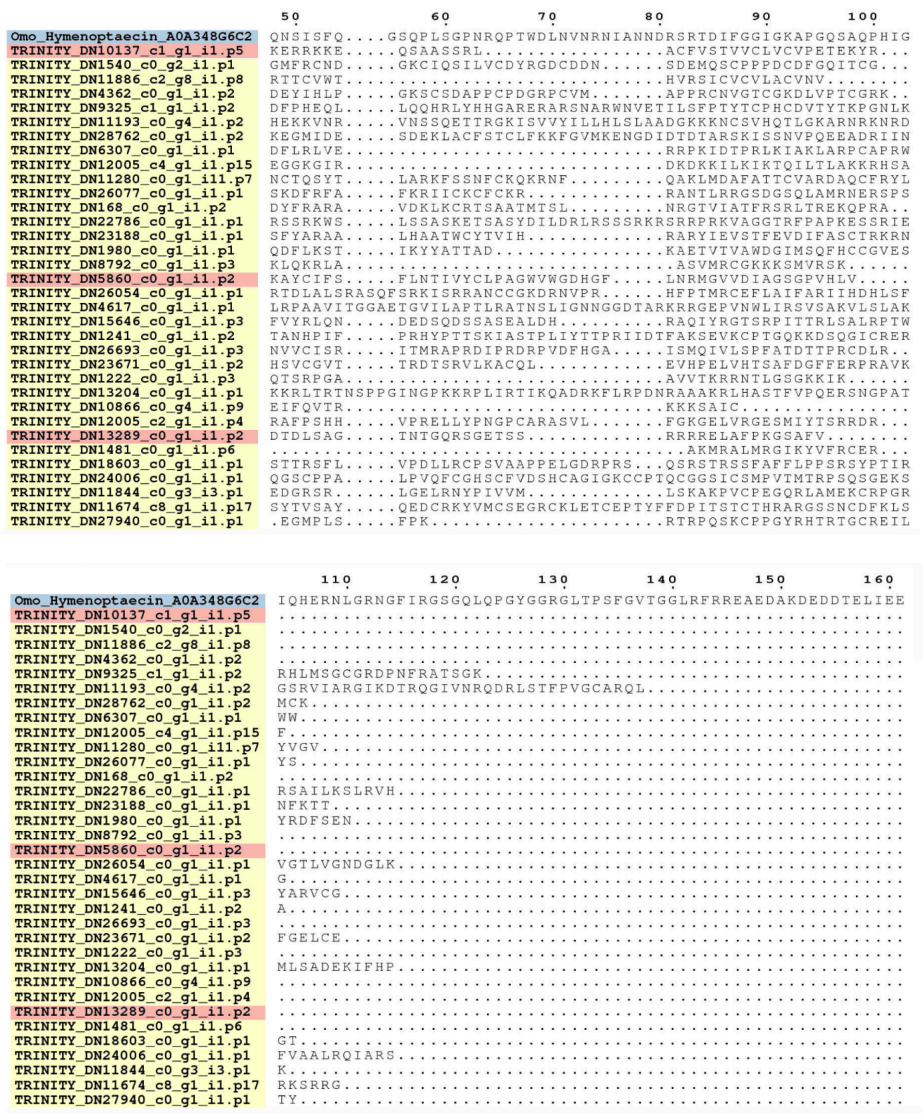


Figura 9. Alinhamento entre himenoptaecina e sequência de referência. Alinhamento global entre a himenoptaecina de referência e as sequências PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do PAM de referência. Fonte: Menk et al., 2023.

Os PAMs do tipo ICK contêm a dobra ICK, que consiste em uma folha β de fita tripla antiparalela unida por três ligações de dissulfeto. Este motivo estrutural é encontrado em toxinas da peçonha e em PAMs de muitos organismos, como fungos, plantas, moluscos e artrópodes (CRAIK, DALY, WAINE, 2001). Sete sequências dos PAMs previstos de *O. chelifer* agrupados com o PAM semelhante ao

ICK da poneratoxina de *Dinoponera quadriceps*. Destes, apenas dois tinham pelo menos seis resíduos de cisteína e tiveram sua estrutura secundária prevista pelo PSIPRED. Apenas uma sequência apresentou pelo menos três fitas e foi escolhida para ter sua estrutura tridimensional modelada (figura 6D), apresentando 16% de identidade em relação ao PAM de referência (figura 10).

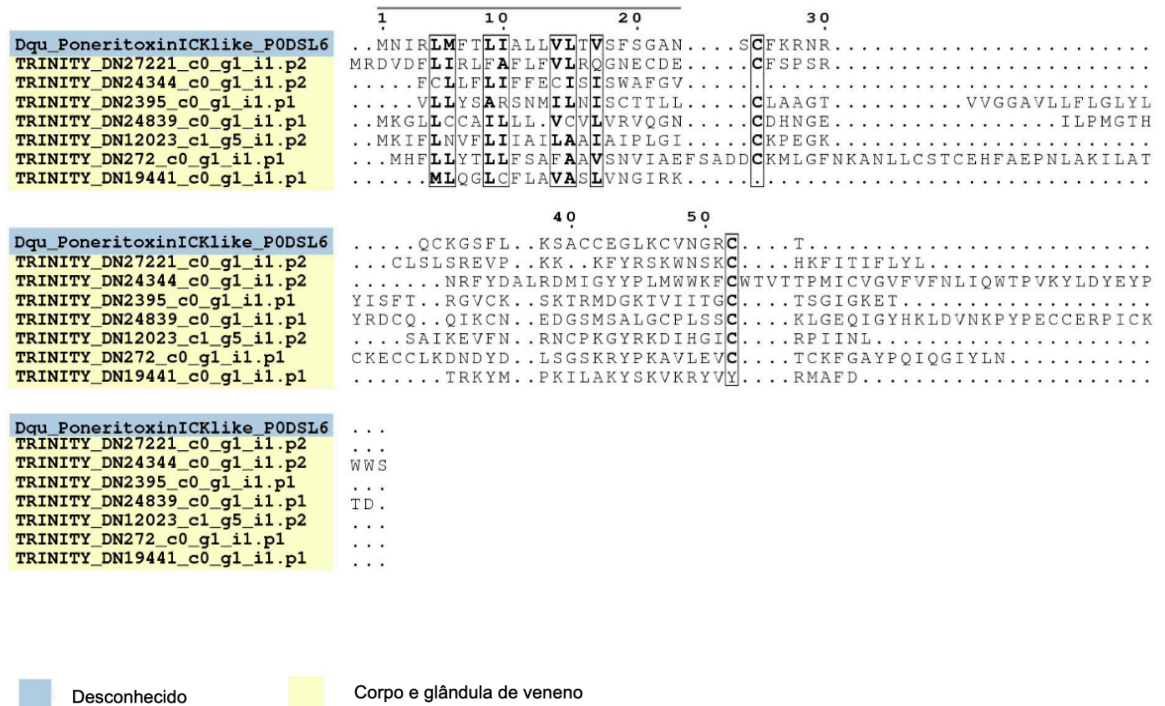


Figura 10. Alinhamento entre PAM do tipo ICK e sequência de referência. Alinhamento global entre o PAM do tipo ICK de referência e as sequências PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do PAM de referência. Fonte: Menk et al., 2023.

Este peptídeo supostamente possui as características necessárias para uma dobra ICK, sendo pelo menos seis resíduos de cisteína e uma folha β de fita tripla antiparalela. No entanto, o número de resíduos de cisteína na região da folha β de fita tripla foi inferior ao esperado (figura 11). Este achado sugere que este peptídeo se assemelha a peptídeos do tipo ICK, talvez mantendo uma proximidade funcional a este grupo.

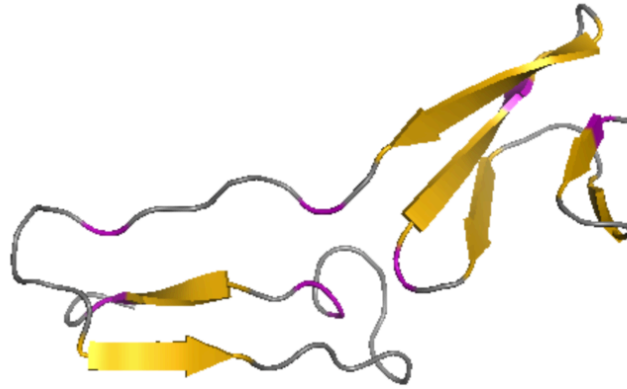


Figura 11. Previsão da estrutura tridimensional do homólogo da toxina tipo ICK da poneratoxina putativa de *O. chelififer* (TRINITY_DN24839_c0_g1_i1.p1) Amarelo: conformação β e roxo: α -hélices. Fonte: Menk et al., 2023.

PAMs ricos em prolina, como a abaecina, são peptídeos catiônicos lineares que possuem uma alta porcentagem (>25%) de resíduos de prolina (GRAF et al., 2014; BULET et al., 1999). Sua capacidade de translocar a membrana bacteriana e matar o microrganismo visando componentes intracelulares está correlacionada com suas estruturas α -helicoidais anfipáticas e o motivo PXP/PXXP característico (CARDOSO et al., 2019; YANG, SHIN, SHIN, 2021). Esses AMPs não líticos mostram um amplo espectro de atividade bacteriana (CARDOSO et al., 2019). Quinze seqüências dos AMPs previstos de *O. chelififer* agrupadas com a abaecina de *O. monticola*. Destes, apenas dois peptídeos tinham > 20% de conteúdo de prolina, e seu alinhamento com a abaecina de *O. monticola* revelou que ambos possuem os motivos PXP/PXXP (figura 12), indicando uma possível proximidade funcional. A seqüência com maior teor de prolina (29,2%) compartilhou uma identidade de 18,75% com o PAM de referência (figura 13) e sua previsão de estrutura tridimensional mostrou um peptídeo linear (figura 6E) exatamente como esperado para um peptídeo semelhante à abaecina. Peptídeos estruturados nem sempre são os mais ativos, e as alças de uma molécula são importantes para sua atividade e citotoxicidade (LANDON et al., 2008).

	1	10	20	30
1	FRV	PPRSP	PGGWKGNFPGQGFNP	PK.....FGK.
2	IPT	PAPVA	PGAITAYPIPQRPAYAP	PNPWVYPKPTTVQPIGR.
3	ISQ	<u>PLPPT</u>	P...SGHVINKRVAPIR	...PVPTVIGLLMKQ

Figura 12. Alinhamento de sequência primária entre duas sequências PAM previstas de *O. chelififer* e a abaecina A0A348G6C5 de *O. monticola*. Resíduos semelhantes estão em caixas. Resíduos idênticos estão em negrito. A linha vermelha indica os motivos PXP/PXXP. Fonte: Menk et al., 2023.

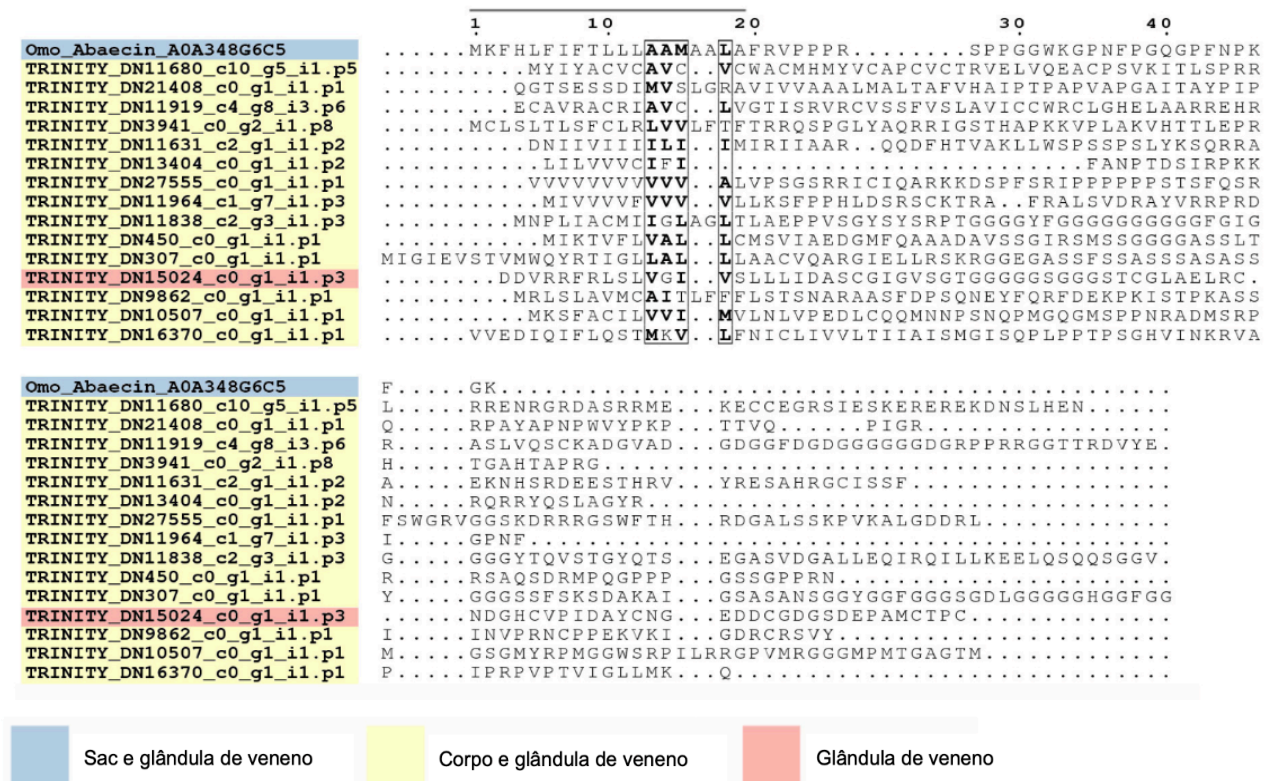


Figura 13. Alinhamento entre abaecina e sequência de referência. Alinhamento global entre a abaecina de referência e as sequências PAM candidatas agrupadas. A linha cinza indica o peptídeo sinal do peptídeo de referência. Fonte: Menk et al., 2023.

5.5 PAMs podem compor o sistema imunológico de *O. chelififer*

A maioria das formigas sociais, incluindo *O. chelififer*, vive em colônias densamente povoadas, onde o alimento é acumulado em um ambiente quente e úmido, facilitando a proliferação de microrganismos patogênicos (QUINET et al., 2012). A primeira linha de defesa do hospedeiro contra esses patógenos é o sistema imune inato, no qual os PAMs são elementos essenciais para eliminar infecções (ZHANG, ZHU, 2012). Para verificar se os PAMs candidatos também podem funcionar no sistema imune inato, e não apenas como toxinas da peçonha, mapeamos os transcritos da glândula da peçonha de *O. chelififer* contra os transcritos montados no corpo da formiga. Observamos que 129 das 136

sequências de PAMs candidatos foram identificados com transcritos do corpo da formiga (Apêndice B), com identidade média de $98,9 \pm 3,5\%$ e valor E máximo de $1,7 \times 10^{-5}$. Isso sugere que esses PAMs candidatos também podem desempenhar um papel na imunidade animal, auxiliando na manutenção de uma colônia saudável.

Também procuramos evidências de tradução de PAMs previstos pesquisando as sequências peptídicas contra um banco de dados de peptídeos *O. monticola* detectados por LC-ESI-MS (TANI et al., 2019). Foi possível recuperar um peptídeo compatível com o banco de dados (TRINITY_DN18722_c0_g1_i1.p2), sugerindo que pelo menos um de nossos PAMs candidatos está traduzido ativamente. Além disso, essa sequência foi a única a ser detectada em todo o secretoma previsto. Isso é interessante devido ao fato de que a maioria dos proteomas não está devidamente preparada para detectar peptídeos. A resolução do equipamento na maioria das vezes não atinge a zona necessária para visualizar proteínas tão pequenas (KIM, ZHONG, PANDEY, 2016).

5.6 Conservação da transcrição de PAMs candidatos no gênero *Odontomachus*

Durante o mapeamento das leituras de aminoácidos de *O. chelififer* com PALADIN (WESTBROOK et al., 2017), usando apenas as sequências maduras dos peptídeos secretomas previstos como referência (APÊNDICE C), verificamos que 132 dos 136 PAMs previstos (97%) estavam presentes na espécie original. Isso sugere que 3% dos PAMs candidatos são gerados apenas durante os procedimentos de montagem ou podem representar artefatos em potencial. Além disso, podemos inferir que a abundância de PAMs candidatos é maior do que o secretoma residual ($P_{\text{Mann-Whitney's U}} = 3,4 \times 10^{-5}$), indicando que o nível basal de expressão desses peptídeos pode ser maior do que o anterior.

Por fim, a verificação desses PAMs candidatos em diferentes espécies do mesmo gênero pode ser uma indicação de que esses PAMs são conservados e, portanto, devem ter alguma função. Encontramos um total de 85 de 136 PAMs previstos em *O. monticola* por mapeamento reverso dos peptídeos contra suas leituras de aminoácidos. Apenas 23 dos 85 PAMs candidatos encontrados nas duas espécies apresentaram maior abundância relativa em *O. monticola* em comparação com *O. chelififer*. Esses achados indicam a possibilidade de que diferentes peptídeos possam realizar tarefas especificamente relacionadas ao nicho e habitat das formigas, sugerindo um possível papel ecológico para essas entidades.

A composição da peçonha de formigas é complexa e suas funções ainda não foram totalmente caracterizadas e diversas toxinas ainda precisam ser identificadas e investigadas, pois as suas moléculas são extremamente promissoras para o desenvolvimento de medicamentos.

6. Conclusão

Verificamos que o número de peptídeos previstos para serem secretados na glândula da peçonha é bastante baixo (1.022 de 125.123). Além disso, prevemos que cerca de 13,3% desses peptídeos podem ser antimicrobianos sob um rigoroso controle de qualidade, e que isso não cria um grupo de proteínas desordenadas, mas proteínas globulares que tendem a ser mais hemolíticas do que o peptídeo médio no secretoma. Isso sugere que esses peptídeos podem estar desempenhando o papel de toxinas na peçonha de *O. chelifera*. Também encontramos evidências de que os PAMs previstos não são apenas toxinas, pois muitos estão presentes no transcriptoma do corpo de *O. chelifera*, que contém a maioria dos PAMs candidatos. A evidência de que esses peptídeos também são traduzidos e transcritos em outras espécies do mesmo gênero fortalece nossas afirmações.

7. Perspectivas futuras

Peptídeos preditos na peçonha de formigas demonstram que há um caminho promissor aplicável na área farmacológica, podendo contribuir para o desenvolvimento de novos antibióticos ou compostos contra patógenos.

Os medicamentos modelados a partir de peptídeos de venenos despertam interesses na indústria farmacêutica e nos pesquisadores, são promissores para o desenvolvimento de novas moléculas.

Como continuidade do trabalho encontra-se a síntese dos peptídeos ou sua produção recombinante, para a realização de testes para verificar suas atividades antimicrobianas.

8. Referências

ABD EL-WAHED, Aida et al. Wasp venom biochemical components and their potential in biological applications and nanotechnological interventions. **Toxins**, v. 13, n. 3, p. 206, 2021.

AGUIAR, Marisa Barbosa de. **Estudo estrutural por RMN do peptídeo policatiônico polybine I de veneno da vespa social Polybia paulista**. 2006.

AGARWAL, Surbhi; SHARMA, Garima; VERMA, Kavita; et al. Pharmacological potential of ants and their symbionts – a review. **Entomologia Experimentalis et Applicata**, v. 170, n. 12, p. 1032–1048, 2022.

AILI, Samira R.; TOUCHARD, Axel; ESCOUBAS, Pierre; et al. Diversity of peptide toxins from stinging ant venoms. **Toxicon**, v. 92, p. 166–178, 2014.

ALTSCHUL, Stephen F.; GISH, Warren; MILLER, Webb; et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403–410, 1990.

AMORIM-CARMO, Bruno; PARENTE, Adriana M. S.; SOUZA, Eden S.; et al. Antimicrobial Peptide Analogs From Scorpions: Modifications and Structure-Activity. **Frontiers in Molecular Biosciences**, v. 9, p. 887763, 2022.

AZEVEDO-SILVA, Marianne; LEMOS, Alessandra S M; GONÇALVES-NETO, Salatiel; et al. Are There Edge Effects on the Genetic Diversity of the Trap-Jaw Ant *Odontomachus chelifer* (Formicidae: Ponerinae) in a Neotropical Savanna Fragment? A First Assessment. **Environmental Entomology**, v. 52, n. 2, p. 279–285, 2023.

BAPTISTA-SAIDEMBERG, Nicoli Barão. Efeito das interações peptídeo-peptídeo e peptídeo-membrana nas atividades funcionais de toxinas peptídicas do veneno da vespa social *Agelaia pallipes pallipes* (Hymenoptera, Vespidae). 2007.

BAPTISTA-SAIDEMBERG, Nicoli Barão. Análise peptidômica dos venenos de vespas sociais neotropicais. 2011.

BARROS, L. C. et al. A new fibrin sealant from *Crotalus durissus terrificus* venom: applications in medicine. **Journal of Toxicology and Environmental Health, Part B**, v. 12, n. 8, p. 553-571, 2009.

BHANDARI, Bikash K.; GARDNER, Paul P.; LIM, Chun Shen. **Razor: annotation of signal peptides from toxins**. [s.l.]: Bioinformatics, 2020. Disponível em: <<http://biorxiv.org/lookup/doi/10.1101/2020.11.30.405613>>. Acesso em: 26 fev. 2023.

BIEBER, Ana Gabriela Delgado; PEREIRA, Eduardo Roberto; BOTTCHEER, Claudia. Flower perianth: first observations on a novel diet item for the ponerine ant *Odontomachus chelifer* (Latreille). **Revista Brasileira de Entomologia**, v. 64, n. 3, p. e20200019, 2020.

BIN, T. E. N. G.; JUNFENG, L. I.; PEIGEN, R. E. N. Peptide drugs application in metabolic diseases and discovery strategies. **Journal of Holistic Integrative Pharmacy**, v. 3, n. 1, p. 24-31, 2022.

BOLTON, B. A taxonomic and zoogeographical census of the extant ant taxa (Hymenoptera: Formicidae). **Journal of Natural History**, v. 29, n. 4, p. 1037–1056, 1995.

BOUZID, Wafa et al. Profiling the venom gland transcriptome of *Tetramorium bicarinatum* (Hymenoptera: Formicidae): The first transcriptome analysis of an ant species. **Toxicon**, v. 70, p. 70-81, 2013.

BUCHAN, Daniel W A; JONES, David T. The PSIPRED Protein Analysis Workbench: 20 years on. **Nucleic Acids Research**, v. 47, n. W1, p. W402–W407, 2019.

BULET, Phillipe; HETRU, Charles; DIMARCQ, Jean-Luc; *et al.* Antimicrobial peptides in insects; structure and function. **Developmental & Comparative Immunology**, v. 23, n. 4–5, p. 329–344, 1999.

CAMACHO, Christiam; COULOURIS, George; AVAGYAN, Vahram; *et al.* BLAST+: architecture and applications. **BMC Bioinformatics**, v. 10, n. 1, p. 421, 2009.

CANTALAPIEDRA, Carlos P; HERNÁNDEZ-PLAZA, Ana; LETUNIC, Ivica; *et al.* eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain

Prediction at the Metagenomic Scale. **Molecular Biology and Evolution**, v. 38, n. 12, p. 5825–5829, 2021.

CARDOSO, Marlon H.; MENEGUETTI, Beatriz T.; COSTA, Bruna O.; *et al.* Non-Lytic Antibacterial Peptides That Translocate Through Bacterial Membranes to Act on Intracellular Targets. **International Journal of Molecular Sciences**, v. 20, n. 19, p. 4877, 2019.

CARVALHO, CAMILA LIMA DE SOUZA. DESENVOLVIMENTO DE ANÁLISE DE TRANSCRIPTOMAS PARA IDENTIFICAÇÃO DE GENES ALVO. 2021.

CHEN, Li; FADAMIRO, Henry Y. Re-investigation of venom chemistry of *Solenopsis* fire ants. I. Identification of novel alkaloids in *S. richteri*. **Toxicon**, v. 53, n. 5, p. 469–478, 2009.

CHEN, Xiaoling *et al.* Evaluation of the bioactivity of a mastoparan peptide from wasp venom and of its analogues designed through targeted engineering. **International journal of biological sciences**, v. 14, n. 6, p. 599, 2018.

COLOGNA, Camila Takeno *et al.* Peptidomic comparison and characterization of the major components of the venom of the giant ant *Dinoponera quadriceps* collected in four different areas of Brazil. **Journal of proteomics**, v. 94, p. 413-422, 2013.

CONESA, Ana; MADRIGAL, Pedro; TARAZONA, Sonia; *et al.* A survey of best practices for RNA-seq data analysis. **Genome Biology**, v. 17, n. 1, p. 13, 2016.

CORREA-VIANA, Geovana Roberto. **Reatividade cruzada: alérgenos do veneno de *P. paulista* e de outros insetos Hymenoptera**. 2018.

CRAIK, David J.; DALY, Norelle L.; WAINE, Clement. The cystine knot motif in toxins and implications for drug design. **Toxicon**, v. 39, n. 1, p. 43–60, 2001.

DAMORIN, Tulio Morato *et al.* **Caracterização bioquímica parcial dos componentes presentes na peçonha bruta da formiga *Odontomachus chelifer***. 2019.

Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**, v. 44, n. D1, p. D7–D19, 2016.

DE AZEVEDO, Ricardo A. et al. Mastoparan induces apoptosis in B16F10-Nex2 melanoma cells via the intrinsic mitochondrial pathway and displays antitumor activity in vivo. **Peptides**, v. 68, p. 113-119, 2015.

DE MELLO SPOSITO, Maria Matilde. Toxina botulínica tipo A: propriedades farmacológicas e uso clínico. **Acta Fisiátrica**, v. 11, n. Supl. 1, p. S7-S44, 2004.

DHAYGUDE, Kishor et al. Transcriptome sequencing reveals high isoform diversity in the ant *Formica exsecta*. **PeerJ**, v. 5, p. e3998, 2017.

DODOU LIMA, Hilania Valéria; DE PAULA CAVALCANTE, Carolina Sidrim; RÁDIS-BAPTISTA, Gandhi. Antifungal In Vitro Activity of Pilosulin- and Ponericin-Like Peptides from the Giant Ant *Dinoponera quadriceps* and Synergistic Effects with Antimycotic Drugs. **Antibiotics**, v. 9, n. 6, p. 354, 2020.

DONOVAN, Gregory; BALDO, Brian. Method of cell inhibition using polypeptides derived from the venom of the Australian jumper ant *Myrmecia pilosula*. Disponível em:
<<https://patentimages.storage.googleapis.com/23/70/48/00c167a94b7b0e/US6294649.pdf>>.

DONOVAN, Gregory R.; BALDO, Brian A.; SUTHERLAND, Struan. Molecular cloning and characterization of a major allergen (Myr p I) from the venom of the Australian jumper ant, *Myrmecia pilosula*. **Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression**, v. 1171, n. 3, p. 272–280, 1993.

DOS SANTOS, Ana P.; DE ARAÚJO, Tamara G.; RÁDIS-BAPTISTA, Gandhi. Nanoparticles functionalized with venom-derived peptides and toxins for pharmaceutical applications. **Current Pharmaceutical Biotechnology**, v. 21, n. 2, p. 97-109, 2020.

EBERHARDT, R. Y.; HAFT, D. H.; PUNTA, M.; *et al.* AntiFam: a tool to help identify spurious ORFs in protein annotation. **Database**, v. 2012, n. 0, p. bas003–bas003, 2012.

ERDŐS, Gábor; PAJKOS, Mátyás; DOSZTÁNYI, Zsuzsanna. IUPred3: prediction of protein disorder enhanced with unambiguous experimental annotation and visualization of evolutionary conservation. **Nucleic Acids Research**, v. 49, n. W1, p. W297–W303, 2021.

ESCOUBAS, Pierre; KING, Glenn F. Venomics as a drug discovery platform. **Expert review of proteomics**, v. 6, n. 3, p. 221-224, 2009.

ERREIRA, S. H. A bradykinin-potentiating factor (BPF) present in the venom of *Bothrops jararaca*. **British journal of pharmacology and chemotherapy**, v. 24, n. 1, p. 163, 1965.

FERNANDES, Itanna Oliveira. **Análise filogenética de *Anochetus* Mayr, 1961 e *Odontomachus Latreille, 1804* (Hymenoptera: Formicidae: Ponerinae) e revisão taxonômica de *Anochetus* para a região Neotropical**. 2017.

FINGERHUT, Legana C H W; MILLER, David J; STRUGNELL, Jan M; *et al.* ampir: an R package for fast genome-wide prediction of antimicrobial peptides. **Bioinformatics**, v. 36, n. 21, p. 5262–5263, 2021.

FONTANA, Renato *et al.* Complexidade e atividade biológica das peçonhas de formigas, em particular de poneromorfos. **As formigas Poneromorfos do Brasil**, p. 271-284, 2015.

FU, Limin; NIU, Beifang; ZHU, Zhengwei; *et al.* CD-HIT: accelerated for clustering the next-generation sequencing data. **Bioinformatics**, v. 28, n. 23, p. 3150–3152, 2012.

GOLDENBERG, Ronald M.; STEEN, Oren. Semaglutide: review and place in therapy for adults with type 2 diabetes. **Canadian journal of diabetes**, v. 43, n. 2, p. 136-145, 2019.

GOTZ, S.; GARCIA-GOMEZ, J. M.; TEROL, J.; *et al.* High-throughput functional annotation and data mining with the Blast2GO suite. **Nucleic Acids Research**, v. 36, n. 10, p. 3420–3435, 2008.

GODOY, Isabela Alves de. Estudos transcriptômicos em Anthozoa e a anotação do transcriptoma de *Ceriantheomorpha brasiliensis* (Cnidaria, Ceriantharia). 2021.

GRABHERR, Manfred G; HAAS, Brian J; YASSOUR, Moran; *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature Biotechnology**, v. 29, n. 7, p. 644–652, 2011.

GRAF, Michael; MARDIROSSIAN, Mario; NGUYEN, Fabian; *et al.* Proline-rich antimicrobial peptides targeting protein synthesis. **Natural Product Reports**, v. 34, n. 7, p. 702–711, 2017.

GU, Hyemin; HAN, Sang Mi; PARK, Kwan-Kyu. Therapeutic effects of apamin as a bee venom component for non-neoplastic disease. *Toxins*, v. 12, n. 3, p. 195, 2020.

GUIDO-PATIÑO, Juan Carlos; PLISSON, Fabien. Profiling hymenopteran venom toxins: Protein families, structural landscape, biological activities, and pharmacological benefits. **Toxicon: X**, v. 14, p. 100119, 2022.

GUIMARÃES, Denise Oliveira; FERRO, Milene; SANTOS, Thamires Silva; *et al.* Transcriptomic and biochemical analysis from the venom gland of the neotropical ant *Odontomachus chelifer*. **Toxicon**, v. 223, p. 107006, 2023.

GUIMARÃES, Maria. Jardineiras fiéis: formigas ajudam sementes a germinar na mata atlântica e no cerrado. *Revista FAPESP*, n. 8, p.100-103, 2012. Disponível em: http://revistapesquisa.fapesp.br/wp-content/uploads/2012/08/100-103_jardineiras.pdf

GUZMAN, Jesus; TÉNÉ, Nathan; TOUCHARD, Axel; *et al.* Anti-Helicobacter pylori Properties of the Ant-Venom Peptide Bicarinalin. **Toxins**, v. 10, n. 1, p. 21, 2017.

HAAS, Brian J; PAPANICOLAOU, Alexie; YASSOUR, Moran; *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. **Nature Protocols**, v. 8, n. 8, p. 1494–1512, 2013.

HAIGHT, Kevin L. Patterns of venom production and temporal polyethism in workers of Jerdon's jumping ant, *Harpegnathos saltator*. **Journal of Insect Physiology**, v. 58, n. 12, p. 1568–1574, 2012.

HABENSTEIN, Jens et al. Transcriptomic, peptidomic, and mass spectrometry imaging analysis of the brain in the ant *Cataglyphis nodus*. **Journal of Neurochemistry**, v. 158, n. 2, p. 391-412, 2021.

HANSEN, Ida K. Ø.; LÖVDAHL, Tomas; SIMONOVIC, Danijela; *et al.* Antimicrobial Activity of Small Synthetic Peptides Based on the Marine Peptide Turgencin A: Prediction of Antimicrobial Peptide Sequences in a Natural Peptide and Strategy for Optimization of Potency. **International Journal of Molecular Sciences**, v. 21, n. 15, p. 5460, 2020.

HANSEN, Kasper D.; BRENNER, Steven E.; DUDOIT, Sandrine. Biases in Illumina transcriptome sequencing caused by random hexamer priming. **Nucleic Acids Research**, v. 38, n. 12, p. e131–e131, 2010.

HERMANN, Henry R.; BLUM, Murray S. The Morphology and Histology of the Hymenopterous Poison Apparatus. I. *Paraponera clavata* (Formicidae). **Annals of the Entomological Society of America**, v. 59, n. 2, p. 397–409, 1966.

HIRAI, YUKO et al. A new mast cell degranulating peptide, polistes mastoparan, in the venom of *Polistes jadwigae*. **Biomedical Research**, v. 1, n. 2, p. 185-187, 1980.

HOFFMAN, Donald R. Ant venoms. **Current Opinion in Allergy & Clinical Immunology**, v. 10, n. 4, p. 342–346, 2010.

HOLDING, Matthew; MARGRES, Mark; MASON, Andrew; *et al.* Evaluating the Performance of De Novo Assembly Methods for Venom-Gland Transcriptomics. **Toxins**, v. 10, n. 6, p. 249, 2018.

HUBER, John T. Biodiversity of hymenoptera. **Insect biodiversity: science and society**, p. 419-461, 2017.

HUERTA-CEPAS, Jaime; SZKLARCZYK, Damian; HELLER, Davide; *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. **Nucleic Acids Research**, v. 47, n. D1, p. D309–D314, 2019.

HUNTER, John D. Matplotlib: A 2D Graphics Environment. **Computing in Science & Engineering**, v. 9, n. 3, p. 90–95, 2007.

HURKA, Sabine; BRINKROLF, Karina; ÖZBEK, Rabia; *et al.* Venomics of the Central European Myrmicine Ants *Myrmica rubra* and *Myrmica ruginodis*. **Toxins**, v. 14, n. 5, p. 358, 2022.

IBARRA-VEGA, Rodrigo; GALVÁN-HERNÁNDEZ, Alan Roberto; SALAZAR-MONGE, Hermenegildo; *et al.* Antimicrobial Compounds from Skin Secretions of Species That Belong to the Bufonidae Family. **Toxins**, v. 15, n. 2, p. 145, 2023.

INAGAKI, Hidetoshi; AKAGI, Masaaki; IMAI, Hirotami T.; *et al.* Molecular cloning and biological characterization of novel antimicrobial peptides, pilosulin 3 and pilosulin 4, from a species of the Australian ant genus *Myrmecia*. **Archives of Biochemistry and Biophysics**, v. 428, n. 2, p. 170–178, 2004.

INAGAKI, Hidetoshi; AKAGI, Masaaki; IMAI, Hirotami T.; *et al.* Pilosulin 5, a novel histamine-releasing peptide of the Australian ant, *Myrmecia pilosula* (Jack Jumper Ant). **Archives of Biochemistry and Biophysics**, v. 477, n. 2, p. 411–416, 2008.

JENSSEN, Havard; HAMILL, Pamela; HANCOCK, Robert EW. Peptide antimicrobial agents. **Clinical microbiology reviews**, v. 19, n. 3, p. 491-511, 2006.

KAUSAR, Saima; ABBAS, Muhammad Nadeem; CUI, Hongjuan. A review on the DNA methyltransferase family of insects: Aspect and prospects. **International journal of biological macromolecules**, v. 186, p. 289-302, 2021.

KATOH, K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. **Nucleic Acids Research**, v. 30, n. 14, p. 3059–3066, 2002.

KAZUMA, Kohei; MASUKO, Keiichi; KONNO, Katsuhiro; *et al.* Combined Venom Gland Transcriptomic and Venom Peptidomic Analysis of the Predatory Ant *Odontomachus monticola*. **Toxins**, v. 9, n. 10, p. 323, 2017.

KIM, Min-Sik; ZHONG, Jun; PANDEY, Akhilesh. Common errors in mass spectrometry-based analysis of post-translational modifications. **PROTEOMICS**, v. 16, n. 5, p. 700–714, 2016.

KING, Malcolm A.; WU, Qi-Xuan; DONOVAN, Gregory R.; *et al.* Flow cytometric analysis of cell killing by the jumper ant venom peptide pilosulin 1. **Cytometry**, v. 32, n. 4, p. 268–273, 1998.

KOEHBACH, Johannes. Structure-Activity Relationships of Insect Defensins. **Frontiers in Chemistry**, v. 5, p. 45, 2017.

KOEHBACH, Johannes; CRAIK, David J. The Vast Structural Diversity of Antimicrobial Peptides. **Trends in Pharmacological Sciences**, v. 40, n. 7, p. 517–528, 2019.

KUNITZ, André Guilherme *et al.* Melitina proveniente do veneno de abelha: processo de purificação, aplicação e avaliação econômica. 2015.

LANDON, Céline; BARBAULT, Florent; LEGRAIN, Michèle; *et al.* Rational design of peptides active against the gram positive bacteria *Staphylococcus aureus*: Rational Design of Peptides. **Proteins: Structure, Function, and Bioinformatics**, v. 72, n. 1, p. 229–239, 2008.

LANGMEAD, Ben; SALZBERG, Steven L. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, v. 9, n. 4, p. 357–359, 2012.

LARABEE, Fredrick J.; SUAREZ, Andrew V. Mandible-Powered Escape Jumps in Trap-Jaw Ants Increase Survival Rates during Predator-Prey Encounters. **PLOS ONE**, v. 10, n. 5, p. e0124871, 2015.

LATREILLE, P. A. **Histoire naturelle des fourmis, et recueil de mémoires et d'observations sur les abeilles, les araignées, les faucheurs, et autres insectes.** Paris: De l'impr. de Crapelet, 1802. Disponível em: <<https://www.biodiversitylibrary.org/bibliography/65810>>. Acesso em: 13 jun. 2023.

LAWRENCE, Travis J; CARPER, Dana L; SPANGLER, Margaret K; *et al.* amPEPpy 1.0: a portable and accurate antimicrobial peptide prediction tool. **Bioinformatics**, v. 37, n. 14, p. 2058–2060, 2021.

LI, Bo; DEWEY, Colin N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. **BMC Bioinformatics**, v. 12, n. 1, p. 323, 2011.

LI, Chenkai; SUTHERLAND, Darcy; HAMMOND, S. Austin; *et al.* AMPLify: attentive deep learning model for discovery of novel antimicrobial peptides effective against WHO priority pathogens. **BMC Genomics**, v. 23, n. 1, p. 77, 2022.

LI, Shuqin; WANG, Yajie; XUE, Zihan; *et al.* The structure-mechanism relationship and mode of actions of antimicrobial peptides: A review. **Trends in Food Science & Technology**, v. 109, p. 103–115, 2021.

LI, Yanmei; XIANG, Qi; ZHANG, Qihao; *et al.* Overview on the recent study of antimicrobial peptides: Origins, functions, relative mechanisms and application. **Peptides**, v. 37, n. 2, p. 207–215, 2012.

LIMA, William Gustavo; DE LIMA, Maria Elena. Therapeutic Prospection of Animal Venoms-Derived Antimicrobial Peptides against Infections by Multidrug-Resistant *Acinetobacter baumannii*: A Systematic Review of Pre-Clinical Studies. **Toxins**, v. 15, n. 4, p. 268, 2023.

LIN, Tzu-Tang; YANG, Li-Yen; LU, I-Hsuan; *et al.* AI4AMP: an Antimicrobial Peptide Predictor Using Physicochemical Property-Based Encoding Method and Deep Learning. **mSystems**, v. 6, n. 6, p. e00299-21, 2021.

LOPES, Kamila Soares; RIOS, Emiliano Ricardo Vasconcelos; LIMA, Camila Nayane De Carvalho; *et al.* The effects of the Brazilian antDinoponera quadriceps venom on chemically induced seizure models. **Neurochemistry International**, v. 63, n. 3, p. 141–145, 2013.

MAGALHÃES, Marta Regina. Caracterização de enzimas em peçonhas animais: identificação de fosfolipases do escorpião *Hadrurus gerstchi* e atividades enzimáticas da arraia *Potamotrygon falkneri*. 2017.

MATUSZEK, Maria A.; HODGSON, Wayne C.; SUTHERLAND, Struan K.; *et al.* Pharmacological studies of jumper ant (*Myrmecia pilosula*) venom: Evidence for the presence of histamine, and haemolytic and eicosanoid-releasing factors. **Toxicon**, v. 30, n. 9, p. 1081–1091, 1992.

MENK, Josilene J. *et al.* Antimicrobial Peptide Arsenal Predicted from the Venom Gland Transcriptome of the Tropical Trap-Jaw Ant *Odontomachus chelifer*. **Toxins**, v. 15, n. 5, p. 345, 2023.

MILJANICH, G. P. Ziconotide: neuronal calcium channel blocker for treating severe chronic pain. **Current medicinal chemistry**, v. 11, n. 23, p. 3029-3040, 2004.

MIRDITA, Milot; SCHÜTZE, Konstantin; MORIWAKI, Yoshitaka; *et al.* ColabFold: making protein folding accessible to all. **Nature Methods**, v. 19, n. 6, p. 679–682, 2022.

MISTRY, Jaina; FINN, Robert D.; EDDY, Sean R.; *et al.* Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. **Nucleic Acids Research**, v. 41, n. 12, p. e121–e121, 2013.

MONTEIRO, Marília Lopes; LIMA, Dânya Bandeira; FREIRE, Katielle Albuquerque; *et al.* Rational design of a trypanocidal peptide derived from *Dinoponera quadriceps* venom. **European Journal of Medicinal Chemistry**, v. 241, p. 114624, 2022.

NECCI, Marco; PIOVESAN, Damiano; TOSATTO, Silvio C. E. Large-scale analysis of intrinsic disorder flavors and associated functions in the protein sequence universe: Large-Scale Analysis of Disorder in Protein Sequences. **Protein Science**, v. 25, n. 12, p. 2164–2174, 2016.

ORIVEL, Jérôme; REDEKER, Virginie; LE CAER, Jean-Pierre; *et al.* Ponericins, New Antibacterial and Insecticidal Peptides from the Venom of the Ant *Pachycondyla goeldii*. **Journal of Biological Chemistry**, v. 276, n. 21, p. 17823–17829, 2001.

PALMA, Mario Sergio. Insect Venom Peptides. *In: Handbook of Biologically Active Peptides*. [s.l.]: Elsevier, 2006, p. 389–396. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/B9780123694423500593>>. Acesso em: 26 fev. 2023.

PALMA, Mario Sergio. Hymenoptera insect peptides. **Handbook of biologically active peptides**, p. 416-422, 2013

PACHECO, Christina et al. Pesquisa translacional na era pós-genômica: avanços na área da transcriptômica. **Saúde em Debate**, v. 43, p. 169-180, 2020.

PASSARINI, Ilaria; ROSSITER, Sharon; MALKINSON, John; *et al.* In Silico Structural Evaluation of Short Cationic Antimicrobial Peptides. **Pharmaceutics**, v. 10, n. 3, p. 72, 2018.

PATERSON FOX, Eduardo Gonçalves; CORREA BUENO, Odair; YABUKI, Antônio Teruyoshi; *et al.* General Morphology and Ultrastructure of the Venom Apparatus and Convoluted Gland of the Fire Ant, *Solenopsis saevissima*. **Journal of Insect Science**, v. 10, n. 24, p. 1–11, 2010.

PEREIRA, Franco Dani Campos. Caracterização molecular e expressão heteróloga do alérgeno fosfolipase A1 do veneno de *Polybia paulista* (Hymenoptera; Vespidae). 2012.

PIEK, Tom; DUVAL, Alain; HUE, Bernard; *et al.* Poneratoxin, a novel peptide neurotoxin from the venom of the ant, *paraponera clavata*. **Comparative Biochemistry and Physiology Part C: Comparative Pharmacology**, v. 99, n. 3, p. 487–495, 1991.

PRICE, Morgan N.; DEHAL, Paramvir S.; ARKIN, Adam P. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. **PLoS ONE**, v. 5, n. 3, p. e9490, 2010.

QUINET, Y; VIEIRA, Rhsf; SOUSA, Mr; *et al.* Antibacterial properties of contact defensive secretions in neotropical *Crematogaster* ants. **Journal of Venomous Animals and Toxins including Tropical Diseases**, v. 18, n. 4, p. 441–445, 2012.

RAHNAMAEIAN, Mohammad; CYTRYŃSKA, Małgorzata; ZDYBICKA-BARABAS, Agnieszka; *et al.* Insect antimicrobial peptides show potentiating functional interactions against Gram-negative bacteria. **Proceedings of the Royal Society B: Biological Sciences**, v. 282, n. 1806, p. 20150293, 2015.

RAIMUNDO, Rafael L G; FREITAS, André V L; OLIVEIRA, Paulo S. Seasonal Patterns in Activity Rhythm and Foraging Ecology in the Neotropical Forest-Dwelling Ant, *Odontomachus chelifer* (Formicidae: Ponerinae). **Annals of the Entomological Society of America**, v. 102, n. 6, p. 1151–1157, 2009.

ROBERT, Xavier; GOUET, Patrice. Deciphering key features in protein structures with the new ENDscript server. **Nucleic Acids Research**, v. 42, n. W1, p. W320–W324, 2014.

RODOVALHO, Cynara M; FERRO, Milene; FONSECA, Fernando Pp; *et al.* Expressed sequence tags from *Atta laevigata* and identification of candidate genes for the control of pest leaf-cutting ants. **BMC Research Notes**, v. 4, n. 1, p. 203, 2011.

SANTOS-JÚNIOR, Célio Dias; PAN, Shaojun; ZHAO, Xing-Ming; *et al.* Macrel: antimicrobial peptide screening in genomes and metagenomes. **PeerJ**, v. 8, p. e10555, 2020.

SAYERS, Eric W; BOLTON, Evan E; BRISTER, J Rodney; *et al.* Database resources of the National Center for Biotechnology Information in 2023. **Nucleic Acids Research**, v. 51, n. D1, p. D29–D38, 2023.

SCHENDEL; RASH; JENNER; *et al.* The Diversity of Venom: The Importance of Behavior and Venom System Morphology in Understanding Its Ecology and Evolution. **Toxins**, v. 11, n. 11, p. 666, 2019.

SMITH, Valerie J.; FERNANDES, Jorge M.O.; KEMP, Graham D.; *et al.* Crustins: Enigmatic WAP domain-containing antibacterial proteins from crustaceans. **Developmental & Comparative Immunology**, v. 32, n. 7, p. 758–772, 2008.

SNELLING, Roy R. A New General Catalogue of the Ants of the World. **American Entomologist**, v. 43, n. 2, p. 127–127, 1997.

STAŃCZEK, Sylwia; CYTRYŃSKA, Małgorzata; ZDYBICKA-BARABAS, Agnieszka. Unraveling the Role of Antimicrobial Peptides in Insects. **International Journal of Molecular Sciences**, v. 24, n. 6, p. 5753, 2023.

SOUZA, Bibiana Monson de. **Estrutura e função de mastoparanos dos venenos de vespas**. 2006.

STEENWYK, Jacob L.; BUIDA, Thomas J.; LI, Yuanning; *et al.* ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. **PLOS Biology**, v. 18, n. 12, p. e3001007, 2020.

THE UNIPROT CONSORTIUM; BATEMAN, Alex; MARTIN, Maria-Jesus; *et al.* UniProt: the Universal Protein Knowledgebase in 2023. **Nucleic Acids Research**, v. 51, n. D1, p. D523–D531, 2023.

TONK, Miray; VILCINSKAS, Andreas. The Medical Potential of Antimicrobial Peptides from Insects. **Current Topics in Medicinal Chemistry**, v. 17, n. 5, p. 554–575, 2016.

TONIONE, Maria Adelena; BI, Ke; TSUTSUI, Neil Durie. Transcriptomic signatures of cold adaptation and heat stress in the winter ant (*Prenolepis imparis*). **Plos one**, v. 15, n. 10, p. e0239558, 2020.

TORRES, Alba F. C.; HUANG, Chen; CHONG, Cheong-Meng; *et al.* Transcriptome Analysis in Venom Gland of the Predatory Giant Ant *Dinoponera quadriceps*: Insights into the Polypeptide Toxin Arsenal of Hymenopterans. **PLoS ONE**, v. 9, n. 1, p. e87556, 2014.

TOUCHARD, Axel; AILLI, Samira; FOX, Eduardo; *et al.* The Biochemical Toxin Arsenal from Ant Venoms. **Toxins**, v. 8, n. 1, p. 30, 2016.

VÉTILLARD, Angélique; BOUZID, Wafa. Fourmis : une chimiothèque de nouveaux anticancéreux. **Biologie Aujourd'hui**, v. 210, n. 2, p. 119–125, 2016.

VIRTANEN, Pauli; GOMMERS, Ralf; OLIPHANT, Travis E.; *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. **Nature Methods**, v. 17, n. 3, p. 261–272, 2020.

WAGHU, Faiza Hanif; BARAI, Ram Shankar; GURUNG, Pratima; *et al.* CAMP_{R3} : a database on sequences, structures and signatures of antimicrobial peptides: Table 1. **Nucleic Acids Research**, v. 44, n. D1, p. D1094–D1097, 2016.

WANANDY, Troy; GUEVEN, Nuri; DAVIES, Noel W.; *et al.* Pilosulins: A review of the structure and mode of action of venom peptides from an Australian ant *Myrmecia pilosula*. **Toxicon**, v. 98, p. 54–61, 2015.

WANG, Binwei; XIAO, Qiaoli; LI, Xun; *et al.* Proteinaceous Venom Expression of the Yellow Meadow Ant, *Lasius flavus* (Hymenoptera: Formicidae). **Toxins**, v. 15, n. 2, p. 106, 2023.

WASKOM, Michael. seaborn: statistical data visualization. **Journal of Open Source Software**, v. 6, n. 60, p. 3021, 2021.

WATERHOUSE, Robert M; SEPPEY, Mathieu; SIMÃO, Felipe A; *et al.* BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. **Molecular Biology and Evolution**, v. 35, n. 3, p. 543–548, 2018.

WESTBROOK, Anthony; RAMSDELL, Jordan; SCHUELKE, Taruna; *et al.* PALADIN: protein alignment for functional profiling whole metagenome shotgun data. **Bioinformatics**, v. 33, n. 10, p. 1473–1478, 2017.

WIESE, M. D.; BROWN, S. G. A.; CHATAWAY, T. K.; *et al.* Original article: *Myrmecia pilosula* (Jack Jumper) ant venom: identification of allergens and revised nomenclature: Allergens in *Myrmecia pilosula* (Jack Jumper) ant venom. **Allergy**, v. 62, n. 4, p. 437–443, 2007.

WU, Qi-xuan; KING, M.A; DONOVAN, G.R; *et al.* Cytotoxicity of pilosulin 1, a peptide from the venom of the jumper ant *Myrmecia pilosula*. **Biochimica et Biophysica Acta (BBA) - General Subjects**, v. 1425, n. 1, p. 74–80, 1998.

XU, Guangxin; CHEN, Li. Biological Activities and Ecological Significance of Fire Ant Venom Alkaloids. **Toxins**, v. 15, n. 7, p. 439, 2023.

YANG, Sung-Tae; SHIN, Song-Yub; SHIN, Sung-Heui. The Central PXXP Motif Is Crucial for PMAP-23 Translocation across the Lipid Bilayer. **International Journal of Molecular Sciences**, v. 22, n. 18, p. 9752, 2021.

YI, Hui-Yu; CHOWDHURY, Munmun; HUANG, Ya-Dong; *et al.* Insect antimicrobial peptides and their applications. **Applied Microbiology and Biotechnology**, v. 98, n. 13, p. 5807–5822, 2014.

ZELEZETSKY, Igor; PAG, Ulrike; ANTCHEVA, Nikolinka; *et al.* Identification and optimization of an antimicrobial peptide from the ant venom toxin pilosulin. **Archives of Biochemistry and Biophysics**, v. 434, n. 2, p. 358–364, 2005.

ZHANG, Zhenting; ZHU, Shunyi. Comparative genomics analysis of five families of antimicrobial peptide-like genes in seven ant species. **Developmental & Comparative Immunology**, v. 38, n. 2, p. 262–274, 2012.

ZHBANNIKOV, Ilya Y; ARBEEV, Konstantin G; YASHIN, Anatoliy I. rqt: an R package for gene-level meta-analysis. **Bioinformatics**, v. 33, n. 19, p. 3129–3130, 2017.

ZHOU, Xiang *et al.* Antennal transcriptome analysis and candidate olfactory genes in *Crematogaster rogenhoferi*. **Bulletin of Entomological Research**, v. 111, n. 4, p. 464–475, 2021.

9 . Apêndices

APÊNDICE A - Peptídeos preditos secretados em glândulas da peçonha de *Odontomachus chelifer*. Os peptídeos com atividade antimicrobiana estão indicados nesta tabela.

Acesso do transcrito	Sequencia	Grupo
TRINITY_DN10404_c0_g1_i1.p2, TRINITY_DN10404_c0_g1_i9.p2	MKRITYAVIVLVIIYICNIFLAISLGAKYCFRFTWPGPSN KTVNCTNDIPCVQPFYNSSVQPNTKNMWDENPNK YSCELNTDNVCIKYTYTYNNAGKTTYFYKITKENQIRF HSRNVYNL	PAM
TRINITY_DN25847_c0_g1_i1.p1	MLSRDIIIIILLSIFGLCFTASIADAVMCNCEHCDEEKVIG SEEETTPAWRVVMYKEDSDDDGNRITCAMDRDFNN RTFPSICYMSCYNHCTRYRVEVEGNDAKKHVLVAYR TNYKLRDGAC	PAM
TRINITY_DN22786_c0_g1_i1.p1	LFLSLSLFFSICLSSTLLFRSARTQSAENNKDVEGQ RLRLVQFAGRSSRKWLSASSASKETSASYDILDRLRS SSRKRSRRPRKVAGGTRFPAPKESSRIERSAILKSLR VH	PAM
TRINITY_DN12961_c0_g1_i1.p1	MARLINVVMIIIGMVLVSVFAAEELYSDKYDDIDVMGIL TNDRLRDQYFKCFMDTGPCTTADMKFYKDIIGEGVV TKCKKCTEKQKENLDTLDWYTKNRLDDWNTFVKKL MEDFKK	PAM
TRINITY_DN11980_c3_g1_i1.p1, TRINITY_DN11980_c3_g1_i2.p1	MGNIYVCSLLVVLATLLVAVQCGDIMRKSIVFDKNTPD VFYCPQHKPIGFKMLVKARPLSRLCQFEGRPIPEDY KSDCYNDVDETEYACKEKYRIMKRFSNPDGDAKAVD NDYVEK	PAM
TRINITY_DN17655_c0_g1_i1.p1	MRRFRAINLIAFCALVLLADDSSATVFYNLKNYLPGFL RSFKQLKTLRYGIVPGTLWCGPGNIASNYSELGVNW KLDTCRAHDSCNDLIRPRNSKYGLYNSSKLCSSLL CQCQLQFY	PAM
TRINITY_DN13371_c0_g1_i1.p1	AVSIGALLSFLGFVACICSCRRNENKTTSANRSLPDI PKDKSREHEDIAESVPQDIYETTETMGDHSELYATDQ DTGQEQISERETQEVSQQSSMTQDTS HQYARFAS	PAM
TRINITY_DN23731_c0_g1_i1.p1	FTMDYRSARTFAFFLIASSTSCGDVRRPRSAENYDAR KIASYINKFPFKYEFPLSSTKYTERIAGGSYANEGQFP FMAVVHRLIDGRRVAQCAGGTIISERWVLTAGHCVAKH PRRF	PAM
TRINITY_DN21289_c0_g1_i1.p1	RRGHFAEMKITPRATPVLLAIILLHDCRTAYTKSSLR VRAVDNDNDFLIAVNVTDCDRVKPFFDSKNISLPAK DLSKTICGGKCCDEEMEKQLKQARADFHNLIIHHS RTLQGLLITTA	PAM
TRINITY_DN14668_c0_g1_i1.p1	CVRACVGAFIILLGYFSYHTARTKLQTLHPGISSRDY KSSKRAFKRADTIDGFHRPRECRGACNNVAANLRG NEITANCEINEPVRVARQSVLTMRRANHPRYSYRING TVDN	PAM
TRINITY_DN18722_c0_g1_i1.p2	MNRNISLLAVVLLVASACAQFYKSGNCPLQSTVSN CTPRCMSDSQCSLNQKCCPNKCGHRSCADPSAVST GSGYKGSYQQDVYCVKCAAYEKQCFDRNTRRE KCTRA	PAM
TRINITY_DN13987_c0_g1_i1.p2	MEHTGKVFLLFTAIFVAVADENKEQDRPKTFRRLIPA DILRDFPGMCFSTKCATIEPTKSWELSPFCGRSTCV	PAM

TRINITY_DN28762_c0_g1_i1.p2	PADDNSGRLFELVEDCGPLPKANPKCKLSDKTNK MKAIVVLAICFAAALGEFTDEEKEKLKTYKESCISET GVDPNVVKNNAKEGMIDESDEKLACFSTCLFKKFGVM KENGIDIDTDARSKISSNVPQEEADRRIINMCK	PAM
TRINITY_DN17243_c0_g1_i1.p1	GMTYIVSLYWILVLLLEWLPVSLCDTVLSLCGRKQRIV AEYDRVRNALRSLKXISWRSWPAERKRKYQLQSRLT EKEREAFPVTAVIDIEAYILCVAAATRKYCVNEDNFKV	PAM
TRINITY_DN1958_c0_g1_i1.p2	MQIRVVFLLLIFSGSVFGKAAWNGVARSPRDNSLNS WDFEYKREKRDDKENNMNTAGTMNNTKTAIVKRDL VDSTLDSTQQLVDNLLDDLEILRTIGSLPI	PAM
TRINITY_DN3547_c0_g1_i1.p1	MSRFLFLLLIYVSTTESCIIQEKEKEDIRVIELDDSSS ESSEQEIVALHLPDRRRREIDSRARNSSCKTDEDCG PGAICLAYLTCVKGKWRITIDSLETTRAAN	PAM
TRINITY_DN9325_c1_g1_i1.p2	NRLTNVVFFLSVLGTMLKMTVCPYKCKRCSLDFPH EQLLQQHRLYHHGARERARSNARWNVETILSFPTYT CPHCDVITYTKPGNLKRHLMSGCGRDPNFRATSGK	PAM
TRINITY_DN1600_c0_g1_i1.p1	RESLCIICVCWITGLVGFPLLGWNTGEMSDHRCM FMKVMDFDYLVFLYFATIIIPALLIAVFYAHYRIVVKQL QQIVTMDPGGGGAAADCRLGGGLRRGRT	PAM
TRINITY_DN26054_c0_g1_i1.p1	HVSTGVCVLLLLFYKCVLPSLTLVSPSGSTLVRGRVD ASRTDLALSRASQFSRKISRRANCCGKDRNVPRHFP TMRCEFLAIFARIIHDHLSFVGLVGNGLK	PAM
TRINITY_DN2760_c0_g2_i1.p2	MRFYSVLSVLLLLVALTCVAEGGSLRARRSFFGDIGN TLNNVGETIKDGMTKGFESVFKRPEPETVTLQPDEI SLESRQLVIVPDRCPNHNHIVNKRCLRQR	PAM
TRINITY_DN10663_c0_g1_i1.p1	MLKMRWLAMIMCLYAGLASARTEEDQQASNTDGIPN SPPVQCFYRQAACPNIINSNECPCKITLANVPEP HAALCCNINNQTLDHARSCMGKQLQELFSKLS	PAM
TRINITY_DN17016_c0_g1_i1.p2	TATTTMMMMMMIMVMMMLHLAADSRYDDTHRA GWRSRKAASGVDFVSSLARTDASSFAASAKTRA NLGRSARRRYGGVQRSTAECVSVTFVWRLAYAART VLQT	PAM
TRINITY_DN28995_c0_g1_i1.p1	RNANISPPPLSLSLFLIIVPISTRRHDCSSYTTDGSV RFRQRLAAAFANLSIATISSCVIASERVGKEQRGKK KEKAQDRRCSVGALAYHGRENVRHFRVGTTPA	PAM
TRINITY_DN13204_c0_g1_i1.p1	CVVLVLVIVVVIVKVTDDTQEASRRASWKKRLTRT NSPPGINGPKKRPLIRTIKQADRKFLRPDNRAAAKRL HASTFVPQERSNGPATMLSADEKIFHP	PAM
TRINITY_DN11980_c3_g10_i1.p1, TRINITY_DN11980_c3_g2_i1.p1	MALDILCACVLLLFYYLTADFDFWKSQGVKGPKEPI PFFGTIARHLLRRLSLGSLRKIYEEYPNEPMVGIFQR GKPALVLRDPELIKQVLIKDFSVFPERN	PAM
TRINITY_DN18603_c0_g1_i1.p1	MLAASPLLVLVLYSPRSAGFFREGSLFLRRLSRTASF FLNLRNHTFSTTRSFLVPDLLRCPVSAAPPELGDRPR SQSRSTRSSFAFFLPPSRSYPTIRGT	PAM
TRINITY_DN3796_c0_g1_i1.p1	MKLLAVFALFCVLAYATADPIPAVYDGPTDELNPFEES ASDEVSENDMPMIRQRRVTCDLLSWNSKWFSFNH SACAAKCLAQRRRGRCRGGICVCR	PAM
TRINITY_DN6260_c0_g2_i1.p1	STTMKTCTIVFVALAALLLCTECVSAMELCPQENCLT PSKCEQEINNLVQCFEQGTSCCSVVKDEFRTHCR HFGGECMNRCPQSLLHETVDCINGQVCCVLV	PAM
TRINITY_DN272_c0_g1_i1.p1	MHFLLYTLLFSAFAAVSNVIAEFSADDCKMLGFNKAN LLCSTCEHFAEPNLAKILATCKECLKNDYDLSGSK RYPKAVLEVCTCKFGAYPQIQGIYLN	PAM
TRINITY_DN2382_c0_g1_i1.p2	VVVIIGSGSRVHVCLYVCVCLSRVAAPCIDKYARV	PAM

	CAQVRRKIHGRNSSQVQPATWASRPLNTSRTSTRFP LRSRNRRTERRRLDPSVGKFLRARRVFPYCPGE	
TRINITY_DN11838_c2_g3_i1.p3	MNPLIACMIIGLAGLTLAEPVSGYSYRPTGGGGYF GGGGGGGGGGFGGGGGYTQVSTGYQTSEGASV DGALLEQIRQILLKEELQSQQSGGV	PAM
TRINITY_DN15852_c0_g1_i1.p2	MNRIAITFALLCALSTLAFNFEDCGSEIGKFGDITISS CTPSSKKCIVNRGDEAHVSIKYTPNKDISNLGARVYT MLQNVPIPFPLDKPNVCKDSN	PAM
TRINITY_DN12031_c2_g3_i1.p5, TRINITY_DN12031_c2_g3_i2.p5	MPSSSSSLLGYGIGMVIISMIIVIDTITVTVADIDTTVT VTVIDDTPSSSHSPSSANTKTRLKHGHYIAEYQPIVE YIRIATDLIMIKRTRKRWKKQRETDARRYTWLRT	PAM
TRINITY_DN307_c0_g1_i1.p1	MIGIEVSTVMWQYRTIGLLALLAACVQARGIELLSK RGEGASSFSSASSASASSYGGGSFSKSDAKAIG SASANSGGYGGFGGGSGDLGGGGHGGFGG	PAM
TRINITY_DN24006_c0_g1_i1.p1	MKLLVCLVFAAISVATGQNRVFAKQGSPPALPVQ FCGHSCFVDSHCAGIGKCCPTQCGGSICMPVTMT RPSQSGEKSFVAALRQIARS	PAM
TRINITY_DN4617_c0_g1_i1.p1	SPSASASTMHRPRGRQSFLFCLFFYTSCEFLWGTG RYISVLRPAAVITGGAETGVILAPTLRATNSLIGNNGG DTARKRRGEPVNWLRVSAKVLSLAKG	PAM
TRINITY_DN11919_c4_g8_i3.p6	ECAVRACRIAVCLVGTISRVCVSSFVSLAVICCWRC LGHELAARREHRRASLVQSKADGVADGGGDFDGD GGGGGGDGRPPRRGGTTRDVYE	PAM
TRINITY_DN11708_c0_g1_i3.p5	MGRDELVLIVFLIITQLVVFSDDIDYTLHDEMRL GEKTGIRKCLAEIGDVTVSMQLDQAELEFPDNNL ACYFKCVMEKGGVNETKLQR	PAM
TRINITY_DN9306_c0_g3_i1.p2	MDHNQLLIAALLATLLHGMIMSTAEPFRLVHDLIQY NIAGAPVAHQKTEWDFDPDIGKQRRVRYQQENGRF GEIAIAKIGMGIGYKGPWGPVD	PAM
TRINITY_DN11674_c8_g1_i1.p17, TRINITY_DN11674_c8_g1_i3.p16, TRINITY_DN11674_c8_g1_i6.p17, TRINITY_DN11674_c8_g1_i8.p16	MWCRLTLLLLVCCCHLAAASKRIRVDGKSYTVSAYQ EDCRKYVMCSEGRCKLETCEPTYFFDPITSTCTHRA RGSSNCDFKLSRKSRRG	PAM
TRINITY_DN5545_c0_g2_i1.p2	PRPVIPCVALACCARESASAPPGESVYERVCEKAKQ ASGLRRREEKRAEERLSRRWPTRRRRRGASSGFRP VGAEAAHQPNSSIERGCS	PAM
TRINITY_DN11969_c3_g2_i1.p11, TRINITY_DN11969_c3_g2_i2.p10	IIALYRGHELRFSSAYLRALFCFAFLLSRTGRASGGS CTSRRRDHRRRIPSDRSRANRRRDIEVPISTVAS HGGSKELSSGAEIDSAPLLVRDTGKKRK	PAM
TRINITY_DN8567_c0_g1_i1.p1	LSLSLCLPFCLIPPLLVLSLSPSPRECRRGRAGV GHVPGPTTVTTKTTQTKTAVVTVANHDLVCVTEQK HPGYHPEGKDRRIA	PAM
TRINITY_DN16791_c0_g1_i1.p1	MKAILSVALLMCLAMVYSYTETNFELNCPIDKENATL LPHPIYCDSFFICQNGEPIPMKCPKGLDFNPELSVCD WPWQAHCKASG	PAM
TRINITY_DN24839_c0_g1_i1.p1	MKGLLCCAILLLVCVLRVQGNCDHNGEILPMGTHYR DCQQIKCNEDGMSALGCPLSSCKLGEQIGYHKLDV NKPYPECCERPICKTD	PAM
TRINITY_DN10676_c0_g2_i1.p1	MSDPTLLVLQFLAAFLACLADYSRAQDEETLVECVE SETGCRRCRDGTCARCAVLLHQGTCVHTCPPGFVA DWSTRDEYMGRICRETGYMFGL	PAM
TRINITY_DN12005_c2_g1_i1.p4	SLSLSLSLSLFLFTFSSLFSDPRVQIVRAIIAISLD ERHSGERAFPSHHVPRELLYPNGPCARASVLFKGG ELVRGESMIYTSRRDR	PAM

TRINITY_DN27555_c0_g1_i1.p1	VVVVVVVVVVALVPSGSRRICQARKKDSPFSTRIPP PPPPSTSFQSRFSWGRVGGSKDRRRGWSWFTHRDG ALSSKPVKALGDDRL	PAM
TRINITY_DN1241_c0_g1_i1.p2, TRINITY_DN1241_c0_g2_i1.p2	MTAFLHGFLMIVLIIINAFTLGISESFERSPTFTANHPIF PRHYPTTSKIASTPLIYTTPRIIDTFAKSEVKCPTGQKK DSQGICRERA	PAM
TRINITY_DN11940_c4_g5_i1.p5	VCVCVCVCVCVCVRTGVYVRACAATRRRWSPAMFG LRHHTSHTSSSPRTPHYHHHFHHHHHGLMAASGTS RSRSIPPSCTR	PAM
TRINITY_DN11680_c10_g5_i1.p5	MYIYACVCAVCVCWACMHMYVCAPCVCTRVELVQE ACPSVKITLSPRRLRRENRRGRDASRRMEKECCEGR SIESKEREREKDNSLHEN	PAM
TRINITY_DN12094_c0_g1_i8.p3	MDSRLLVIVLLIVSILFVCSSGGKSAGGFLSKIVSCIIF SLAVDRNFSCDESVCVILVATVTVVCARWCERYQIL RVRVLRQR	PAM
TRINITY_DN11280_c0_g1_i11.p7, TRINITY_DN11280_c0_g1_i5.p7, TRINITY_DN11280_c0_g1_i6.p7	MLCDMFDQIDAVAFVALLINQYSSCRPISHLHHIRCAL NCTQSYTLARKFSSNFCKQKRNFQAKLMDAFATTCV ARDAQCFRYLYVGV	PAM
TRINITY_DN10507_c0_g1_i1.p1	MKSFACILVIMVNLVLPEDLCQQMNNPSNQPMGQG MSPPNRADMSRPMGSGMYRPMGGWSRPILRRGPV MRGGMPMTGAGTM	PAM
TRINITY_DN5860_c0_g1_i1.p2	PPIDDEYMGGICAAFLFQLSFATTSTTIVSGAMAERC NFKAYCIFSFLNTIVYCLPAGWVWGDHGFLNRMGVV DIAGSGPVHLV	PAM
TRINITY_DN11013_c2_g1_i1.p3, TRINITY_DN11013_c2_g1_i3.p3, TRINITY_DN11013_c2_g2_i1.p1	MNGHVANVVPVLLLAEAIGLAMMPTTLRHQSRRGLF SFWVSASQSSGEVLDDGGDESRLVDDALARSRQA EYEILDEIRDPSVGEDEEPPGDGARHSLCAVRPSGA A	PAM
TRINITY_DN12094_c0_g1_i18.p4, TRINITY_DN12094_c0_g1_i4.p2, TRINITY_DN12094_c0_g1_i5.p3, TRINITY_DN12094_c0_g1_i7.p4	MDSRLLVIVLLIVSILFVCSSGGKSAGGFLSKIVSCIIF SLTVDRNFSCDEPVLRYPCCNQSDGCLRSLVSLPNS QSANFS	PAM
TRINITY_DN4750_c0_g1_i1.p4	MITRWRLIFVLILSPSIVLSRLYLHGVSRRNNSMFFLL EITARKRRRNFEIARDTASLLASKVFCGRRSAAAFSQ LSLCT	PAM
TRINITY_DN24344_c0_g1_i1.p2	FCLLFLIFFECISISWAFGVNRFYDALRDMIGYYPLMW WKFCWTVTTPMICVGVFVFNLIQWTPVKYLDYEYPW WS	PAM
TRINITY_DN23671_c0_g1_i1.p2	GTRALAALALSCIQSGCIVRRGRCCAMHSVCGVTTR DTSRVLKACQLEVHPELVHTSAFDGFFERPRAVKFG ELCE	PAM
TRINITY_DN12012_c2_g6_i1.p4	MMMMMMMMMMMMMMMMTESENESESEKGRSMKM FLLPNVFAVVAANATIRRKKCDKCTFGGEKKRLSRE TALFRISSARRAADERIPLIVLTVSQFFSNPVSA	PAM
TRINITY_DN15024_c0_g1_i1.p3	DDVRRFRLSLVGIVSLLIDASCIGVSGTGGGGGSG GGSTCGLAELRCNDGHCVPIDAYCNGEDDCGDGSD EPAMCTPC	PAM
TRINITY_DN1980_c0_g1_i1.p1	IVLVLLLECWWIGLACGFRDAEKGTDQDFLKSTIKYYATT ADKAETVTVAWDGMISQFHCCGVESYRDFSEN	PAM
TRINITY_DN15646_c0_g1_i1.p3	MRCVRCWWLVKCVFACLTRGGRTGGVVFYRLQ NDEDSQDSSASEALDHRAQIYRGTSRPITRRLSALRP TWYARVCG	PAM
TRINITY_DN1540_c0_g2_i1.p1	AAVFAVLLCVTDVTAIRSGDGCNGMFRNCNDGKCIQ SILVCDYRGDCDDNSDEMQSCPPDCDFGQITCG	PAM

TRINITY_DN11631_c2_g1_i1.p2	DNIIVIIIIIMIRIIAARQQDFHTVAKLLWSPSSPSLYKS QRRAAEKNHSRDEESTHRVYRESAHRGCISF	PAM
TRINITY_DN11193_c0_g4_i1.p2, TRINITY_DN11193_c0_g4_i4.p2	RYFHCSLRFPLCLASDFIVCFDLSRPKKEKKLVISKRY KINTKHEKKVNRVNSSQETTRGKISVVYILLHLSLAAD GKKKNCSVHQTGKARNRKNRDGSRVIARGIKDTRQ GIVNRQDRLSTFPVGCARQL	PAM
TRINITY_DN26077_c0_g1_i1.p1	IHPTVFSVLFWLGYCNSAINPCIYALFSKDFRFAFKRII CKCFCKRRANTLRRGSDGSQLAMRNERSPSYS	PAM
TRINITY_DN26693_c0_g1_i1.p3	MEENEVRGILVPVLAACIGPALSITSCARNVVCISRIT MRAPRDIPRDRPVDFHGAISMQIVLSPFATDTTPRCD LR	PAM
TRINITY_DN4362_c0_g1_i1.p2	MKLIIVLMAIAVVLSTTVTGQTDDEYIHLPGKSCSDAP PCPDGRPCVMAPPRCNVGTGCKDLVPTCGRK	PAM
TRINITY_DN23188_c0_g1_i1.p1	FARIFVFFSLLTSLVKMATASTGIGVSFYARAALHAAT WCYTVIHRARYIEVSTFEVDIFASCTRKRNNFKTT	PAM
TRINITY_DN450_c0_g1_i1.p1	MIKTVFLVALLCMSVIAEDGMFQAAADAVSSGIRSMS SGGGGASSLTRRSQAQSDRMPQGGPPGSSGPPRN	PAM
TRINITY_DN9862_c0_g1_i1.p1	MRLSLAVMCAITLFFFLSTSNARAASFDPQNEYFQR FDEKPKISTPKASSIINVPRNCPPEKVKIGDRCRSVY	PAM
TRINITY_DN9582_c1_g1_i1.p2	SLSLALSFFFFLTTRPKTIVAASVAVGDDIPRLGERKR CEKRRLKSTRRPGMKPNFRASAPQPSRFARSERV	PAM
TRINITY_DN11844_c0_g3_i3.p1, TRINITY_DN11844_c0_g3_i5.p1	MKFYVCILLVIMILSLVPEDLCQEEDGRSRLGELRNYP IVVMSLKAKPVCPEGQRLAMEKCRPGRK	PAM
TRINITY_DN22691_c0_g1_i1.p1	WHPRTATVLLFFSLVSSPLPRGPGSHLLCIHSLMSR DQALPTEFFAPTNGVLGRSPHRLSPACSSCA	PAM
TRINITY_DN29460_c0_g1_i1.p2	FKFSRNTLICVYVPVFCVYAVNAVYTMCVRVCARVC VNADCAIVIFLRLHPAIRRVGRAEERDVERKR	PAM
TRINITY_DN11122_c0_g2_i1.p4	LSLSLFLFLSPLITLCLLFLPLCHRKGGGQERGIKEG ARGERRVPSAFDPVILVGGGAPP	PAM
TRINITY_DN168_c0_g1_i1.p2	NSTFILDYDRITGMIVALFVIATVTSSSRDYFRARAVDK LKCRSAATMTSLNRGTVIATFRSRLTREKQPRA	PAM
TRINITY_DN10137_c1_g1_i1.p5	FFFFFFFFFLLCDSPVPNLNCRWKERRKKEQSAAS SRLACFVSTVVCLVCPETEKYR	PAM
TRINITY_DN13289_c0_g1_i1.p2	MRATWLVLILAEILPPLSLLGEAVCHRNDTDLASAGTNT GQRSGETSSRRRRELAFPKGSFAV	PAM
TRINITY_DN6307_c0_g1_i1.p1	ARQTALVLFGDFIVARARDKTSDLLRTADFLRLVERRP KIDTPRLKIAKLARPCAPRWWW	PAM
TRINITY_DN12066_c4_g1_i1.p1	MKHSVKQLFLFLVFVFLRFLKPLVRHKLRTKQPVEFI RNETFFIRGTTKFRSGRFRETSAA	PAM
TRINITY_DN12925_c0_g1_i1.p2	MKISNLLAFVAIIVALFVSVTSADPAPEAVKLFHRHRT VGRIGAVVALRPKPKSTTTVKPTTPKKG	PAM
TRINITY_DN2395_c0_g1_i1.p1	VLLYSARSNMILNISCTLLCLAAGTVVGGAVLLFLGL YLYISFTRGVCKSKTRMDGKTVIITGCTSGIGKET	PAM
TRINITY_DN21408_c0_g1_i1.p1	QGTSESSDIMVSLGRAVIVAAALMALTAHVHAIPTPA PVAPGAITAYPIQRPAYAPNPWVYKPTTVQPIGR	PAM
TRINITY_DN10891_c1_g1_i1.p4, TRINITY_DN10891_c1_g1_i1.p3	SEFFHNYDFLCVVCVYSRAPSGRAVERVQQVQRK AKKGQREGHPVFSQQGILLSLFTGLIKRAAT	PAM
TRINITY_DN12023_c1_g5_i1.p2	MKIFLNVFLIIAIAIPLGICKPEGKSAIKEVFNRNCP KGYRKDIHGICRPIINL	PAM
TRINITY_DN27221_c0_g1_i1.p2	MRDVDFLIRLFAFLVLRQGNCECFSPSRCLSLSR EVPKFKFYRSKWNSKCHKFITIFLYL	PAM

TRINITY_DN7631_c0_g1_i1.p2	VTEYNLYNLISTMFVIAIISSDVPTSRAKKRVNRLAND CLRDTFHSQRSARREVLNRNSGRNYVRCDD	PAM
TRINITY_DN3941_c0_g2_i1.p8	MCLSLTLSFCLRLVVLFTFTRRQSPGLYAQRRIQSTH APKKVPLAKVHTTLEPRHTGAHTAPRG	PAM
TRINITY_DN27940_c0_g1_i1.p1, TRINITY_DN27940_c0_g2_i1.p1	MAPARLLLLLIVSLMLVFANEGMPLSFPKTRTPQSKC PPGYRHTRTGCREILTY	PAM
TRINITY_DN11964_c1_g7_i1.p3	MIVVVVVFVVVLLKSFPPLHDSRCKTRAFRALSVDR AYVRRPRDIGPNF	PAM
TRINITY_DN16370_c0_g1_i1.p1	VVEDIQIFLQSTMKVLFNICLIVLTIIAISMGISQPLPPT PSGHVINKRVAPIRPVPTVIGLLMKQ	PAM
TRINITY_DN12005_c4_g1_i1.p15, TRINITY_DN12005_c4_g4_i1.p15	MLCVCMCVCAIQVDEREKEGGKGIKDKKILKIKTQ ILTAKKRHSF	PAM
TRINITY_DN20021_c0_g1_i1.p2	MRLSLYIVVFLLIIMATTAFIVPIKLRWKQKAKVGVQL RRPKTVHLPQLPIQ	PAM
TRINITY_DN8792_c0_g1_i1.p3	DIRDWRLAILIFLSTREFTDGTMSLKLQKRLAASVM RCGKKKSMVRSK	PAM
TRINITY_DN11886_c2_g8_i1.p8	EREREREGENVCMVCVCVCVCARACVCTRVRTTC VWTHVRSICVCLACVNV	PAM
TRINITY_DN19441_c0_g1_i1.p1	MLQGLCFLAVASLVNGIRKTRKYMFKILAKYSKVRY VYRMAFD	PAM
TRINITY_DN24173_c0_g1_i1.p2	MILVTLVVLFLGWKLRTIFELSSSSSAFLVWKHAG CCVPRI	PAM
TRINITY_DN1222_c0_g1_i1.p3	PLNARNVASVFLSLLLEALRADQTSRPGAAVTKRR NTLGSQKKIK	PAM
TRINITY_DN11861_c3_g2_i1.p5	MVFNLLLVNILLIYLTHIFGLRACHALVSTSRVLRVVI	PAM
TRINITY_DN10866_c0_g4_i1.p9	TTNSSVPVNGLVISVLCYFAAPFSPIVSRVPRFEIFQV TRKKKSAIC	PAM
TRINITY_DN13404_c0_g1_i1.p2	LILVVVICIFANPTDSIRPKKNRQRRYQSLAGYR	PAM
TRINITY_DN11793_c4_g1_i1.p4, TRINITY_DN11793_c4_g2_i1.p4	MSVGLLTQAFLLIICREIVKTKMKKTIALLWN	PAM
TRINITY_DN1481_c0_g1_i1.p6	MYLMLGMALIAMCFNLMQEEVIKMRALMRGIKYVF RCER	PAM
TRINITY_DN11610_c0_g2_i1.p3	RASSVLAAPGVFPFCFAVSYISLITAGLSPRRRADRR VSRDFT	PAM
TRINITY_DN4260_c0_g1_i1.p3, TRINITY_DN4260_c0_g2_i1.p3	MLSVPHTGLYSVAIQGHAVVVLVVLVGLSKLLLD KLFCGKAKATK	PAM
TRINITY_DN6759_c0_g1_i1.p7, TRINITY_DN6759_c0_g2_i1.p7	MTLLAWIFLFSVDGPGSPGRIPHAAFRFGLLK	PAM
TRINITY_DN11202_c0_g2_i3.p14	MTQMDMAMAVTTVATVVGAAHVVKVILASFL	PAM
TRINITY_DN17524_c0_g1_i1.p3	DLPPALLLPLFLLHFFLFRLSLSPSPAPVRVHVVRS VAR	PAM
TRINITY_DN12050_c5_g5_i1.p4	MTVRTLSNIILVVALISLVTLIPAKLKKAILNPAMVI	PAM

APÊNDICE B - Resultado do mapeamento cruzado da busca de nucleotídeos usando os transcritos que codificam os 136 PAMs previstos da glândula da peçonha de *O. chelifer* contra o transcriptoma de seu corpo. A pesquisa referenciada usou BlastN.

ID da transcrição da glândula <i>O. chelifer</i>	Tamanho da sequência	ID transcrito do corpo de <i>O.</i> <i>chelifer</i>	Similaridade (%)
TRINITY_DN12031_c2_g3_i1	2902	TRINITY_DN992_c1_g2_i2	99,89
TRINITY_DN12031_c2_g3_i2	2890	TRINITY_DN992_c1_g2_i1	99,89
TRINITY_DN11122_c0_g2_i1	579	TRINITY_DN981_c0_g1_i9	100
TRINITY_DN1241_c0_g2_i1	672	TRINITY_DN979_c0_g1_i5	100
TRINITY_DN1241_c0_g1_i1	750	TRINITY_DN979_c0_g1_i4	100
TRINITY_DN15852_c0_g1_i1	292	TRINITY_DN9659_c0_g1_i2	99,32
TRINITY_DN1958_c0_g1_i1	469	TRINITY_DN958_c0_g1_i9	99,07
TRINITY_DN450_c0_g1_i1	418	TRINITY_DN899_c0_g1_i1	100
TRINITY_DN17524_c0_g1_i1	255	TRINITY_DN8975_c2_g3_i1	100
TRINITY_DN11980_c3_g10_i1	830	TRINITY_DN8908_c0_g1_i2	100
TRINITY_DN11980_c3_g2_i1	816	TRINITY_DN8908_c0_g1_i1	100
TRINITY_DN12005_c2_g1_i1	2056	TRINITY_DN863_c0_g5_i1	100
TRINITY_DN28762_c0_g1_i1	375	TRINITY_DN82_c0_g1_i4	99,73
TRINITY_DN13204_c0_g1_i1	305	TRINITY_DN8074_c0_g1_i3	100
TRINITY_DN2395_c0_g1_i1	222	TRINITY_DN804_c0_g1_i1	100
TRINITY_DN2760_c0_g2_i1	1002	TRINITY_DN800_c0_g1_i1	100
TRINITY_DN14668_c0_g1_i1	345	TRINITY_DN7956_c0_g1_i1	100
TRINITY_DN6307_c0_g1_i1	282	TRINITY_DN7423_c0_g1_i1	100
TRINITY_DN21289_c0_g1_i1	371	TRINITY_DN729_c1_g3_i1	100
TRINITY_DN5545_c0_g2_i1	355	TRINITY_DN7267_c0_g2_i1	100
TRINITY_DN1600_c0_g1_i1	314	TRINITY_DN7216_c0_g1_i1	100
TRINITY_DN17243_c0_g1_i1	336	TRINITY_DN703_c2_g1_i1	100
TRINITY_DN12012_c2_g6_i1	1674	TRINITY_DN6652_c3_g1_i1	99,82
TRINITY_DN4750_c0_g1_i1	1872	TRINITY_DN6611_c0_g1_i2	100
TRINITY_DN16370_c0_g1_i1	205	TRINITY_DN646_c0_g4_i1	100
TRINITY_DN12094_c0_g1_i4	1422	TRINITY_DN632_c0_g1_i14	92,81
TRINITY_DN12094_c0_g1_i5	3041	TRINITY_DN632_c0_g1_i14	92,79

TRINITY_DN12094_c0_g1_i7	3062	TRINITY_DN632_c0_g1_i14	96,06
TRINITY_DN12094_c0_g1_i8	1532	TRINITY_DN632_c0_g1_i14	92,76
TRINITY_DN12094_c0_g1_i18	3425	TRINITY_DN632_c0_g1_i14	96,69
TRINITY_DN9325_c1_g1_i1	810	TRINITY_DN615_c0_g2_i1	100
TRINITY_DN11193_c0_g4_i1	1608	TRINITY_DN5909_c0_g1_i2	100
TRINITY_DN11193_c0_g4_i4	1587	TRINITY_DN5909_c0_g1_i2	100
TRINITY_DN11631_c2_g1_i1	1108	TRINITY_DN587_c2_g1_i1	100
TRINITY_DN11202_c0_g2_i3	1537	TRINITY_DN5748_c0_g1_i1	99,92
TRINITY_DN24173_c0_g1_i1	390	TRINITY_DN5560_c0_g1_i2	100
TRINITY_DN11680_c10_g5_i1	4552	TRINITY_DN52_c8_g1_i1	99,97
TRINITY_DN11940_c4_g5_i1	1406	TRINITY_DN515_c0_g10_i1	100
TRINITY_DN18722_c0_g1_i1	1359	TRINITY_DN4996_c0_g1_i2	99,85
TRINITY_DN22786_c0_g1_i1	615	TRINITY_DN498_c1_g1_i1	100
TRINITY_DN18603_c0_g1_i1	841	TRINITY_DN489_c0_g2_i3	99,88
TRINITY_DN9582_c1_g1_i1	392	TRINITY_DN4815_c0_g1_i1	100
TRINITY_DN11861_c3_g2_i1	895	TRINITY_DN4733_c0_g2_i1	100
TRINITY_DN1980_c0_g1_i1	216	TRINITY_DN4600_c0_g1_i2	100
TRINITY_DN272_c0_g1_i1	948	TRINITY_DN460_c0_g1_i2	100
TRINITY_DN11793_c4_g1_i1	1105	TRINITY_DN4518_c0_g2_i1	99,72
TRINITY_DN11793_c4_g2_i1	1108	TRINITY_DN4518_c0_g2_i1	100
TRINITY_DN1481_c0_g1_i1	496	TRINITY_DN4405_c0_g2_i2	98,8
TRINITY_DN4362_c0_g1_i1	614	TRINITY_DN4151_c0_g2_i1	100
TRINITY_DN6759_c0_g1_i1	875	TRINITY_DN4044_c0_g1_i1	99,54
TRINITY_DN6759_c0_g2_i1	878	TRINITY_DN4044_c0_g1_i1	99,88
TRINITY_DN7631_c0_g1_i1	204	TRINITY_DN3883_c1_g1_i1	100
TRINITY_DN8792_c0_g1_i1	591	TRINITY_DN385_c0_g1_i4	99,44
TRINITY_DN9862_c0_g1_i1	460	TRINITY_DN372_c0_g4_i1	100
TRINITY_DN24839_c0_g1_i1	887	TRINITY_DN3702_c0_g1_i2	100
TRINITY_DN4617_c0_g1_i1	730	TRINITY_DN3662_c0_g1_i1	79,38
TRINITY_DN11919_c4_g8_i3	5144	TRINITY_DN3525_c2_g1_i4	99,98
TRINITY_DN24006_c0_g1_i1	384	TRINITY_DN3485_c0_g1_i2	100
TRINITY_DN11964_c1_g7_i1	970	TRINITY_DN3317_c0_g1_i1	100
TRINITY_DN26693_c0_g1_i1	730	TRINITY_DN327_c0_g3_i2	80

TRINITY_DN307_c0_g1_i1	447	TRINITY_DN319_c0_g1_i46	99,55
TRINITY_DN11610_c0_g2_i1	280	TRINITY_DN3173_c1_g1_i2	99,64
TRINITY_DN10676_c0_g2_i1	431	TRINITY_DN3015_c0_g1_i5	100
TRINITY_DN13404_c0_g1_i1	260	TRINITY_DN28943_c0_g1_i1	100
TRINITY_DN28995_c0_g1_i1	325	TRINITY_DN2776_c0_g1_i2	100
TRINITY_DN17655_c0_g1_i1	357	TRINITY_DN26944_c0_g1_i1	99,72
TRINITY_DN3796_c0_g1_i1	1357	TRINITY_DN26725_c0_g1_i1	100
TRINITY_DN21408_c0_g1_i1	426	TRINITY_DN26693_c0_g1_i1	100
TRINITY_DN12961_c0_g1_i1	572	TRINITY_DN2663_c2_g1_i1	98,14
TRINITY_DN11013_c2_g2_i1	1749	TRINITY_DN264_c2_g1_i3	100
TRINITY_DN11013_c2_g1_i3	2785	TRINITY_DN264_c2_g1_i2	100
TRINITY_DN11013_c2_g1_i1	2767	TRINITY_DN264_c2_g1_i1	100
TRINITY_DN25847_c0_g1_i1	562	TRINITY_DN25488_c0_g1_i1	100
TRINITY_DN22691_c0_g1_i1	205	TRINITY_DN25445_c0_g1_i1	100
TRINITY_DN27221_c0_g1_i1	318	TRINITY_DN25251_c0_g1_i1	77,03
TRINITY_DN11674_c8_g1_i3	6711	TRINITY_DN249_c0_g2_i6	100
TRINITY_DN11674_c8_g1_i1	11824	TRINITY_DN249_c0_g2_i5	99,99
TRINITY_DN11674_c8_g1_i8	7449	TRINITY_DN249_c0_g2_i3	99,98
TRINITY_DN11674_c8_g1_i6	11086	TRINITY_DN249_c0_g2_i1	100
TRINITY_DN4260_c0_g1_i1	548	TRINITY_DN248_c0_g1_i1	100
TRINITY_DN4260_c0_g2_i1	672	TRINITY_DN248_c0_g1_i1	100
TRINITY_DN23671_c0_g1_i1	237	TRINITY_DN24747_c0_g1_i1	100
TRINITY_DN29460_c0_g1_i1	218	TRINITY_DN2465_c1_g2_i1	100
TRINITY_DN10507_c0_g1_i1	699	TRINITY_DN245_c0_g1_i3	99,09
TRINITY_DN11844_c0_g3_i3	593	TRINITY_DN245_c0_g1_i21	96,27
TRINITY_DN11844_c0_g3_i5	976	TRINITY_DN245_c0_g1_i21	91,9
TRINITY_DN10891_c1_g1_i1	1576	TRINITY_DN23575_c0_g1_i1	100
TRINITY_DN168_c0_g1_i1	684	TRINITY_DN22904_c0_g1_i1	99,38
TRINITY_DN20021_c0_g1_i1	475	TRINITY_DN22372_c0_g1_i1	100
TRINITY_DN10404_c0_g1_i1	2419	TRINITY_DN21960_c0_g1_i1	99,91
TRINITY_DN10404_c0_g1_i9	2289	TRINITY_DN21960_c0_g1_i1	99,89
TRINITY_DN11708_c0_g1_i3	4873	TRINITY_DN2093_c0_g1_i2	99,98
TRINITY_DN2382_c0_g1_i1	1275	TRINITY_DN1999_c0_g2_i1	99,92

TRINITY_DN12925_c0_g1_i1	1049	TRINITY_DN1898_c0_g1_i4	100
TRINITY_DN11886_c2_g8_i1	1657	TRINITY_DN1881_c0_g3_i1	100
TRINITY_DN23731_c0_g1_i1	346	TRINITY_DN1824_c0_g1_i4	100
TRINITY_DN13371_c0_g1_i1	328	TRINITY_DN18121_c0_g1_i1	100
TRINITY_DN11980_c3_g1_i1	3389	TRINITY_DN1801_c0_g1_i7	99,92
TRINITY_DN11980_c3_g1_i2	3790	TRINITY_DN1801_c0_g1_i6	100
TRINITY_DN26077_c0_g1_i1	214	TRINITY_DN17875_c0_g1_i1	100
TRINITY_DN8567_c0_g1_i1	265	TRINITY_DN1678_c2_g1_i1	100
TRINITY_DN24344_c0_g1_i1	231	TRINITY_DN15893_c0_g1_i1	100
TRINITY_DN10663_c0_g1_i1	1043	TRINITY_DN1563_c0_g1_i5	100
TRINITY_DN13987_c0_g1_i1	381	TRINITY_DN15490_c0_g1_i1	100
TRINITY_DN11280_c0_g1_i5	3243	TRINITY_DN147_c0_g2_i17	100
TRINITY_DN11280_c0_g1_i6	3758	TRINITY_DN147_c0_g2_i17	100
TRINITY_DN11280_c0_g1_i11	3684	TRINITY_DN147_c0_g2_i17	100
TRINITY_DN17016_c0_g1_i1	343	TRINITY_DN14572_c0_g1_i1	100
TRINITY_DN23188_c0_g1_i1	462	TRINITY_DN1411_c1_g2_i1	100
TRINITY_DN1540_c0_g2_i1	210	TRINITY_DN1400_c0_g1_i1	100
TRINITY_DN19441_c0_g1_i1	334	TRINITY_DN13982_c0_g1_i1	99,68
TRINITY_DN11969_c3_g2_i1	5923	TRINITY_DN13556_c0_g1_i1	95,82
TRINITY_DN11969_c3_g2_i2	5217	TRINITY_DN13556_c0_g1_i1	95,82
TRINITY_DN27555_c0_g1_i1	880	TRINITY_DN13542_c0_g1_i1	100
TRINITY_DN12050_c5_g5_i1	846	TRINITY_DN1333_c0_g1_i7	100
TRINITY_DN12023_c1_g5_i1	402	TRINITY_DN1305_c0_g1_i8	98,51
TRINITY_DN27940_c0_g1_i1	425	TRINITY_DN1227_c0_g1_i1	99,52
TRINITY_DN27940_c0_g2_i1	430	TRINITY_DN1227_c0_g1_i1	98,35
TRINITY_DN15646_c0_g1_i1	508	TRINITY_DN12152_c0_g1_i1	100
TRINITY_DN9306_c0_g3_i1	2808	TRINITY_DN1207_c0_g1_i4	99,96
TRINITY_DN10866_c0_g4_i1	4237	TRINITY_DN1199_c1_g1_i1	99,79
TRINITY_DN12066_c4_g1_i1	568	TRINITY_DN1134_c0_g4_i2	90,24
TRINITY_DN11838_c2_g3_i1	3157	TRINITY_DN1131_c6_g1_i1	99,96
TRINITY_DN26054_c0_g1_i1	468	TRINITY_DN11264_c0_g2_i1	100
TRINITY_DN12005_c4_g1_i1	3934	TRINITY_DN11250_c0_g1_i1	100
TRINITY_DN12005_c4_g4_i1	3963	TRINITY_DN11250_c0_g1_i1	100

TRINITY_DN1222_c0_g1_i1	636	TRINITY_DN1093_c0_g5_i1	99,51
TRINITY_DN3941_c0_g2_i1	4515	TRINITY_DN1080_c0_g2_i1	100
TRINITY_DN6260_c0_g2_i1	649	TRINITY_DN10_c0_g1_i1	99,85
TRINITY_DN5860_c0_g1_i1	252	No Blast Hit	
TRINITY_DN13289_c0_g1_i1	323	No Blast Hit	
TRINITY_DN3547_c0_g1_i1	1208	No Blast Hit	
TRINITY_DN15024_c0_g1_i1	244	No Blast Hit	
TRINITY_DN16791_c0_g1_i1	437	No Blast Hit	
TRINITY_DN10137_c1_g1_i1	496	No Blast Hit	

APÊNDICE C - Mapeamento dos PAMs potenciais de *O. chelifer* comparadas as leituras de proteínas convertidas do transcriptoma de *O. monticola* usando PALADIN (WESTBROOK et al., 2017).

ID Proteína	Glândula de <i>O. chelifer</i>	Corpo <i>O. chelifer</i>	<i>O. monticola</i>
TRINITY_DN307_c0_g1_i1.p1	6306	22057	2223
TRINITY_DN28762_c0_g1_i1.p2	3397	36569	467
TRINITY_DN15852_c0_g1_i1.p2	3183	28602	5
TRINITY_DN18722_c0_g1_i1.p2	2950	7380	586
TRINITY_DN4362_c0_g1_i1.p2	2886	8890	215
TRINITY_DN12961_c0_g1_i1.p1	2765	33817	61
TRINITY_DN6260_c0_g2_i1.p1	2117	505	7
TRINITY_DN12925_c0_g1_i1.p2	1788	542	27
TRINITY_DN2382_c0_g1_i1.p2	1757	1064	229
TRINITY_DN10891_c1_g1_i1.p3	1709	93	125
TRINITY_DN10891_c1_g1_i1.p4	1652	109	135
TRINITY_DN8792_c0_g1_i1.p3	1616	1491	2023
TRINITY_DN3547_c0_g1_i1.p1	1529	0	66415
TRINITY_DN272_c0_g1_i1.p1	1114	1124	117
TRINITY_DN10404_c0_g1_i1.p2	1072	142	13
TRINITY_DN11980_c3_g10_i1.1	1060	107	210
TRINITY_DN11980_c3_g2_i1.p1	1022	110	219
TRINITY_DN10404_c0_g1_i9.p2	1021	131	18
TRINITY_DN2760_c0_g2_i1.p2	809	1877	96
TRINITY_DN11631_c2_g1_i1.p2	785	414	0
TRINITY_DN10507_c0_g1_i1.p1	720	13554	0
TRINITY_DN9862_c0_g1_i1.p1	710	1858	0
TRINITY_DN450_c0_g1_i1.p1	349	2257	0
TRINITY_DN12031_c2_g3_i1.p5	344	187	321
TRINITY_DN12031_c2_g3_i2.p5	340	193	300
TRINITY_DN22786_c0_g1_i1.p1	328	415	73
TRINITY_DN4750_c0_g1_i1.p4	290	309	318

TRINITY_DN24839_c0_g1_i1.p1	289	3401	194
TRINITY_DN12005_c2_g1_i1.p4	286	89	58
TRINITY_DN3796_c0_g1_i1.p1	278	2461	22
TRINITY_DN11844_c0_g3_i5.p1	273	1056	0
TRINITY_DN11674_c8_g1_i8.p6	244	1004	635
TRINITY_DN11844_c0_g3_i3.p1	240	1069	0
TRINITY_DN11674_c8_g1_i3.p6	235	1021	637
TRINITY_DN24006_c0_g1_i1.p1	235	969	119
TRINITY_DN11674_c8_g1_i1.p7	226	1028	665
TRINITY_DN11674_c8_g1_i6.p7	217	1037	636
TRINITY_DN25847_c0_g1_i1.p1	205	590	96
TRINITY_DN10663_c0_g1_i1.p1	184	436	167
TRINITY_DN9306_c0_g3_i1.p2	166	354	24
TRINITY_DN11940_c4_g5_i1.p5	158	129	267
TRINITY_DN8567_c0_g1_i1.p1	132	126	296
TRINITY_DN11708_c0_g1_i3.p5	129	283	77
TRINITY_DN11886_c2_g8_i1.p8	128	33	12
TRINITY_DN12012_c2_g6_i1.p4	102	88	20
TRINITY_DN11680_c10_g5_i1.5	100	74	0
TRINITY_DN11969_c3_g2_i2.p0	82	0	0
TRINITY_DN11013_c2_g1_i3.p3	81	62	3
TRINITY_DN11838_c2_g3_i1.p3	75	96	103
TRINITY_DN11013_c2_g2_i1.p1	73	81	9
TRINITY_DN11969_c3_g2_i1.p1	72	0	0
TRINITY_DN11980_c3_g1_i2.p1	72	344	191
TRINITY_DN11013_c2_g1_i1.p3	71	61	6
TRINITY_DN9325_c1_g1_i1.p2	70	60	76
TRINITY_DN1241_c0_g2_i1.p2	67	157	0
TRINITY_DN11980_c3_g1_i1.p1	59	291	178
TRINITY_DN1241_c0_g1_i1.p2	58	155	0
TRINITY_DN9582_c1_g1_i1.p2	51	18	6

TRINITY_DN28995_c0_g1_i1.p1	48	22	78
TRINITY_DN16791_c0_g1_i1.p1	44	4	0
TRINITY_DN27555_c0_g1_i1.p1	44	27	6
TRINITY_DN10137_c1_g1_i1.p5	39	0	14
TRINITY_DN20021_c0_g1_i1.p2	39	203	19
TRINITY_DN11122_c0_g2_i1.p4	37	22	28
TRINITY_DN3941_c0_g2_i1.p8	37	18	19
TRINITY_DN4260_c0_g2_i1.p3	36	152	0
TRINITY_DN14668_c0_g1_i1.p1	33	68	2
TRINITY_DN27940_c0_g1_i1.p1	33	182	26
TRINITY_DN10676_c0_g2_i1.p1	32	61	32
TRINITY_DN27940_c0_g2_i1.p1	31	170	23
TRINITY_DN11610_c0_g2_i1.p3	30	26	0
TRINITY_DN12005_c4_g4_i1.p5	27	25	8
TRINITY_DN11964_c1_g7_i1.p3	25	10	5
TRINITY_DN1980_c0_g1_i1.p1	25	58	127
TRINITY_DN23188_c0_g1_i1.p1	25	19	0
TRINITY_DN4260_c0_g1_i1.p3	25	161	0
TRINITY_DN1958_c0_g1_i1.p2	24	0	0
TRINITY_DN12094_c0_g1_i8.p3	23	7	0
TRINITY_DN12005_c4_g1_i1.p5	21	22	5
TRINITY_DN17243_c0_g1_i1.p1	21	97	18
TRINITY_DN26054_c0_g1_i1.p1	21	14	0
TRINITY_DN10866_c0_g4_i1.p9	20	28	0
TRINITY_DN17016_c0_g1_i1.p2	20	21	10
TRINITY_DN11861_c3_g2_i1.p5	19	2	0
TRINITY_DN13987_c0_g1_i1.p2	18	10	313
TRINITY_DN21289_c0_g1_i1.p1	18	49	58
TRINITY_DN5545_c0_g2_i1.p2	18	26	98
TRINITY_DN15646_c0_g1_i1.p3	16	16	127
TRINITY_DN13204_c0_g1_i1.p1	15	8	0

TRINITY_DN4617_c0_g1_i1.p1	15	3	12
TRINITY_DN11280_c0_g1_i11.7	14	46	0
TRINITY_DN11280_c0_g1_i5.p7	14	45	0
TRINITY_DN1600_c0_g1_i1.p1	14	8	4
TRINITY_DN2395_c0_g1_i1.p1	14	125	41
TRINITY_DN11280_c0_g1_i6.p7	13	41	0
TRINITY_DN11919_c4_g8_i3.p6	13	5	3
TRINITY_DN23671_c0_g1_i1.p2	13	7	0
TRINITY_DN23731_c0_g1_i1.p1	13	81	0
TRINITY_DN12023_c1_g5_i1.p2	12	76	0
TRINITY_DN19441_c0_g1_i1.p1	12	7	0
TRINITY_DN1222_c0_g1_i1.p3	11	4	0
TRINITY_DN17655_c0_g1_i1.p1	11	80	24
TRINITY_DN21408_c0_g1_i1.p1	11	10	9
TRINITY_DN13371_c0_g1_i1.p1	10	10	54
TRINITY_DN18603_c0_g1_i1.p1	10	60	24
TRINITY_DN5860_c0_g1_i1.p2	10	4	8
TRINITY_DN1481_c0_g1_i1.p6	9	5	0
TRINITY_DN26693_c0_g1_i1.p3	9	4	0
TRINITY_DN12094_c0_g1_i18.4	8	6	0
TRINITY_DN13289_c0_g1_i1.p2	8	0	0
TRINITY_DN24344_c0_g1_i1.p2	8	13	9
TRINITY_DN27221_c0_g1_i1.p2	8	0	0
TRINITY_DN12094_c0_g1_i4.p2	7	6	0
TRINITY_DN12094_c0_g1_i7.p4	7	9	0
TRINITY_DN15024_c0_g1_i1.p3	7	1	4
TRINITY_DN6307_c0_g1_i1.p1	7	6	0
TRINITY_DN11193_c0_g4_i4.p2	6	1	0
TRINITY_DN12066_c4_g1_i1.p1	6	0	0
TRINITY_DN11193_c0_g4_i1.p2	5	3	0
TRINITY_DN12094_c0_g1_i5.p3	5	10	0

TRINITY_DN1540_c0_g2_i1.p1	5	35	25
TRINITY_DN16370_c0_g1_i1.p1	5	13	0
TRINITY_DN168_c0_g1_i1.p2	5	2	0
TRINITY_DN26077_c0_g1_i1.p1	5	7	5
TRINITY_DN11793_c4_g2_i1.p4	4	2	2
TRINITY_DN29460_c0_g1_i1.p2	4	25	44
TRINITY_DN22691_c0_g1_i1.p1	3	4	1
TRINITY_DN24173_c0_g1_i1.p2	3	9	8
TRINITY_DN7631_c0_g1_i1.p2	3	43	9
TRINITY_DN11793_c4_g1_i1.p4	2	1	0
TRINITY_DN6759_c0_g1_i1.p7	2	5	0
TRINITY_DN13404_c0_g1_i1.p2	1	1	0
TRINITY_DN6759_c0_g2_i1.p7	0	5	0
