



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
DEPARTAMENTO DE BIOCÊNCIAS
CURSO DE BACHAREL EM BIOTECNOLOGIA

Lara Vitória da Costa Bezerra

**Prospecção *in silico* e caracterização físico-química de peptídeos antimicrobianos
(AMPs) de *Linum usitatissimum* L., *Oryza sativa* L. e *Triticum aestivum* L.**

MOSSORÓ

2024

Lara Vitória da Costa Bezerra

Prospecção *in silico* e caracterização físico-química de peptídeos antimicrobianos (AMPs) de *Linum usitatissimum* L., *Oryza sativa* L. e *Triticum aestivum* L.

Monografia apresentada a Universidade Federal Rural do Semi-Árido como requisito para obtenção do título de Bacharel em Biotecnologia.

Orientadora: Profa. Dra. Isadora Louise Alves da Costa Ribeiro Quintans.

MOSSORÓ

2024

© Todos os direitos estão reservados a Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996 e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tomar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata. A mesma poderá servir de base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) sejam devidamente citados e mencionados os seus créditos bibliográficos.

B574p Bezerra, Lara Vitória da Costa.
Prospecção in silico e caracterização físico-química de peptídeos antimicrobianos (AMPs) de *Linum usitatissimum* L., *Oryza sativa* L. e *Triticum aestivum* L. / Lara Vitória da Costa Bezerra. - 2024.
58 f. : il.

Orientadora: Isadora Louise Alves da Costa Ribeiro Quintans .
Monografia (graduação) - Universidade Federal Rural do Semi-árido, Curso de Biotecnologia, 2024.

1. AMPs. 2. Bioinformática. 3. Plantas. I. Quintans , Isadora Louise Alves da Costa Ribeiro , orient. II. Título.

Ficha catalográfica elaborada por sistema gerador automático em conformidade com AACR2 e os dados fornecidos pelo(a) autor(a).
Biblioteca Campus Mossoró / Setor de Informação e Referência
Bibliotecária: Keina Cristina Santos Sousa e Silva
CRB: 15/120

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

Lara Vitória da Costa Bezerra

**Prospecção *in silico* e caracterização físico-química de peptídeos antimicrobianos
(AMPs) de *Linum usitatissimum* L., *Oryza sativa* L. e *Triticum aestivum* L.**

Monografia apresentada a Universidade
Federal Rural do Semi-Árido como requisito
para obtenção do título de Bacharel em
Biotecnologia.

Defendida em: 17 / 04 / 2024.

BANCA EXAMINADORA

Documento assinado digitalmente
 ISADORA LOUISE ALVES DA COSTA RIBEIRO QU
Data: 25/04/2024 09:17:43-0300
Verifique em <https://validar.itl.gov.br>

Profa. Dra. Isadora Louise Alves da Costa Ribeiro Quintans (UFERSA)
Presidente

Documento assinado digitalmente
 AMANDA GONDIM DE OLIVEIRA
Data: 25/04/2024 11:43:46-0300
Verifique em <https://validar.itl.gov.br>

Profa. Dra. Amanda Gondim de Oliveira (UFERSA)
Membro Examinador

Documento assinado digitalmente
 JAMES LUCAS DA COSTA LIMA
Data: 25/04/2024 10:58:04-0300
Verifique em <https://validar.itl.gov.br>

Prof. Dr. James Lucas da Costa Lima (UFERSA)
Membro Examinador

Dedico este trabalho aos meus pais, Maria Leonora e João Batista; minha irmã, Lorena e meu amigo Daniel (In Memoriam).

AGRADECIMENTOS

A minha família pelo apoio de sempre, em especial minha irmã e mãe, que sempre me incentivaram e apoiaram no decorrer do curso.

À minha Orientadora, Profa. Dra. Isadora Louise, expresso minha profunda gratidão por todas as oportunidades valiosas que me ofereceu e por seu constante apoio, que foi fundamental para o desenvolvimento do trabalho e meu desempenho na graduação.

Aos membros da banca, Profa. Dra. Amanda e Prof. Dr. James, pelas contribuições no trabalho, que foram fundamentais para o desenvolvimento do trabalho.

Aos professores que tive ao longo da graduação que de alguma forma me incentivaram.

Aos meus amigos por sempre estarem presentes no decorrer da graduação, especialmente Vanessa, Thyanne e Atarissis, cujo apoio e companhia foram fundamentais para superar os desafios e celebrar as conquistas ao longo dessa jornada.

Aos membros do Laboratório de Engenharia Genética, Genômica e Proteômica (LEGPRO), em especial Gabriel Macedo e Pablo Leandro.

À UFERSA pela bolsa PICI.

ÉPIGRAFE

“Só se pode alcançar um grande êxito quando nos mantemos fiéis a nós mesmos.”
Friedrich Nietzsche

“Vocês não podem esperar construir um mundo sem melhorar as pessoas.
Cada um de nós deve trabalhar na sua própria melhora.”
Marie Curie

RESUMO

Os peptídeos antimicrobianos (AMPs) estão presentes em plantas, animais e microrganismos e são importantes na defesa contra patógenos (por ex. atividade antifúngica e antibacteriana). Além disso, podem participar da regulação e do crescimento vegetal, adaptação a condições ambientais, comunicação celular, entre outros. As espécies vegetais *Linum usitatissimum* L. (linho), *Oryza sativa* L. (arroz) e *Triticum aestivum* L. (trigo) apresentam grande potencial econômico, agrônomo e medicinal, porém, estudos referentes à presença de AMPs nestas plantas ainda é limitado. Levando esse fator em consideração, essa pesquisa teve por objetivo identificar, caracterizar e realizar a predição de estruturas tridimensionais de AMPs do arroz, trigo e linho, com a utilização de ferramentas de bioinformática. Com isso, o objetivo deste trabalho foi obter um perfil para cada classe de AMPs presentes nas espécies estudadas. Para atingir estes objetivos, inicialmente foi realizada uma busca nos bancos de dados PlantPepDB e Phytozome por sequências de AMPs de *Linum usitatissimum*, *Oryza sativa* e *Triticum aestivum*. Esta pesquisa resultou em 50, 18 e 19 sequências de AMPs, de linho, trigo e arroz, respectivamente, totalizando 87 sequências. Posteriormente, os dados foram submetidos a análise das propriedades físico-químicas com o biopython e as sequências foram verificadas quanto a presença de domínios conservados com o CDD e HMMER. Os resultados indicaram que o LTP foi a classe de AMPs mais abundante entre todas as espécies. Essas sequências foram então modeladas no ColabFold e seus aminoácidos foram analisados através das ferramentas WebLogo e M-Coffee. Este conhecimento possibilitará uma compreensão maior referente a diversidade dos peptídeos antimicrobianos presentes no arroz, trigo e linho.

Palavras-chave: AMPs; bioinformática; plantas.

ABSTRACT

Antimicrobial peptides (AMPs) are found in plants, animals, and microorganisms and play a crucial role in defense against pathogens, exhibiting antimicrobial activities such as antifungal and antibacterial properties. Additionally, they may participate in plant growth regulation, adaptation to environmental conditions, cellular communication, among other functions. Species such as *Linum usitatissimum* L. (flax), *Oryza sativa* L. (rice), and *Triticum aestivum* L. (wheat) hold significant economic, agronomic, and medicinal potential. However, studies regarding the presence of AMPs in these plants remain limited. Taking this into consideration, this research aimed to identify, characterize, and predict the three-dimensional structures of AMPs from rice, wheat, and flax using bioinformatics tools. The objective was to establish a profile for each class of AMPs present in the studied species. To achieve these objectives, an initial search was conducted in the PlantPepDB and Phytozome databases for AMP sequences from *Linum usitatissimum*, *Oryza sativa*, and *Triticum aestivum*. This search yielded 50, 18, and 19 AMP sequences from flax, wheat, and rice, respectively, totaling 87 sequences. Subsequently, the data underwent analysis of physicochemical properties using Biopython, and the sequences were screened for the presence of conserved domains using CDD and HMMER. The results indicated that lipid transfer protein (LTP) was the most abundant class of AMPs among all species. These sequences were then modeled using ColabFold, and their amino acids were analyzed using WebLogo and M-Coffee tools. This knowledge will provide a better understanding of the diversity of antimicrobial peptides present in rice, wheat, and flax.

Keywords: AMPs; bioinformatics; plants.

LISTA DE FIGURAS

Figura 1 - Classificação dos AMPs.....	18
Figura 2 - Fluxograma do processo de análise de sequências de AMPs.....	31
Figura 3 - Representação esquemática dos resíduos de cisteína dos AMPs de snakina (Traes_5DL_BEB7ADA1D.1 - trigo), heveína (PPepDB_2144 - trigo), defensina (Lus10019226 - linho), tionina (PPepDB_3504 - trigo), LTP (Lus10000082 - linho), knottin (LOC_Os02g07440.1 - arroz) e α -hairpinina (PPepDB_2068 - trigo).....	34
Figura 3 - Defensina do linho (Lus10019227) no ChimeraX.....	39
Figura 4 - Alinhamento par-a-par de sequências da α -hairpinina através do M-Coffee. O (*) indica a presença de resíduos altamente conservados e (:) são os resíduos semelhantes.....	46
Figura 5A - Visualização dos aminoácidos conservados de LTP (I), defensina (II) e tionina (III) com o WebLogo.....	46
Figura 5B - Visualização dos aminoácidos conservados de snakina (I) e knottin (II) com o WebLogo.....	47
Figura 6 - AMPs das espécies vegetais com maior pLDDT.....	48
Figura 7 - Representação das classes de AMPs no ChimeraX.....	49

LISTA DE GRÁFICOS

Gráfico 1a - Datos estadísticos de <i>L. usitatissimum</i>	36
Gráfico 1b - Datos estadísticos de <i>T. aestivum</i>	39
Gráfico 1c - Datos estadísticos de <i>O. sativa</i>	41

LISTA DE TABELAS

Tabela 1a - Características gerais dos AMPs: aminoácidos, peso molecular e carga.....	21
Tabela 1b - Características gerais dos AMPs: estrutura tridimensional e resíduos cisteína.....	22
Tabela 2 - Domínios conservados das sequências no CDD e HMMER.....	31
Tabela 3 - Total de AMPs de <i>L. usitatissimum</i> , <i>T. aestivum</i> e <i>O. sativa</i>	32
Tabela 4 - Caracterização físico-química de <i>Linum usitatissimum</i>	35
Tabela 5 - Caracterização físico-química de <i>Triticum aestivum</i>	40
Tabela 6 - Caracterização físico-química de <i>Oryza sativa</i>	43

LISTA DE ABREVIATURAS E SIGLAS

ABA	Ácido abscísico (do inglês, <i>Abscisic acid</i>)
AMP	Peptídeo antimicrobiano (do inglês, <i>Antimicrobial peptide</i>)
BLAST	Basic Local Alignment Search Tool
CCK	Nó de cistina cíclico (do inglês, <i>Cyclic cystine knot</i>)
CDD	Banco de dados de domínio conservado (do inglês, <i>Conserved Domain Database</i>)
CRP	Resíduos ricos em cisteína (do inglês, <i>Cysteine-rich residues</i>)
GASA	Genes Reguladores Estimulados por Ácido Giberélico (do inglês, <i>Gibberellic Acid-Stimulated Regulatory Genes</i>)
GO	Gene Ontology
HMM	Modelo Oculto de Markov (do inglês, <i>Hidden Markov Model</i>)
LC-MS/MS	Cromatografia Líquida Acoplada à Espectrometria de Massas (do inglês, <i>Liquid chromatography–mass spectrometry</i>)
LTP	Proteína de Transferência de Lipídios (do inglês, <i>Lipid Transfer Protein</i>)
PDB	Protein Data Bank
pI	Ponto isoelétrico (do inglês, <i>Isoelectric point</i>)
ROS	Espécies Reativas de Oxigênio (do inglês, <i>Reactive oxygen species</i>)

LISTA DE SÍMBOLOS

@	Arroba
©	Copyright
®	Marca registrada
%	Porcentagem
\$	Cifrão

SUMÁRIO

1. INTRODUÇÃO.....	16
2. FUNDAMENTAÇÃO TEÓRICA.....	18
2.1. AMPs vegetais.....	18
2.2. Linho (<i>Linum usitatissimum</i> L.).....	21
2.3. Arroz (<i>Oryza sativa</i> L.).....	22
2.4. Trigo (<i>Triticum aestivum</i> L.).....	22
2.5. Ferramentas de bioinformática.....	22
2.5.1. PlantAMP.....	24
3. HIPÓTESES.....	25
4. JUSTIFICATIVA.....	25
5. OBJETIVOS.....	26
5.1. OBJETIVO GERAL.....	26
5.2. OBJETIVOS ESPECÍFICOS.....	26
6. MATERIAIS E MÉTODOS.....	26
6.1. Identificação dos AMPs vegetais.....	26
6.2. Prospecção <i>in silico</i> e caracterização físico-química.....	27
6.3. Modelagem <i>in silico</i>	27
7. RESULTADOS E DISCUSSÃO.....	27
8. CONSIDERAÇÕES FINAIS.....	41
REFERÊNCIAS.....	44

1. INTRODUÇÃO

Peptídeos antimicrobianos (AMPs) são moléculas pequenas de baixa massa molecular (10 kDa), essenciais na defesa contra patógenos (Slavokhotova; Rogozhin, 2020). Estão presentes em diversos organismos tais como plantas, microrganismos e animais (Slavokhotova; Rogozhin, 2020).

A expressão dos AMPs ocorre a partir de genes específicos, que são regulados através de mecanismos constitutivos ou induzidos através de fatores externos específicos (Lei *et al.*, 2019). Dentre os mecanismos de ação dos AMPs pode-se citar a destruição das membranas celulares, interrupção da atividade enzimática, interferência da divisão celular e inibição da síntese da parede celular (Ramazi *et al.*, 2022).

Os AMPs vegetais apresentam determinadas características como carga positiva, hidrofobicidade, e são ricos em motivos de cisteína que são cruciais para a estabilidade da estrutura tridimensional dessas moléculas e reconhecimento de alvos, tais como as membranas de patógenos (Santos-Silva *et al.*, 2020). O estudo dessas características essenciais dos AMPs vegetais é crucial para a compreensão de seu mecanismo de ação e, consequentemente, atividade antiviral, antibacteriana, antifúngica, além da regulação e crescimento vegetal (Santos-Silva *et al.*, 2020; Bakare, 2022).

Os AMPs vegetais evoluíram de maneira distinta em comparação com outros tipos de peptídeos resultando em uma maior quantidade de pontes dissulfeto devido à presença de resíduos ricos em cisteína (CRPs), contribuindo para uma estrutura mais compacta e resistente à degradação térmica, química e enzimática (Tam *et al.*, 2015). Os AMPs vegetais são classificados em oito famílias (defensina, tionina, proteína de transferência de lipídios, peptídeos semelhantes a heveína, peptídeos do tipo “knottin”, α -hairpinina, snakina e ciclotídeo), de acordo com suas características estruturais e sequências (Li *et al.*, 2021). Entretanto, as sequências de nucleotídeos entre as diferentes classes de AMPs vegetais são bastante variáveis, dificultando sua predição *in silico* utilizando métodos convencionais, tais como a ferramenta BLAST (do inglês “Basic Local Alignment Search Tool”) (Santos-Silva *et al.*, 2020). Com o objetivo de reduzir essas limitações decorrentes do BLAST, existem ferramentas que podem ser utilizadas, como o CDD (do inglês “Conserved Domain Database”) ou HMM (do inglês “Hidden Markov Model”) (Santos-Silva *et al.*, 2020).

Sequências de AMPs estão disponíveis em bancos de dados como o UniProtKB (<https://www.uniprot.org/>) (Boutet *et al.*, 2007), CAMPR4 (<http://www.bicnirrh.res.in/antimicrobial/>) (Gawde *et al.*, 2023) e PlantPepDB

(<http://14.139.61.8/PlantPepDB/>) (Das *et al.*, 2020), sendo este último específico para AMPs de plantas. Esses bancos de dados fornecem informações taxonômicas, microbiológicas e físico-químicas sobre quase todos os AMPs (Torrent *et al.*, 2012; Quintans *et al.*, 2022). O número de sequências de AMPs descritos até agora ainda é pequeno, especialmente em plantas (Waghu *et al.*, 2016). Os peptídeos antimicrobianos vegetais representam apenas 16% das sequências de AMPs disponíveis no Uniprot, onde os LTPs são a classe de AMPs mais abundante (Clara Pestana-Calsa *et al.*, 2010).

A comunidade científica tem trabalhado para desenvolver estratégias para identificar novos AMPs de plantas e melhor entender suas estruturas e funções, utilizando análises *in silico* e biologia molecular (Pestana-Calsa & Calsa JR, 2011; Quintans *et al.*, 2022). A abordagem *in silico* consiste na utilização de métodos computacionais para auxiliar no desenvolvimento e avanço das pesquisas laboratoriais (Ekins; Mestres; Testa, 2007). Neste contexto, Quintans (2019) desenvolveu um pipeline com o objetivo de realizar a busca por peptídeos antimicrobianos (AMPs) em bancos de dados públicos de sequências de DNA, resultando em 16.870 AMPs de 1.003 espécies (Quintans, 2019). Entretanto estas sequências ainda não foram experimentalmente validadas. Sendo assim, a caracterização dessas sequências utilizando ferramentas de bioinformática irá contribuir para a expansão do conhecimento sobre AMPs vegetais, tornando-se necessária a prospecção *in silico* e caracterização físico-química destas sequências.

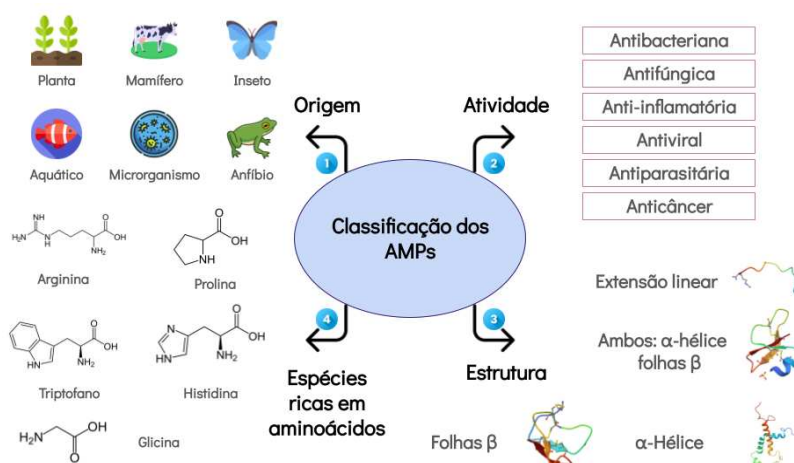
2. FUNDAMENTAÇÃO TEÓRICA

2.1. AMPs vegetais

As plantas são organismos sésseis que desenvolveram estratégias de sobrevivência para se defender contra patógenos e pragas, através de barreiras, produção de metabólitos secundários e compostos antimicrobianos (Adams, 2005; Li *et al.*, 2020). As barreiras mecânicas ou físicas de proteção são constituídas pelo desenvolvimento de espinhos, cerdas e tricomas e formação de uma cutícula cerosa que protege a planta dos insetos (War *et al.*, 2012). Já a barreira química é constituída por metabólitos secundários que atuam na defesa da planta contra diferentes estresses (War *et al.*, 2012). Por exemplo, o arroz tem como metabólitos secundários os ácidos fenólicos, flavonóides, terpenóides, esteróides e alcalóides, que possuem propriedades antimicrobianas, antioxidantes, citotóxicas, entre outras (Wang *et al.*, 2018).

Assim como os metabólitos secundários, o peptídeos antimicrobianos vegetais são essenciais para o sistema de defesa das plantas (Figura 1), pois fornecem resistência contra muitos patógenos (Goyal, Mattoo, 2014), bem como atuam em resposta a diferentes estresses ambientais abióticos, tais como seca, salinidade (Olga *et al.*, 2020). Os AMPs atuam no sistema imune inato através da atividade antimicrobiana de amplo espectro (Tantong *et al.*, 2016).

Figura 1 - Classificação dos AMPs



Fonte: Adaptado de Huan *et al.*, (2020)

A purotionina é pertencente a família tionina dos AMPs, sendo esta classificada somente após a descoberta da α -purotionina, como α -tionina, β -tionina e γ 1/ γ 1-tionina (Tam

et al., 2015). O primeiro AMP vegetal, intitulado purotionina, foi isolado e obtido a partir do endosperma do trigo (*Triticum aestivum* L.) (De Caley 1972). Para as plantas, os AMPs são de extrema importância porque as mesmas não apresentam imunidade adaptativa e os AMPs desempenham a função de defesa contra infecções bacterianas, fúngicas, entre outras (Kumar; Kizhakkedathu; Straus, 2018).

As famílias de AMPs vegetais são classificadas com base na similaridade entre suas sequências e estruturas tridimensionais em: proteínas de transferência de lipídeos (LTPs), defensinas, tioninas, peptídeos do tipo heveína, peptídeos do tipo “knottins”, α -hairpinina, snakinas e ciclotídeos (Lima *et al.*, 2022).

As proteínas de transferência de lipídios (LTPs) são os maiores peptídeos antimicrobianos vegetais, variando de 70 a 90 aminoácidos (Salminen; Blomqvist; Edqvist, 2016). Ainda assim, são consideradas moléculas pequenas, visto que, são considerados AMPs peptídeos que tenham um número inferior a 100 aminoácidos. Os LTPs, assim como todos os AMPs vegetais, são ricos em resíduos de cisteína (Cys) e restritos às plantas, apresentando em estrutura principal 8 cisteínas estabilizadas por quatro pontes dissulfeto conservadas (Salminen; Blomqvist; Edqvist, 2016). Os LTPs se destacam principalmente pelo seu desempenho em respostas a estresses bióticos e abióticos, tradução de sinal e crescimento e desenvolvimento das plantas (Liu *et al.*, 2015). Inicialmente, os peptídeos de transferência de lipídios (LTPs) foram classificados em LTP1 e LTP2, com base no peso molecular de 9 e 7 kDa, respectivamente (Santos-Silva *et al.*, 2020). Mas foi proposto um novo sistema de classificação com 5 tipos principais e 5 tipos adicionais, sendo LTP1, LTP2, LTPc, LTPd e LTPg e LTPe, LTPf, LTPh, LTPj e LTPk, respectivamente (Edstam *et al.*, 2011).

As defensinas vegetais apresentam cerca de 45-54 resíduos de aminoácidos e estão presentes nas sementes, folhas, raízes, cascas, tubérculos, vagens, frutas e órgãos florais (Parisi *et al.*, 2019). Suas sequências de aminoácidos variam muito, com exceção aos resíduos de cisteína conservados (De Oliveira Carvalho; Moreira, 2011). A localização das defensinas nas plantas evidencia a função de proteção contra patógenos, pois são expressos na parede celular e camadas externas de sementes, raízes, xilema, entre outros (De Oliveira Carvalho; Moreira, 2011; Terras *et al.*, 1995).

As tioninas apresentam peso molecular de até 5 kDa, têm três ou quatro pontes dissulfeto e são agrupadas em 5 grupos de acordo com o tamanho, número de resíduos e pontes dissulfeto (Azmi; Hussain, 2021). São encontrados apenas em plantas e têm como função a defesa contra patógenos, mas também atuam na maturação das sementes e

apresentam atividade citotóxica e anticancerígena (Guzmán-Rodríguez *et al.*, 2015). Estes peptídeos se assemelham muito estruturalmente às defensinas (Pelegriani; Franco, 2005).

Os peptídeos semelhantes a heveína receberam esse nome graças à seringueira (*Hevea brasiliensis*), no qual foi isolado inicialmente esse AMP (Archer, 1960). Essa família apresenta cerca de 29 - 45 resíduos de aminoácidos ricos em cisteína e glicina e são capazes de se ligar à quitina (Slavokhotova *et al.*, 2017). Apresentam atividade antifúngica, sendo capazes de inibir o crescimento e desenvolvimento de fungos que contêm quitina, atividade antibacteriana e desempenham um papel relevante no sistema de defesa do trigo (Porto *et al.*, 2012).

Os peptídeos do tipo “knottin” apresentam cerca de 30 aminoácidos, com 6 resíduos de cisteína que formam um nó de cisteína através das ligações dissulfeto conservadas (Tam *et al.*, 2015). Além dos organismos vegetais, esses peptídeos estão presentes em insetos, artrópodes, moluscos e aracnídeos (Molesini *et al.*, 2017). Esses peptídeos são estudados com o principal propósito o desenvolvimento de medicamentos, tendo em vista sua ampla afinidade e seletividade, constituindo particularidades desejáveis para um fármaco (Kintzing; Cochran, 2016).

As α -hairpininas apresentam 4 resíduos de cisteína e estão presentes principalmente nas sementes de espécies vegetais como o trigo-sarraceno (*Fagopyrum esculentum*), com a finalidade de proteção contra insetos (Slavokhotova; Rogozhin, 2020). Apresenta atividade inibidora de tripsina, presente no peptídeo C2 isolado da semente de abóbora (*Cucurbita maxima*), atividade antifúngica, entre outros (Slavokhotova; Rogozhin, 2020).

Snakinas são peptídeos antimicrobianos presentes apenas em plantas, que apresentam 12 resíduos de cisteína e um domínio conservado intitulado GASA (Genes Reguladores Estimulados por Ácido Giberélico). O GASA desempenha um papel fundamental na regulação da germinação das sementes, resposta a estresse biótico e abiótico, desenvolvimento floral e alongamento do caule. Snakin/GASA representa os peptídeos antimicrobianos que apresentam uma maior quantidade de resíduos Cys, dentre as outras famílias (Cheng *et al.*, 2019; Su *et al.*, 2020). Snakina-1 ou SN-1 apresenta atividade antifúngica e antibacteriana e formam seis ligações dissulfeto (Almasia *et al.*, 2017). Snakina-2 ou SN2, expresso em folhas e flores, mata os patógenos através da formação de poros na membrana e faz parte do sistema de defesa de *Solanum lycopersicum* (tomate) (Herbel; Sieber-Frank; Wink, 2017).

Os ciclotídeos são peptídeos circulares que apresentam uma estrutura denominada nó de cistina cíclico (CCK), responsável por garantir estabilidade à molécula (Ireland *et al.*,

2010). Essa estrutura ou motivo CCK é formado a partir de três ligações dissulfeto com seis cisteínas, que evitam a degradação química, térmica e enzimática (Park *et al.*, 2017), podendo ser utilizada como molde para o desenho racional de fármacos (Craik *et al.* 2012). Além do potencial farmacológico, os ciclotídeos apresentam atividade inseticida, anti-helmíntica, antibacteriana e anticancerígena (Park *et al.*, 2017).

A Tabela 1a apresenta as características gerais de AMPs referente a quantidade de aminoácidos, peso molecular e carga para diferentes famílias de AMPs vegetais (LTP, defensina, tionina, heveína, “knottin”, α -hairpinina e snakina).

A carga dos peptídeos, em sua maioria, é positiva (Quintans *et al.*, 2022) e o peso molecular apresenta variações, sendo a classe LTP a maior dentre os outros valores e a α -hairpinina a menor, assim como o comprimento das sequências, que varia de 28-35 (Slavokhotova; Rogozhin, 2020).

Tabela 1a - Características gerais dos AMPs: aminoácidos, peso molecular e carga.

Família	Aminoácidos	Peso molecular (KDa)	Carga	Referências
LTP	70-90	7-9	Positiva	Quintans <i>et al.</i> , 2022
Defensina	45-54	2-5	Positiva	Quintans <i>et al.</i> , 2022
Tionina	45-47	5	Positiva	Quintans <i>et al.</i> , 2022
Heveína	29-45	4.7	Positiva	Quintans <i>et al.</i> , 2022 Slavokhotova <i>et al.</i> , 2017
Knottin	30-50	< 6	Positiva	Attah <i>et al.</i> , 2022 Han <i>et al.</i> , 2023 Quintans <i>et al.</i> , 2022
α -Hairpinina	28-35	2-4	Positiva	Slavokhotova; Rogozhin, 2020 Tam <i>et al.</i> , 2015
Snakina	60	6.9	Positiva	Quintans <i>et al.</i> , 2022

Já a Tabela 1b destaca as características dos AMPs com relação à estrutura tridimensional e a quantidade de resíduos de cisteína para cada família.

É essencial conhecer a conformação estrutural e os resíduos de cisteína presentes nas sequências, pois é através dessas informações que é possível diferenciar as diversas classes de AMPs existentes. Por exemplo, LTP é um AMP que sua estrutura tem como característica a presença de 4 α -hélices (Quintans *et al.*, 2022) e é essencial que a molécula tenha essa conformação.

Tabela 1b - Características gerais dos AMPs: estrutura tridimensional e resíduos cisteína.

Família	Estrutura tridimensional	Resíduos cisteína	Referências
LTP	Conjunto de 4 α -hélices interligadas por laços flexíveis	X{3}-C-X{9}-C-X{12}-CC-X{18} -C-X{1}-C-X{23}-C-X{15}-C-X{4} }	Lay; Anderson, 2005 Quintans <i>et al.</i> , 2022
Defensina	Folha β com 3 fitas e α -hélice alinhada	X{3}-C-X{10}-C-X{5}-C-X{3}-C-X{9}-C-X{8}-C-X{1}-C-X{3}-C	Lay; Anderson, 2005 Quintans <i>et al.</i> , 2022
Tionina	2 α -hélices dispostas de forma antiparalela e 2 fitas β também em conformidade antiparalela	X{2}-CC-X{7}-C-X{3}-C-X{8}-C-X{3}-C-X{1}-C-X{8}-C-X{6}	Lay; Anderson, 2005 Quintans <i>et al.</i> , 2022
Heveína	Folha β composta por três fitas, e uma α -hélice curta conectando a segunda à terceira fita	X{3}-C-X{4}-C-X{4}-CC{5}-C-X{6}-C-X{2}	Lay; Anderson, 2005 Quintans <i>et al.</i> , 2022
Knottin	Folhas β e α -hélice que formam um padrão de nó	X{1}-C-X{6}-C-X{8}-CC-X{3}-C-X{10}-C-X{3}	Quintans <i>et al.</i> , 2022 Tam <i>et al.</i> , 2015
α -Hairpinina	2 α -hélices conectadas por uma região mais flexível, formando uma alça. Hélice - alça - hélice	X{1}-C-X{3}-C-X{1}-C-X{3}-C-X{1}	Slavokhotova; Rogozhin, 2020 Tam <i>et al.</i> , 2015
Snakina	Padrão helicoidal composto por 1 α -hélice curta e 1 α -hélice 310, intercaladas por alças rigidamente estabilizadas	C-X{3}-C-X{3}-C-X{7,11}-C-X{3}-C-X{2}-CC-X{2}-C-X{1,3}-C-X{11}-C-X{1,2}-C-X{11,14}-KCP	Porto; Franco, 2013 Quintans <i>et al.</i> , 2022

Legenda: X (Qualquer aminoácido, exceto cisteína), C (Cisteína)

As famílias de AMPs vegetais desempenham funções cruciais relacionadas à defesa das plantas, podendo destacar as cultivares do trigo, arroz e linho, devido à sua importância agronômica e econômica. A investigação dos AMPs dessas cultivares apresenta grande relevância nos estudos *in silico*, oferecendo uma abordagem eficaz e eficiente para explorar suas propriedades.

2.2. Linho (*Linum usitatissimum* L.)

O linho ou linhaça (*Linum usitatissimum* L.) é uma planta herbácea, dicotiledônea e pertencente à família Linaceae, que é responsável pela produção do óleo de linhaça a partir das sementes, e a fibra através do linho, sendo utilizada tanto na indústria alimentícia, quanto na indústria têxtil (Bemiller *et al.*, 1993). Com relação à composição, a linhaça apresenta 40-45% de ácidos graxos e 20-25% de proteína (Saleem *et al.*, 2020), representando um excelente alimento funcional. A linhaça tem sido bastante utilizada na nutrição de humanos e animais, promovendo inúmeros benefícios à saúde, tais como redução dos níveis de colesterol (Dupasquier *et al.*, 2006), diminuição da pressão arterial (Rodriguez-Leyva *et al.*, 2013), redução nos níveis de glicose em pessoas com pré-diabetes ou diabetes tipo 2 (Soltanian; Janghorbani, 2018), entre outros. Sendo assim, esta planta pode ser considerada uma fonte inexplorada para a prospecção de AMPs.

Além disso, algumas variedades de linho são resistentes a inúmeros patógenos, tais como *Fusarium oxysporum* (Kanapin *et al.*, 2021). *Fusarium* sp. apresenta resistência a grande parte dos antifúngicos, sendo essencial o estudo sobre a relevância dos peptídeos antimicrobianos na redução das infecções fúngicas causadas por esse gênero (Denardi *et al.*, 2021). Com relação à infecções causadas por fungos, *Fusarium culmorum* é um patógeno que causa acúmulo de micotoxinas e podridão do pé e raiz da planta, levando a formação de Espécies Reativas de Oxigênio (ROS, do inglês “Reactive oxygen species”) e inúmeros danos celulares no linho (Boba *et al.*, 2022). Até o momento, não há muitas informações sobre o mecanismo de infecção do *Fusarium culmorum* no linho, mas acredita-se que o ABA (do inglês “Absciscic acid”) esteja envolvido na defesa (Boba *et al.*, 2022). Sendo assim, torna-se necessário explorar os genes que podem estar envolvidos no sistema de defesa desta espécie, e se os AMPs podem ser um tipo de mecanismo de defesa até então pouco estudado.

2.3. Arroz (*Oryza sativa* L.)

O arroz (*Oryza sativa* L.) é uma planta herbácea, monocotiledônea e pertencente à família Poaceae, amplamente cultivado e com um bom valor nutritivo, constituído principalmente por lipídios, minerais, vitaminas, fibras, entre outros (Zhao; Lin; Chen, 2020).

A produção dos grãos do arroz é afetada pela presença de bactérias e fungos, que causam prejuízos em seu cultivo (Tantong *et al.*, 2016). Dentre as espécies que prejudicam o cultivo do arroz, pode-se destacar: *Fusarium* spp., *Rhizoctonia solani* e *Curvularia* spp. que causam danos com as manchas nas mudas no arroz e podridão nas sementes (Verma *et al.*, 2018). Com relação a pragas que atacam as lavouras, no arroz e no trigo, *Rhizopertha dominica* é o inseto que causa um maior dano às sementes (Machado *et al.*, 2006). Em um estudo realizado por Shelenkov e colaboradores (2020), a partir da busca por potenciais AMPs vegetais, foi possível determinar que as famílias Asteraceae, Fabaceae e Poaceae apresentaram uma maior abundância em relação a presença de defensinas (Shelenkov; Slavokhotova; Odintsova, 2020).

2.4. Trigo (*Triticum aestivum* L.)

O trigo (*Triticum aestivum* L.) é uma planta herbácea, monocotiledônea e pertencente à família Poaceae, constituído por ácidos fenólicos, ferro (Fe), zinco (Zn) que são importantes para a saúde humana em decorrência do seu alto valor nutricional (Lephuthing *et al.*, 2017). Além do seu alto poder nutricional, o trigo apresenta AMPs que são responsáveis pela defesa da planta contra diversos patógenos e as primeiras defensinas vegetais foram isoladas a partir do *T. aestivum* (Nawrot *et al.*, 2014).

A podridão radicular comum e cercosporiose são doenças presentes no trigo causadas pelo fungo *Bipolaris sorokiniana* e *Cercospora* spp., respectivamente. Pesquisas mostram que alguns AMPs vegetais são capazes de afetar as membranas celulares dos fungos e inibir o crescimento dos mesmos (Su *et al.*, 2020).

Dessa forma, abordagens computacionais, como a bioinformática, pode contribuir para o avanço sobre os AMPs do arroz, trigo e linho, além de proporcionar uma compreensão mais profunda referente às suas propriedades estruturais.

2.5. Ferramentas de bioinformática

A bioinformática é definida pela utilização de ferramentas computacionais tendo como principal propósito à análise, captura e interpretação de dados biológicos (Bayat, 2002). É possível destacar que, atualmente, a bioinformática é essencial para o desenvolvimento da pesquisa e da biologia moderna, visto que a quantidade de dados existentes necessita de novos métodos de visualização, análise, modelagem e previsão de dados (Rhee; Dickerson; Xu 2006). Nesse contexto, a bioinformática proporciona um vasto conhecimento em relação à identificação de AMPs vegetais, busca por domínios conservados e análise de estruturas tridimensionais de cultivares como *L. usitatissimum*, *T. aestivum* e *O. sativa*. A diversificação das classes de AMPs e composição variada, torna-se necessário a utilização de ferramentas computacionais, por exemplo o HMMER (Eddy, 1992). O HMMER é baseado em Modelos Ocultos de Markov (HMM), um perfil que analisa sequências biológicas a partir de modelos estatísticos e pode ser utilizado para diversas finalidades como: alinhamento de sequências através de um perfil-HMM, identificação de sequências homólogas, buscar domínios conservados em sequências, entre outros (Eddy, 1992). Um domínio conservado pode ser definido como uma região de uma proteína que se mantém relativamente inalterada ao longo da evolução e, por isso, esses domínios estar associados à características funcionais de determinada proteína (Forslund; Pekkari; Sonnhammer, 2011). Se um AMP apresenta o domínio “Gamma-thionin”, ele pode pertencer a classe das defensinas e essas informações podem ser obtidas a partir da ferramenta HMMER juntamente com uma base de dados Pfam-A, obtida no banco de dados InterPro (Mcdowall; Hunter, 2011).

Além disso, os bancos de dados e algoritmos são recursos utilizados com o propósito de obter sequências e estruturas de proteínas através do NCBI (do inglês “National Center for Biotechnology Information”) e PDB (do inglês “Protein Data Bank”), analisar dados biológicos como o Biopython, realizar anotações pelo Gene Ontology (GO), entre outros (Pevzner; Shamir, 2011). O PDB (<https://www.rcsb.org/>) é um banco de dados primário que detalha informações referentes aos dados dos estudos das proteínas envolvendo Cristalografia de raios x e Ressonância Magnética Nuclear (NMR), além de fornecer formas de visualizar e baixar as proteínas em diferentes formatos, o banco de dados apresenta recursos externos de classificação de estrutura proteica, o CATH (<https://www.cathdb.info/>), flexibilidade da estrutura proteica, o PDBflex (<https://pdfflex.org/>), entre outros (Burley *et al.*, 2019). O NCBI é um repositório de banco de dados que integra informações referentes à bancos de dados de DNA e proteínas, além de detalhar dados sobre a taxonomia, estrutura, entre outros

(Bayat, 2002). O NCBI disponibiliza o acesso a uma das ferramentas mais utilizadas da bioinformática, BLAST (Johnson *et al.*, 2008).

O BLAST é uma ferramenta bastante utilizada que baseia-se em algoritmos heurísticos para realizar alinhamentos locais entre sequências (Ranganathan; Nakai; Schonback, 2018). As sequências de busca do BLAST também podem ser comparadas com sequências de estruturas conhecidas no Protein Data Bank (PDB) (Ranganathan; Nakai; Schonback, 2018). Entretanto, a ferramenta mais utilizada para realizar a anotação de sequências de estudos de genômica e transcriptômica, o BLAST, pode falhar em reconhecer as sequências de AMPs vegetais nos dados analisados (Santos-Silva *et al.*, 2020), visto que apresenta uma sensibilidade limitada para identificar sequências que são altamente divergentes ou com baixa similaridade. Com isso, ao realizar o alinhamento local da sequência e as comparações, o BLAST divide a sequência e os AMPs, por serem moléculas pequenas altamente variáveis, podem não apresentar bons alinhamentos (Santos-Silva *et al.*, 2020), dificultando assim a identificação e predição com o BLAST. Outras ferramentas seriam mais indicadas, tais como HMMER, tendo em vista que é um software baseado em HMM, sendo mais úteis na detecção de sequências divergentes de forma mais sensível que o BLAST (Eddy, 1992), levando a um aumento na identificação de AMPs vegetais (Santos-Silva *et al.*, 2020). Uma vez que as sequências são identificadas, é necessário realizar a busca por domínios conservados.

A busca por domínios conservados é realizada com o propósito de confirmar sequências de AMPs preditas. Quintans (2019) realizou uma busca por domínio conservado de AMPs vegetais e para a defensina obteve 56.49% das sequências com CDD (Quintans, 2019). O CDD dispõe de uma compilação de alinhamentos com domínios de proteína conservados e que faz uso do RPS-BLAST no seu mecanismo de busca, para realizar comparações entre sequências e os domínios contidos no banco de dados (Marchler-Bauer *et al.*, 2002). Contribuindo para uma abordagem mais eficiente na análise de dados biológicos, além da busca por domínios conservados, pode ser utilizado o Python (Van Rossum *et al.*, 1995)..

Python é uma linguagem de programação que tornou-se bastante popular e é muito utilizada em estudos referentes à biologia de sistemas, filogenia e biologia estrutural, através da análise de dados e resoluções de questões referentes à biologia, bioinformática, entre outros (Ekmekci; Mcanany; Mura 2016; Van Rossum *et al.*, 1995). A sua popularidade é decorrente da facilidade de manuseio por ser uma linguagem de programação bastante intuitiva e, além disso, o Python tem muitos fóruns que promovem a retirada de possíveis dúvidas e erros presentes no código (Ekmekci; Mcanany; Mura, 2016). Com o Python, os

pesquisadores podem automatizar processos de análise de dados, incluindo a realização de alinhamentos entre sequências através de pacotes do Biopython, tornando o processo mais eficiente e acessível (Cock *et al.*, 2009).

Biopython é uma biblioteca do Python que apresenta inúmeras aplicações na bioinformática, através de seus diferentes pacotes e funções; podendo trabalhar com sequências biológicas, estimar propriedades de DNA e proteínas, entre outros (Bassi, 2007). Além do biopython, existem outras bibliotecas importantes na bioinformática, como o Pandas, NumPy, Pysam e BioPerl (Li *et al.*, 2009; Mckinney, 2012; O’Neil, 2017; Stajich *et al.*, 2002). O Pandas é outra biblioteca Python utilizada para análise de dados, permitindo o carregamento, manipulação, análise e combinação de dados (O’Neil, 2017).

As ferramentas de bioinformática contribuem para que seja possível a realização da predição e caracterização físico-química dos AMPs vegetais, sendo uma alternativa bastante viável para a predição de novas estruturas moleculares. Compreender as propriedades dos AMPs (carga, composição e distribuição de aminoácidos) é importante, pois são através dessas informações que pode-se obter o conhecimento a respeito do modo de interação do AMP com a membrana (Pirtskhalava *et al.*, 2021). Por exemplo, geralmente os AMPs são carregados positivamente, ou seja, catiônicos e essa característica facilita a fixação dos mesmos na membrana dos microrganismos (Chung *et al.*, 2020). Isso ocorre em decorrência das diferenças entre as cargas do peptídeo e da membrana da bactéria (carga negativa, aniônica), causando uma ruptura da membrana e consequente morte celular bacteriana (Li *et al.*, 2017).

Em um estudo realizado por Moyer e colaboradores (2021), foi realizada a identificação de AMPs do tecido aéreo de *Amaranthus tricolor* através de abordagens *in silico* e em bancada (Moyer *et al.*, 2021). Os experimentos de bancada consistiram na proteômica bottom-up LC-MS/MS (Cromatografia Líquida Acoplada à Espectrometria de Massas) e ensaios de bioatividade (Moyer *et al.*, 2021). As classes de AMPs identificadas foram: snakina, LTP e defensinas e o Atr-LTP1 apresentou atividade contra todas as espécies bacterianas testadas (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Enterobacter cloacae*), com exceção a *Pseudomonas aeruginosa* (Moyer *et al.*, 2021).

2.5.1. PlantAMP

O PlantAMP é um banco de dados com uma interface simples e prática, que está em desenvolvimento pelo nosso grupo de pesquisa e é focado em peptídeos antimicrobianos de plantas, o qual dispõe de 1535 sequências das quais 595 já encontram-se preditas. É importante ressaltar que o PlantAMP ainda não foi publicado (Quintans *et al.*, 2023).

É um banco de dados essencial para estudos envolvendo AMPs vegetais, pois garante um aumento no número de sequências depositadas, visto que os peptídeos antimicrobianos de plantas não são muito documentados, além de fornecer um recurso para pesquisa. O PlantAMP é bastante importante para a pesquisa envolvendo peptídeos, pois existem poucos bancos de dados específicos para AMPs de plantas, podendo destacar o PlantPepDB (<http://14.139.61.8/PlantPepDB/index.php>) (Das *et al.*, 2020), porém em alguns momentos apresenta uma certa instabilidade ao acessar.

3. HIPÓTESES

A prospecção *in silico* e caracterização físico-química em sequências do arroz, trigo e linho permitirá uma maior compreensão sobre as diferentes classes de AMPs presentes nestas espécies vegetais, bem como suas funções biológicas.

4. JUSTIFICATIVA

O arroz, trigo e a linhaça são culturas amplamente utilizadas como alimento por grande parte da população, sendo bastante comercializados pela indústria. Essas culturas de cereais apresentam muitos benefícios à saúde, como por exemplo, a linhaça que apresenta redução nos níveis de colesterol (Dupasquier *et al.*, 2006), e o arroz e trigo têm efeitos anti-hipertensivos (Capriotti *et al.*, 2015). Alguns peptídeos obtidos de cereais são capazes de inibir a oxidação de lipídios, e torna-se viável determinar se esses cereais (arroz, trigo e linhaça) também apresentam essa vantagem (López-García *et al.*, 2022).

O arroz e o trigo são pertencentes à mesma família, Poaceae e o linho pertence à família Linaceae. A família Poaceae é constituída por espécies que têm capacidade de resistir a ambientes desfavoráveis e que apresentam mecanismos de defesa que estão relacionados à produção de AMPs, sendo as defensinas as mais estudadas (Odintsova; Slezina; Istomina, 2020). Já a família Linaceae destaca-se pela diversidade ecológica e morfológica (De Santana

Lopes *et al.*, 2018). Estas famílias compreendem espécies com um amplo potencial para pesquisas envolvendo os AMPs vegetais.

Estudos com relação aos peptídeos antimicrobianos do linho, arroz e trigo são pouco explorados pela comunidade científica, apesar de serem culturas amplamente difundidas. Há evidências de que essas plantas são ótimos candidatos, pois são consideradas opções dietéticas saudáveis que, segundo estudos, melhoram os perfis lipídicos no sangue, principalmente o *Linum usitatissimum* L. (Moyer; Brechbill; Hicks, 2021). Além disso, os AMPs do linho, trigo e arroz apresentam atividade antifúngica e antibacteriana que podem auxiliar no cultivo dessas culturas, uma vez que possibilitaria resistência a algumas doenças que afetam essas espécies e, conseqüentemente, haveria menos necessidade do uso de defensivos químicos.

Sendo assim, torna-se necessário a pesquisa referente à prospecção e caracterização *in silico* de AMPs de *Linum usitatissimum* L., *Oryza sativa* L. e *Triticum aestivum* L., visando enriquecer o banco de dados PlantAMP, que está em desenvolvimento pelo nosso grupo de pesquisa e determinar a estrutura tridimensional dessas plantas para posterior utilidade biotecnológica referente à predição de patógenos vegetais que possam ser alvos para o tratamento com peptídeos antimicrobianos. Essa abordagem pode ser realizada através de ferramentas de engenharia genética que permite a superexpressão de AMPs em diferentes espécies vegetais com o objetivo de ter um controle sobre fungos, bactérias e outros patógenos de plantas (Iqbal *et al.*, 2019). Vale destacar que para realizar essa aplicação, seria necessário inicialmente a extração dos AMPs, identificação dos peptídeos, realização da Cromatografia Líquida Acoplada à Espectrometria de Massas e ensaios de bioatividade (Moyer; Brechbill; Hicks, 2021), sendo uma metodologia complexa, tendo em vista a quantidade de equipamentos e infraestrutura adequada para os ensaios em laboratório.

5. OBJETIVOS

5.1. OBJETIVO GERAL

Identificar, caracterizar e realizar a predição de estruturas tridimensionais de AMPs presentes no arroz, trigo e linho, através de ferramentas de bioinformática.

5.2. OBJETIVOS ESPECÍFICOS

- Classificar os AMPs vegetais do arroz, trigo e linho de acordo com a família;
- Realizar a prospecção *in silico* dos AMPs;

- Aprimorar novos conhecimentos com relação às bibliotecas do Python;
- Caracterizar os AMPs do linho, arroz e trigo;
- Realizar a predição de estrutura tridimensional com o ColabFold;
- Enriquecer o banco de dados de peptídeos antimicrobianos de plantas, o PlantAMP;
- Identificar motivos proteicos para cada classe de AMP.

6. MATERIAIS E MÉTODOS

6.1. Identificação dos AMPs vegetais

As sequências de AMPs provenientes de *Linum usitatissimum* L. foram coletadas através do banco de dados Phytozome, e de *Oryza sativa* L. e *Triticum aestivum* L. foram obtidas através do Phytozome e PlantPepDB. Para realizar a limpeza dos dados, excluindo qualquer sequência repetida, foi utilizado a biblioteca Pandas. Posteriormente, as sequências foram classificadas e adicionadas em um arquivo fasta, sendo separadas de acordo com a família de AMP vegetal correspondente - LTPs, defensina, heveína, snakín/GASA, tioninas, “knottins”, α -hairpininas.

6.2. Prospecção *in silico* e caracterização físico-química

A presença de domínios conservados de AMPs vegetais foi realizada através do RPS-BLAST contra o CDD (Wang *et al.*, 2023) do NCBI e o HMMER (Versão 3.4) (Eddy, 1992).

Com o objetivo de excluir os possíveis pontos de clivagem e o peptídeo sinal das sequências, foi utilizado o servidor SignalP 6.0 (<https://services.healthtech.dtu.dk/services/SignalP-6.0/>) (Almagro *et al.*, 2019).

As propriedades físico-químicas dos peptídeos antimicrobianos foram preditas por meio da biblioteca Biopython (Cock *et al.*, 2009) no Google Colab. A biblioteca pandas também foi utilizada com o objetivo de analisar os dados resultantes da caracterização físico-química e 3 gráficos foram gerados no Matplotlib (Hunter, 2007). A estatística descritiva dos dados foi realizada através do método describe() no pandas, destacando a média e desvio padrão. Os resíduos de aminoácidos conservados dentre os peptídeos prospectados foram analisados por meio das ferramentas M-Coffee (versão 11.00) e WebLogo (Crooks *et al.*, 2004).

6.3. Modelagem *in silico*

As seqüências peptídicas foram avaliadas por modelagem *in silico* de suas estruturas tridimensionais por meio do ColabFold (Mirdita *et al.*, 2022), tendo em vista se a estrutura referente apresenta modelagem de acordo com as conformidades da classe de cada AMP. Por fim, os dados dos AMPs serão publicizados na base de dados PlantAMP.

A Figura 2 representa o fluxograma da metodologia realizada no trabalho.

Figura 2 - Fluxograma do processo de análise de seqüências de AMPs.



Fonte: Autoria própria.

7. RESULTADOS E DISCUSSÃO

As seqüências de AMPs obtidas nos bancos de dados PlantPepDB e Phytozome foram salvas em formato fasta. Com o propósito de verificar se as seqüências tinham domínios conservados, foi realizada a busca por domínios conservados através do HMMER e CDD. A busca de seqüências nos bancos de dados resultou em um total de 127 seqüências, sendo 61 do linho (*Linum usitatissimum*), 18 do trigo (*Triticum aestivum*) e 48 de arroz (*Oryza sativa*) (Tabela 2). Das 127 seqüências, 87 apresentaram domínios conservados. Foram identificadas um total de 9 domínios conservados em 9 seqüências distintas do linho, enquanto 16 domínios conservados foram encontrados em 18 seqüências únicas do trigo, utilizando o programa Batch CD-Search (Marchler-Bauer *et al.*, 2002). As seqüências de arroz não apresentaram domínios conservados identificáveis a partir do CDD (Marchler-Bauer *et al.*, 2002). A

realização da busca por domínios conservados utilizando o Pfam através do HMMER (v. 3.4) resultou em um maior número de domínios conservados em todas as espécies, sendo 50 para o linho, 18 trigo e 19 arroz (Tabela 2). A partir dos resultados referentes ao domínio conservado, foi possível salvar as sequências separadas, no formato fasta, de acordo com o domínio.

Tabela 2 - Domínios conservados das sequências no CDD e HMMER.

Planta	Domínio conservado	CDD	HMMER	Classe de AMP
<i>Linum usitatissimum</i>	Gamma-thionin	9	8	Defensina
	Tryp_alpha_amyl	-	42	LTP
<i>Triticum aestivum</i>	Thionin	3	3	Tionina
	Tryp_alpha_amyl	3	3	LTP
	Antimicrobial21	-	2	α -Hairpinina
	Chitin_bind_1 (ChtBD1)	1	1	Heveína
	Gamma-thionin	4	4	Defensina
	GASA	5	5	Snakina
<i>Oryza sativa</i>	Tryp_alpha_amyl	-	4	LTP
	Thionin	-	2	Tionina
	SLR1-BP	-	7	Knottin
	Gamma-thionin	-	5	Defensina
	Chitin_bind_1 (ChtBD1)	-	1	Heveína

Fonte: Autoria própria.

Algumas sequências apresentaram domínio conservado em ambas as ferramentas, sendo elas: Lus10000290, Lus10019226, Lus10029546, Lus10029544, Lus10004290, Lus10004289, Lus10039622, Lus10019227 (*Linum usitatissimum*); PPepDB_1564, PPepDB_1565, PPepDB_3504, PPepDB_2144, PPepDB_2809, PPepDB_2810, PPepDB_3809, PPepDB_3863 (*Triticum aestivum*).

Com relação às famílias de AMPs vegetais (Tabela 3), foi possível encontrar defensinas e LTPs para o linho; snakinas, LTPs, defensinas, heveína, α -Hairpininas e tioninas foram identificadas para o trigo e LTPs, defensinas, heveína, tioninas e peptídeos do tipo knottin para o arroz, totalizando 87 sequências que foram utilizadas para as etapas posteriores

de caracterização, alinhamento de sequências, análise dos padrões de conservação com a ferramenta WebLogo e, por fim, a modelagem dos AMPs.

Tabela 3 - Total de AMPs de *L. usitatissimum*, *T. aestivum* e *O. sativa*.

Planta	Classe de AMP	Total
<i>Linum usitatissimum</i>	LTP	42
	Defensina	8
<i>Triticum aestivum</i>	LTP	3
	Tionina	3
	α -Hairpinina	2
	Heveína	1
	Defensina	4
	Snakina	5
<i>Oryza sativa</i>	LTP	4
	Tionina	2
	Knottin	7
	Defensina	5
	Heveína	1
Total		87

Fonte: Autoria própria.

Os AMPs vegetais apresentam muitos resíduos do aminoácido cisteína e isso é bastante importante, já que essas estruturas geralmente são estabilizadas por pontes dissulfeto, garantindo uma maior estabilidade a molécula (Moyer; Brechbill; Hicks, 2021).

A Figura 3 destaca os resíduos de cisteína presentes nas diferentes classes de AMPs identificadas, sendo selecionada uma sequência representativa de cada classe para destacar o número de cisteínas e pontes dissulfeto presentes em cada uma delas. Esta abordagem visa

ilustrar as características gerais das diferentes classes de AMPs, utilizando amostras específicas como exemplos.

Analisando as estruturas, observa-se a formação das pontes dissulfeto em cada uma delas. A snakina apresenta 12 cisteínas, resultando em 5 pontes dissulfeto, a heveína possui 10 cisteínas, também formando 5 pontes dissulfeto; a defensina, tionina e knottin contêm 8 cisteínas cada, resultando em 4 pontes dissulfeto em cada caso; LTP tem 8 cisteínas, resultando em 3 pontes dissulfeto; enquanto a α -hairpinina, com 4 cisteínas, forma apenas 1 ponte dissulfeto.

Esses resultados referentes ao número de cisteína e pontes de hidrogênio formadas nas sequências de snakina, defensina e tionina foram encontradas em estudos com AMPs vegetais, no qual Tam e colaboradores (2015) fizeram uma pesquisa com o número total de cisteína de diferentes classes e as pontes dissulfeto formadas nas estruturas (Tam *et al.*, 2015). Nesse estudo, a snakina que apresentou um total de 12 cisteínas e 5 pontes dissulfeto foi Snakin-1, heveína foi EAFP1, NaD1 para defensina, β -Purothionin para tionina (Tam *et al.* 2015). Para a α -hairpinina, os resultados estão de acordo com o trabalho realizado por Slavokhotova e Rogozhin (2020), no qual relata que esses AMPs apresentam 4 cisteínas (Slavokhotova; Rogozhin, 2020). Na literatura, LTPs apresentam 8 cisteínas e 4 pontes dissulfeto (Salminen; Blomqvist; Edqvist, 2016), divergindo apenas o valor referente a formação de pontes dissulfeto, que foram encontradas 3.

Já para os peptídeos do tipo knottin, o resultado referente ao número de cisteínas e pontes dissulfeto divergiu com os estudos realizados por Molesini e colaboradores (2017), que determina que os AMPs do tipo knottin têm 6 resíduos de cisteína e 3 pontes dissulfeto, que formam um nó de cisteína (Molesini *et al.*, 2017).

Figura 3 - Representação esquemática dos resíduos de cisteína dos AMPs de snakina (Traes_5DL_BEB7ADA1D.1 - trigo), heveína (PPepDB_2144 - trigo), defensina (Lus10019226 - linho), tionina (PPepDB_3504 - trigo), LTP (Lus10000082 - linho), knottin (LOC_Os02g07440.1 - arroz) e α -hairpinina (PPepDB_2068 - trigo).



Fonte: Autoria própria.

Foram realizadas a caracterização físico-química (Tabela 4, 5 e 6) e a modelagem *in silico* de 87 sequências e 7 diferentes famílias de AMPs, excluindo aquelas com um peso molecular superior a 10 KDa e sem a presença de domínios conservados.

A caracterização físico-química foi executada no Biopython, considerando os seguintes parâmetros: comprimento total da sequência, peso molecular (em KDa), índice de instabilidade, ponto isoelétrico e carga líquida. Os resultados referentes à média e desvio padrão encontram-se no Material Suplementar.

Para o linho só foram encontradas as classes LTPs e defensinas e os resultados estão disponibilizados na Tabela 4. A média referente ao número total de aminoácidos das sequências de defensinas está de acordo com os valores encontrados na literatura, que é de 45-54 (Tabela 1a). Diferente do peso molecular, que na literatura encontra-se com os seguintes valores 2-5 (Tabela 1a) e a média foi um pouco mais alta. Com relação ao índice de instabilidade, apenas 3 das 8 sequências de AMPs das defensinas são estáveis, representando apenas 37,5% das sequências com um valor de índice de instabilidade menor que 40. Isso justifica que a média tenha sido superior a 40. AMPs considerados instáveis são aqueles que apresentam um índice superior a 40, sendo conhecidos por ter um ciclo de vida curto, ou seja,

proteínas que degradam rapidamente (Guruprasad; Reddy; Pandit, 1990). O ideal é ter um maior número de peptídeos estáveis em decorrência do seu ciclo de vida. Todas as 8 sequências apresentaram carga positiva e, de acordo com o ponto isoelétrico e a média ($pI > 7$), elas são alcalinas.

A classe LTP do linho foi encontrado um número maior de sequências, sendo um total de 42. A média referente ao número total de aminoácidos das sequências de LTPs está de acordo com os valores encontrados na literatura, que é de 70-90 (Tabela 1a), assim como o peso molecular. A média do índice de instabilidade da classe LTP também foi alto e 50% dos AMPs foram consideradas estáveis, já que 21 sequências apresentaram um índice de instabilidade menor que 40. A carga líquida variou entre positiva e negativa, com um total de 29 sequências positivas e 13 negativas. Das 42 sequências, 13 resultaram em um $pI < 7$, sendo classificadas como ácidas e foram as sequências que apresentaram carga negativa, pois em um ambiente com pH neutro ou alcalino, as proteínas com $pI < 7$ tendem a ter uma predominância de grupos carboxílicos ($-COOH$) ou outros grupos ácidos na sua estrutura, resultando em carga líquida negativa (Tokmakov; Kurotani; Sato, 2021). Isso significa que o ponto isoelétrico é fortemente influenciado pela composição química das sequências. Um total de 5 sequências apresentaram um $pI \approx 7$, sendo classificadas como neutras e 24 alcalinas.

Tabela 4 - Caracterização físico-química de *Linum usitatissimum*.

ID	Família	Comprimento (aa)	Peso molecular (KDa)	Índice de instabilidade	Ponto isoelétrico (pI)	Carga líquida
Lus10000290	Defensina	50	5.52	21.35	8.20	1.40
Lus10019226	Defensina	54	5.91	44.69	8.76	3.47
Lus10029546	Defensina	50	5.39	17.10	7.72	0.40
Lus10029544	Defensina	47	5.33	65.16	9.54	7.43
Lus10004290	Defensina	49	5.57	49.42	8.97	4.58
Lus10004289	Defensina	54	6.09	46.83	9.49	7.43
Lus10039622	Defensina	47	5.35	37.46	8.20	1.35
Lus10019227	Defensina	81	9.09	62.67	10.34	8.52
Lus10029445	LTP	96	9.98	34.29	8.68	3.34
Lus10019029	LTP	85	8.95	64.03	8.82	4.33
Lus10022745	LTP	90	9.13	53.21	7.74	0.44
Lus10019770	LTP	94	9.79	17.72	6.96	-0.54
Lus10015279	LTP	90	8.73	27.74	9.07	5.39

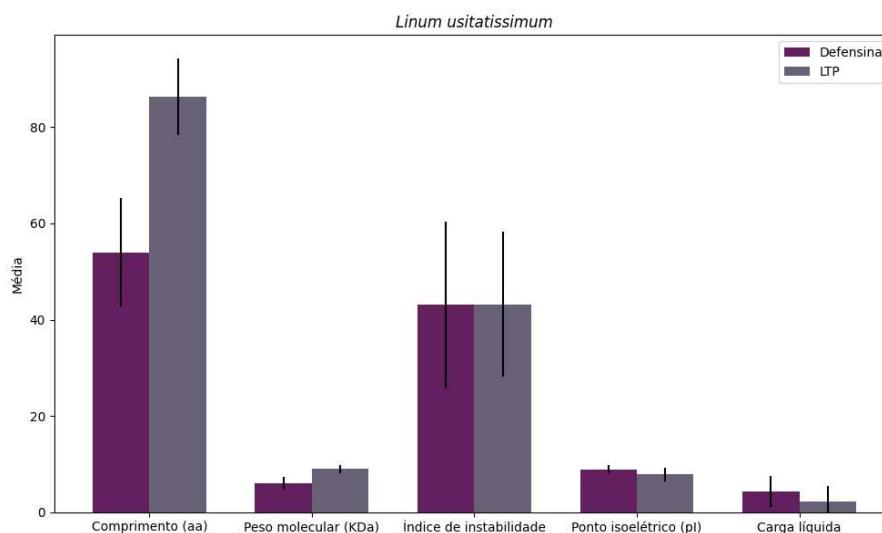
Lus10002741	LTP	76	8.05	52.67	9.25	6.38
Lus10010482	LTP	88	9.11	34.48	6.49	-0.89
Lus10033574	LTP	74	7.59	52.41	8.94	4.40
Lus10009911	LTP	90	9.48	31.66	6.41	-0.55
Lus10029131	LTP	69	7.29	58.25	9.25	6.40
Lus10014167	LTP	90	9.16	52.49	6.73	-0.56
Lus10005763	LTP	70	7.16	73.17	4.37	-1.64
Lus10016582	LTP	80	8.43	55.01	6.75	-0.59
Lus10025151	LTP	93	9.73	33.77	4.93	-3.62
Lus10025866	LTP	90	9.51	47.08	8.45	2.04
Lus10042512	LTP	92	9.88	31.13	5.06	-1.93
Lus10032575	LTP	87	9.41	50.89	9.50	8.03
Lus10013030	LTP	69	7.24	50.73	9.08	5.40
Lus10005010	LTP	74	7.90	63.52	8.89	4.08
Lus10033573	LTP	83	8.88	68.23	5.18	-3.53
Lus10025234	LTP	93	9.47	38.99	9.35	7.34
Lus10005011	LTP	80	8.36	34.26	8.86	4.52
Lus10004677	LTP	93	9.48	31.98	8.48	2.39
Lus10039511	LTP	75	7.85	57.36	8.92	4.39
Lus10017615	LTP	74	7.65	58.17	5.08	-2.53
Lus10027345	LTP	86	9.35	31.17	8.17	1.38
Lus10026418	LTP	89	9.01	19.94	9.05	5.34
Lus10031282	LTP	92	9.84	34.60	6.03	-0.93
Lus10007280	LTP	90	8.97	23.07	8.92	4.44
Lus10019030	LTP	85	8.90	40.43	6.65	-0.89
Lus10017616	LTP	91	9.60	57.10	8.77	4.35
Lus10028003	LTP	92	9.95	55.99	7.70	0.42
Lus10001703	LTP	90	9.41	28.16	8.36	1.71
Lus10028002	LTP	92	9.95	55.99	7.70	0.42
Lus10000082	LTP	94	9.88	24.22	7.47	0.07
Lus10016357	LTP	94	9.75	20.55	7.73	0.47
Lus10009872	LTP	97	9.94	30.99	8.47	2.34
Lus10029226	LTP	90	8.94	20.57	8.92	4.44
Lus10038233	LTP	90	9.53	49.22	8.45	2.04
Lus10015278	LTP	92	9.03	35.13	9.41	7.40

Lus10016323	LTP	76	8.07	49.42	9.41	7.38
Lus10012384	LTP	92	9.86	64.31	6.95	-0.54

Fonte: Autoria própria.

O Gráfico 1a corresponde aos dados estatísticos (média e desvio padrão) das sequências do linho, onde o eixo x representa os valores referentes a média de cada parâmetro e o eixo y corresponde aos parâmetros (comprimento, peso molecular, índice de instabilidade, ponto isoelétrico e carga). Ao analisar os resultados do gráfico, é possível observar que os AMPs da classe LTP tendem a ter um comprimento médio maior em comparação com as defensinas, esse resultado corrobora com os valores encontrados na literatura. Assim como o peso molecular, que também foi superior. Quanto ao índice de instabilidade, as defensinas exibiram um desvio padrão maior em comparação com LTP. Além disso, vale destacar que a média do índice de instabilidade obtido pode ser decorrente da presença de um maior número de AMPs instáveis, pois foi de 43.085. Os IDs com as maiores sequências que apresentaram carga líquida negativa foram: Lus10033573 (-3.53, linho), LOC_Os04g15740.1 (-3.94, arroz).

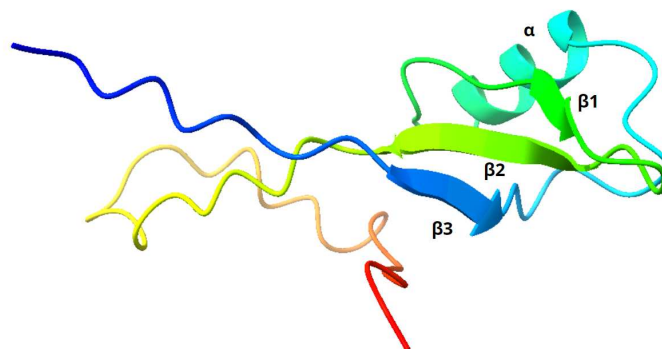
Gráfico 1a - Dados estatísticos de *L. usitatissimum*.



Fonte: Autoria própria.

O valor de comprimento de sequência mais alto para as defensinas foi de 81 (Lus10019227), apesar de ser superior aos dados esperados (45-54), a sequência apresentou uma modelagem característica de AMP da classe defensina (Figura 3), onde tem uma folha β com 3 fitas e α -hélice alinhada (Tabela 1b).

Figura 3 - Defensina do linho (Lus10019227) no ChimeraX.



Fonte: ChimeraX (2024).

O trigo foi a espécie vegetal que teve uma maior distribuição de classes. Para o trigo, foram encontradas as classes LTP, snakina, defensina, heveína, α -hairpinina e tionina, e os resultados estão disponibilizados na Tabela 5.

Para a classe LTP, a média do comprimento da sequência e peso molecular estão de acordo com a literatura (Tabela 1a), que varia de 70-90 (comprimento) e 7-9 KDa (peso molecular). Em contrapartida, a média referente ao índice de instabilidade foi alta, pois apenas uma sequência, dentre as 3 do linho, é estável. Todas as sequências do linho são alcalinas e com média de carga líquida positiva. O desvio padrão referente ao comprimento das sequências do linho foi 0, pois todos os valores eram iguais, sem dispersão em torno da média.

As médias referentes ao comprimento e peso molecular das snakinas foram abaixo do valor encontrado na literatura (Tabela 1a). A média do índice de instabilidade resultou em um valor considerável bom, tendo em vista que sequências consideradas instáveis têm um índice acima de 40. Quanto ao pI, é possível observar 4 sequências alcalinas e 1 ácida e todas com carga positiva.

As médias pertencentes ao comprimento e peso molecular das defensinas estão de acordo com a literatura (Tabela 1a). Em contrapartida, a média referente ao índice de instabilidade foi bastante superior ao esperado. Todas as sequências são alcalinas e com carga positiva.

Para a classe heveína, foi obtida apenas 1 sequência no qual seus valores estão próximos com os dados da literatura (Tabela 1a). O comprimento total da sequência resultou em 44, e os valores padrão são 29-45. O peso molecular obtido foi de 4.45 e o valor de referência é de 4.7. A sequência de heveína apresentou um índice de instabilidade de 24.64, ponto isoelétrico de 8.41 e carga positiva de 2.35.

A partir dos dados é possível inferir que o comprimento e peso molecular de α -hairpinina foram ideais, ao comparar com os valores da Tabela 1a. Das 2 sequências de α -hairpinina, de acordo com o pI, é possível concluir que uma sequência é alcalina e a outra é ácida. Além disso, uma sequência apresentou um índice de instabilidade de 29.34 (estável) e a outra foi de 50.59 (instável) e ambas com carga positiva.

O comprimento das 3 sequências de tionina foi o mesmo (45 aa), por isso o desvio padrão resultou em 0. A média resultante para as tioninas foi um pouco abaixo do encontrado na literatura (5 KDa) (Tabela 1a). Todas as sequências resultaram em um índice de instabilidade < 40, sendo então AMPs estáveis e alcalinas (pI > 7).

Tabela 5 - Caracterização físico-química de *Triticum aestivum*.

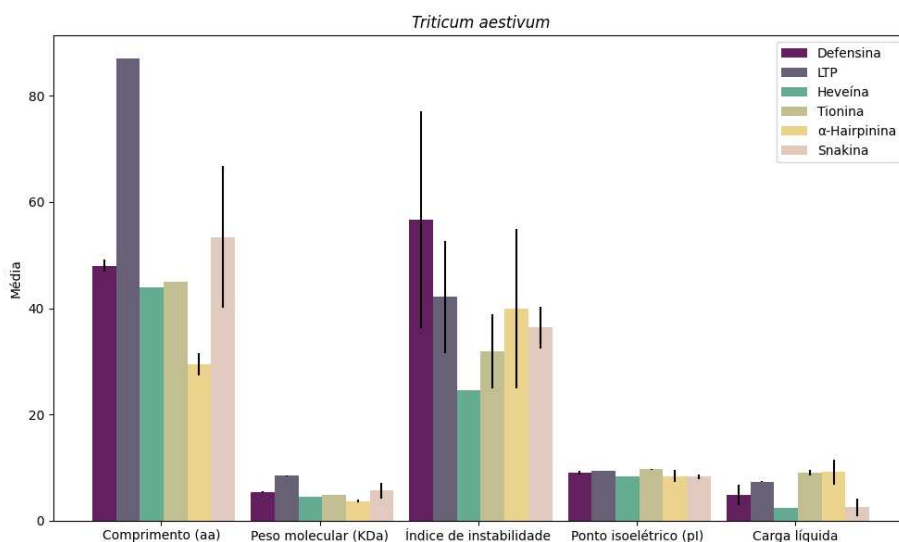
ID	Família	Comprimento (aa)	Peso molecular (KDa)	Índice de instabilidade	Ponto isoeletrico (pI)	Carga líquida
Traes_5AL_23DA4A352.1	Snakina	44	4.55	39.39	8.15	1.36
Traes_5DL_9EC9EBA1F.1	Snakina	66	7.02	40.07	8.49	3.32
Traes_5DL_BEB7ADA1D.1	Snakina	66	6.99	37.89	8.65	4.32
Traes_5AL_2F742634B.1	Snakina	55	5.93	31.20	8.54	3.31
Traes_5AL_120F47252.1	Snakina	36	3.81	33.36	7.67	0.37
PPepDB_3863	LTP	87	8.46	40.86	9.35	7.38
PPepDB_3873	LTP	87	8.35	53.33	9.35	7.34
PPepDB_3874	LTP	87	8.53	32.36	9.33	7.38
PPepDB_2809	Defensina	49	5.59	45.41	8.66	3.40
PPepDB_2810	Defensina	49	5.54	43.87	8.73	3.43
PPepDB_3809	Defensina	47	5.15	50.45	9.12	5.35
PPepDB_3991	Defensina	47	5.24	86.92	9.49	7.35

PPepDB_214 4	Heveína	44	4.45	24.64	8.41	2.35
PPepDB_206 1	α -Hairpinina	28	3.52	29.34	9.17	3.54
PPepDB_206 8	α -Hairpinina	31	3.87	50.59	7.63	0.18
PPepDB_156 4	Tionina	45	4.83	39.54	9.75	9.35
PPepDB_156 5	Tionina	45	4.93	25.71	9.54	8.34
PPepDB_350 4	Tionina	45	4.86	30.41	9.75	9.35

Fonte: Autoria própria.

No Gráfico 1b, é possível destacar que LTP foi a que obteve uma maior média de comprimento (87) e peso molecular (8.44). O índice de instabilidade foi a característica que obteve um desvio padrão alto, sendo defensina a que teve o maior. O ponto isoelétrico, peso molecular e carga líquida não apresentaram muita variação em relação ao desvio padrão.

Gráfico 1b - Dados estatísticos de *T. aestivum*.



Fonte: Autoria própria.

Para o arroz, foram encontradas as classes LTP, defensina, heveína, tionina e knottin, e os resultados estão disponibilizados na Tabela 6. Para a classe LTP, o comprimento e peso molecular estão de acordo com as características dessa classe. Das 4 sequências de LTP, todas apresentam ponto isoelétrico característico de sequências alcalinas, com carga líquida positiva e duas sequências estáveis.

Com base na estatística dos dados, é possível concluir que todas as defensinas foram aniônicas e ácidas, com 50% de AMPs estáveis. A média da carga líquida resultou em um valor negativo, pois todas as sequências de defensina apresentaram carga negativa.

Para a classe heveína, foi obtida apenas 1 sequência no qual seus valores estão diferentes com relação aos dados da literatura (Tabela 1a), no que diz respeito o comprimento da sequência, peso molecular e índice de instabilidade. O comprimento da sequência foi de 64 e o peso molecular de 6.43 KDa. Essa sequência é aniônica, ácida e instável.

O comprimento da sequência e peso molecular das tioninas estão de acordo com os valores de referência (Tabela 1a). Além disso, a carga das sequências é positiva, ponto isoeletrico característico de sequências alcalinas e média de índice de instabilidade em conformidade.

Dos 6 peptídeos do tipo “knottin”, 6 são aniônicos e 1 catiônico. Além disso, 6 sequências são ácidas e 1 alcalina. Todas as sequências apresentaram um índice de instabilidade muito alto e com grandes variações em relação ao comprimento e peso molecular.

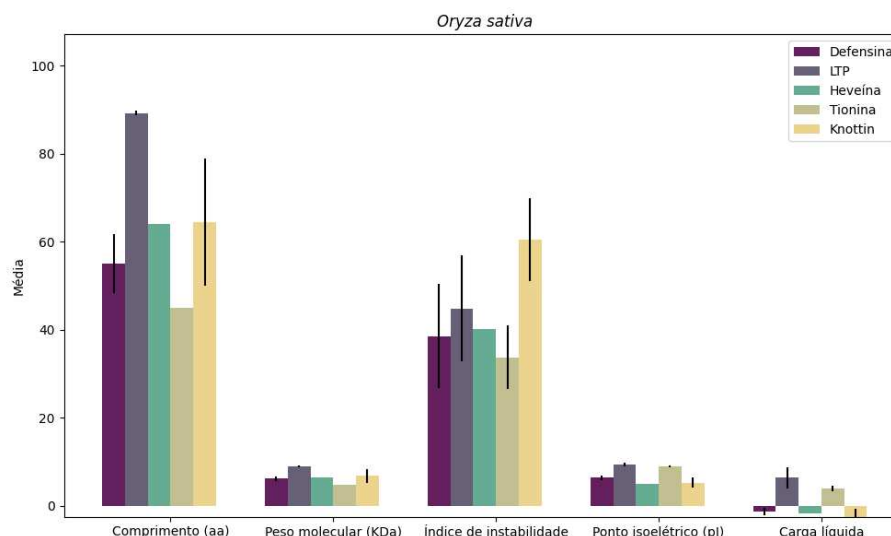
Tabela 6 - Caracterização físico-química de *Oryza sativa*.

ID	Família	Comprimento (aa)	Peso molecular (KDa)	Índice de instabilidade	Ponto isoeletrico (pI)	Carga líquida
PPepDB_1562	LTP	89	9.08	55.47	9.01	4.38
PPepDB_3877	LTP	89	8.68	35.69	9.41	7.35
PPepDB_3906	LTP	89	8.83	33.23	9.91	9.37
PPepDB_3949	LTP	90	9.18	54.97	8.92	4.35
LOC_Os06g22880.1	Defensina	52	5.83	25.23	6.94	-0.58
LOC_Os02g07550.1	Defensina	67	7.12	54.66	5.66	-1.61
LOC_Os06g23060.1	Defensina	51	5.76	43.90	6.28	-1.47
LOC_Os01g40220.1	Defensina	52	5.75	29.19	6.94	-0.58
LOC_Os0	Defensina	53	6.25	39.86	6.02	-2.39

1g61360.1						
LOC_Os04g41640.1	Heveína	64	6.43	40.12	5.01	-1.64
PPepDB_3920	Tionina	45	4.80	28.65	9.10	4.43
PPepDB_3921	Tionina	45	4.80	38.80	8.90	3.44
LOC_Os02g20130.1	Knottin	64	6.51	62.88	4.58	-3.91
LOC_Os11g39910.1	Knottin	75	7.91	74.59	6.00	-0.96
LOC_Os03g56682.1	Knottin	51	5.54	49.21	4.50	-3.56
LOC_Os04g15740.1	Knottin	59	6.37	62.53	4.70	-3.94
LOC_Os02g07440.1	Knottin	50	5.02	70.56	4.68	-1.64
LOC_Os04g31250.1	Knottin	91	9.46	53.65	4.88	-3.88
LOC_Os02g53600.1	Knottin	61	6.64	53.37	7.69	0.35

Fonte: Autoria própria.

No Gráfico 1c é possível visualizar que o desvio padrão do comprimento e índice de instabilidade variaram bastante, com exceção a LTP (comprimento). A maior média do índice de instabilidade foi para os peptídeos do tipo knottin de 60.97. Como a carga líquida apresentou valores positivos e negativos, é possível observar no gráfico alguns blocos no eixo y negativo.

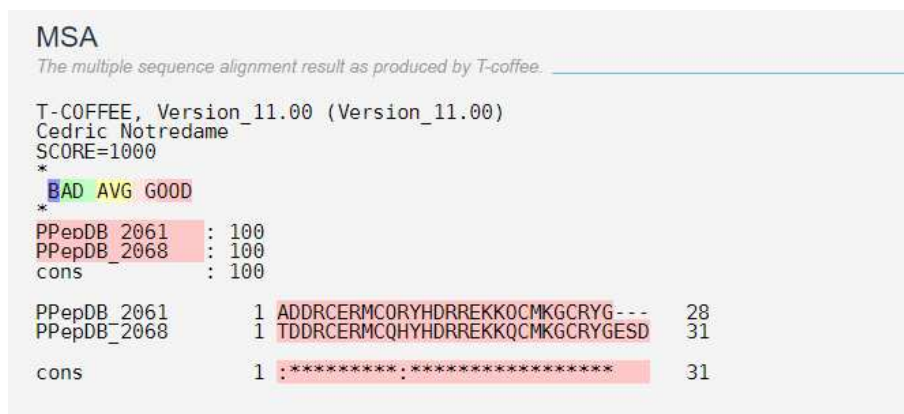
Gráfico 1c - Dados estatísticos de *O. sativa*.

Fonte: Autoria própria.

Para a análise de domínios conservados, inicialmente foi realizado o alinhamento de sequências com o programa M-Coffee para as seguintes classes: LTP, defensina, tionina, α -hairpinina, knottin, snakina e heveína. Cada classe foi analisada individualmente.

Com relação ao alinhamento de sequências, todos os scores do foram altos (α -hairpinina: 1000, snakina: 1000, tionina: 959, heveína: 799, LTP: 768, defensina: 639, knottin: 626), destacando as snakinas e α -hairpininas, que obtiveram uma pontuação de 1000 e uma notável conservação das sequências (Figura 4). As sequências altamente conservadas das snakinas e α -hairpininas podem significar uma importância na estrutura e função destes AMPs, pois sugere que esses peptídeos têm a manutenção de padrões específicos de aminoácidos ao longo da evolução.

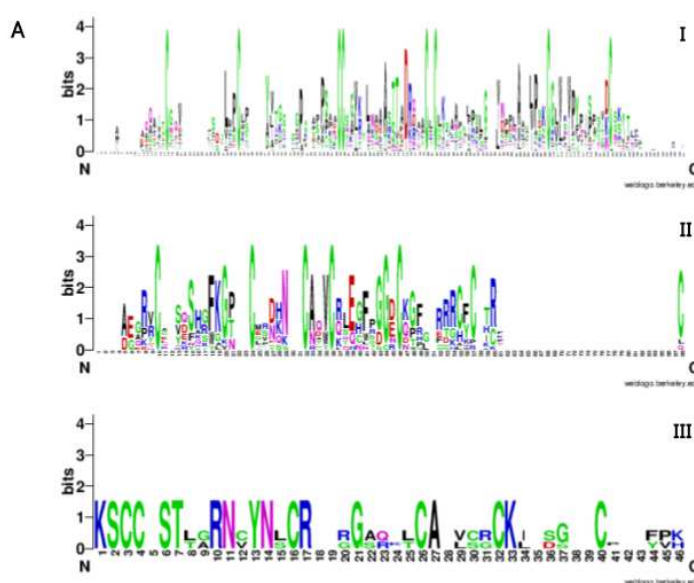
Figura 4 - Alinhamento par-a-par de sequências da α -hairpinina através do M-Coffee. O (*) indica a presença de resíduos altamente conservados e (:) são os resíduos semelhantes.



Fonte: Autoria própria.

Os alinhamentos foram submetidos ao programa WebLogo para obter a representação gráfica das sequências alinhadas, com exceção das classes α -hairpinina e heveína (apresentam somente 2 sequências), uma vez que a ferramenta só gera a representação de alinhamentos com mais de 3 sequências. Sendo assim, foi possível gerar um gráfico do padrão de aminoácidos conservados nas sequências de LTP, tionina e defensina (Figura 5A), snakina e knottin (Figura 5B).

Figura 5A - Visualização dos aminoácidos conservados de LTP (I), defensina (II) e tionina (III) com o WebLogo.

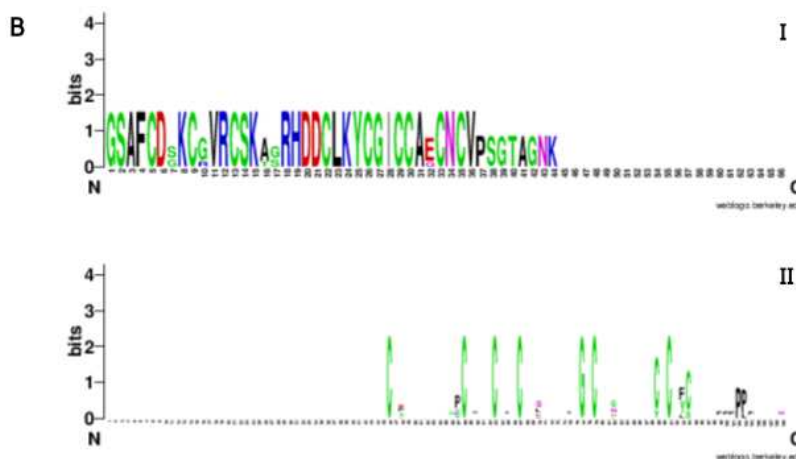


Fonte: Autoria própria.

Quanto maior foi o número de sequências, menor será a visibilidade do gráfico padrão ilustrado com o WebLogo, pois isso leva a uma maior diversidade de aminoácidos presentes em cada posição da sequência, isso ocorreu com as sequências do linho e das defensinas, mas mesmo assim ainda é possível verificar a presença das cisteínas, dentre os aminoácidos.

Já os aminoácidos da tionina e snakina são os mais visíveis, pois o gráfico é realizado com um menor número de sequências. Por fim, o peptídeo do tipo knottin (Figura 5B - II) é possível visualizar principalmente as cisteínas. O espaço presente em todos os padrões do WebLogo representam os gaps do alinhamento, ou seja, a pontuação negativa em um alinhamento e a ausência de um aminoácido. A classe knottin foi a família com um maior número de gaps e também a que obteve o menor score dentre as outras classes.

Figura 5B - Visualização dos aminoácidos conservados de snakina (I) e knottin (II) com o WebLogo.



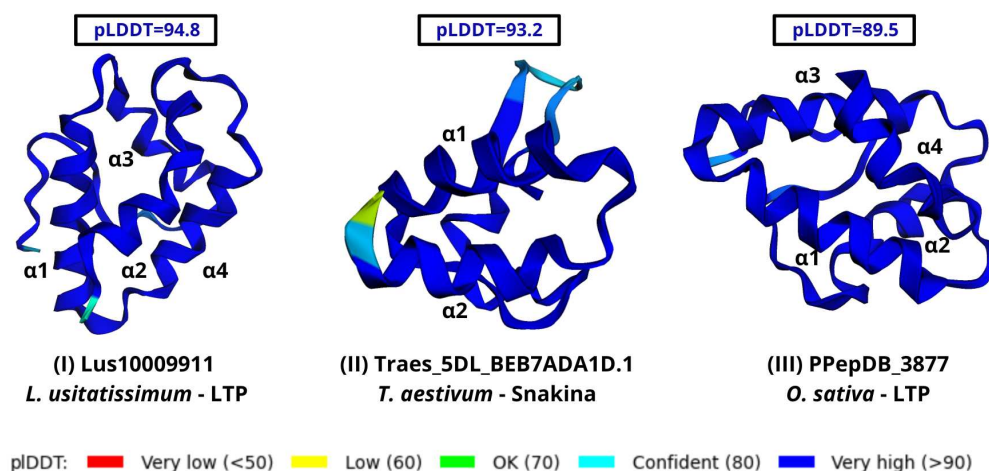
Fonte: Autoria própria.

A modelagem *in silico* foi realizada através do ColabFold (Figura 5) (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>). O ColabFold utiliza redes neurais profundas para modelagem e retorna 5 modelos tridimensionais diferentes de uma estrutura e esses modelos fazem uso da métrica pLDDT (Mirdita *et al.* 2022).

A métrica pLDDT foi utilizada para verificar a qualidade das sequências (Mirdita *et al.* 2022). O pLDDT tem 5 escalas: Muito alto (cor: azul, pLDDT > 90), confiante (cor: azul claro, pLDDT: 80), ok (cor: verde, pLDDT: 70), baixo (cor: amarelo, pLDDT: 60) e muito baixo (cor: vermelho, pLDDT < 50) (Mirdita *et al.* 2022).

As sequências que apresentaram uma métrica de confiabilidade muito alta pertenciam às classes de LTP (*L. usitatissimum*) e snakina (*T. aestivum*), totalizando 9 sequências, sendo 6 do linho e 3 do trigo. Em conjunto com as classificações "confiante" e "ok", foi possível identificar um total de 76 sequências com alta qualidade, o que representa 87,36% do total de 87 sequências modeladas. Apenas 12,64% das sequências foram classificadas como baixo ou muito baixo em termos de confiabilidade.

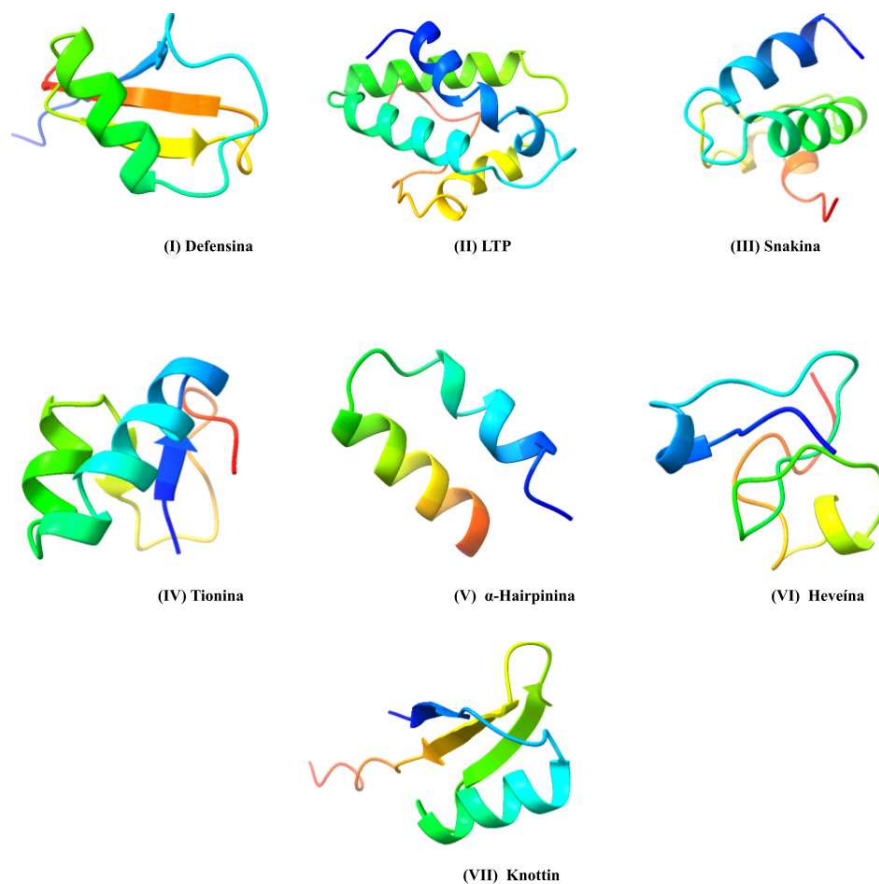
Figura 6 - AMPs das espécies vegetais com maior pLDDT.



Fonte: Autoria própria.

A Figura 7 representa a modelagem *in silico* dos AMPs de cada classe encontrada no trabalho. A estrutura I representa a defensina Lus10029546 (pLDDT: 83.8), composta por uma folha β com 3 fitas nas cores azul, laranja e amarelo e α -hélice alinhada (verde). Já a estrutura II consiste num LTP (pLDDT: 90.2), com α -hélices interligadas. A molécula III é uma snakina Traes_5AL_2F742634B.1 (pLDDT: 92.1) e sua estrutura é composta por 2 α -hélices, porém uma é curta. A molécula IV é a tionina PPepDB_1565 (pLDDT: 75.3), que apresenta tanto α -hélice quanto folha β , sendo duas de cada. A α -hairpinina PPepDB_2061 (pLDDT: 84.8) (V) é constituída por 2 α -hélices que são interligadas por uma alça. A molécula V não apresentou uma conformação que é possível identificar como ela está estruturada, mas apresenta o domínio ChtBD1, característico da heveína. A heveína é caracterizada por apresentar uma folha β com 3 fitas e 1 α -hélice curta, diferente da imagem. Os peptídeos do tipo “knottin” LOC_Os02g07440.1 (pLDDT: 76.7) apresentam folha β e α -hélice, assim como na Figura 7. Todas as estruturas encontram-se em conformidade com a representação desses peptídeos na literatura (Tabela 1b), com exceção da heveína.

Figura 7 - Representação das classes de AMPs no ChimeraX.



Fonte: ChimeraX (2024).

A análise das sequências de *L. usitatissimum*, *T. aestivum* e *O. sativa* revelou padrões distintos de conservação de domínios entre as espécies, com destaque para a presença de domínios conservados em LTP em todas as espécies. Observou-se também que o aumento do número de sequências resultou em uma menor visibilidade dos padrões ilustrados com o WebLogo, devido à maior diversidade de aminoácidos em cada posição da sequência. No entanto, mesmo com essa diversidade, foi possível verificar a presença das cisteínas, especialmente nos peptídeos do tipo knottin. Além disso, a análise de confiabilidade destacou a predominância de sequências de alta qualidade nas classes de LTP e snakina. Esses resultados sugerem uma conservação funcional dessas sequências ao longo da evolução, indicando uma potencial importância biotecnológica desses peptídeos em diversas espécies de plantas, sendo necessário estudos futuros para comprovar.

8. CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo a caracterização físico-química das sequências de *L. usitatissimum*, *T. aestivum* e *O. sativa*, juntamente com a modelagem *in silico* e análise dos resíduos de aminoácidos presentes nessas sequências. Como resultado, foi observado que a família de AMP mais abundante e presente em todas as espécies foi a LTP. Além disso, todas as 87 sequências analisadas apresentaram os resíduos de cisteína característicos das respectivas classes, sendo possível observar os padrões através do WebLogo. As informações referentes às propriedades físico-químicas, tais como comprimento, peso molecular, índice de instabilidade, pI e carga líquida, e os modelos tridimensionais obtidos a partir do ColabFold das sequências, serão publicizadas no banco de dados PlantAMP.

REFERÊNCIAS

- ADAMS, Paul C. **The boundless self: Communication in physical and virtual spaces.** Syracuse University Press, 2005.
- ALMAGRO ARMENTEROS, José Juan et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. **Nature biotechnology**, v. 37, n. 4, p. 420-423, 2019.
- ALMASIA, Natalia Inés et al. Successful production of the potato antimicrobial peptide Snakin-1 in baculovirus-infected insect cells and development of specific antibodies. **BMC biotechnology**, v. 17, p. 1-11, 2017. <https://doi.org/10.1186/s12896-017-0401-2>.
- ARCHER, B. L. The proteins of Hevea brasiliensis latex. 4. Isolation and characterization of crystalline hevein. **Biochemical Journal**, v. 75, n. 2, p. 236, 1960. <https://doi.org/10.1042/bj0750236>.
- ATTAH, Francis Alfred et al. Nutritional and Pharmaceutical Applications of Under-Explored Knottin Peptide-Rich Phytomedicines. **Plants**, v. 11, n. 23, p. 3271, 2022. <https://doi.org/10.3390/plants11233271>.
- AZMI, Sarfuddin; HUSSAIN, Mohd Kamil. Analysis of structures, functions, and transgenic of phytopeptides defensin and thionin: a review. **Beni-Suef University Journal of Basic and Applied Sciences**, v. 10, n. 1, p. 1-11, 2021. <https://doi.org/10.1186/s43088-020-00093-5>.
- BAKARE, Olalekan Olanrewaju et al. Plant antimicrobial peptides (PAMPs): Features, applications, production, expression, and challenges. **Molecules**, v. 27, n. 12, p. 3703, 2022. <https://doi.org/10.3390/molecules27123703>.
- BASSI, Sebastian. A primer on python for life science researchers. **PLoS computational biology**, v. 3, n. 11, p. e199, 2007. <https://doi.org/10.1371/journal.pcbi.0030199>.
- BAYAT, Ardeshtir. Bioinformatics.(Science, Medicine, and the Future). **British Medical Journal**, v. 324, n. 7344, p. 1018-1023, 2002.
- BEMILLER, JAMES N. et al. Aloe, chia, flaxseed, okra, psyllium seed, quince seed, and tamarind gums. In: Industrial gums. **Academic Press**, 1993. p. 227-256. <https://doi.org/10.1016/B978-0-08-092654-4.50013-9>.
- BOBA, Aleksandra et al. Absciscic Acid—Defensive Player in Flax Response to Fusarium culmorum Infection. **Molecules**, v. 27, n. 9, p. 2833, 2022. <https://doi.org/10.3390/molecules27092833>.
- BOUTET, Emmanuel et al. UniProtKB/Swiss-Prot: the manually annotated section of the UniProt KnowledgeBase. In: **Plant bioinformatics: methods and protocols**. Totowa, NJ: Humana Press, 2007. p. 89-112. https://doi.org/10.1007/978-1-59745-535-0_4.
- BURLEY, Stephen K. et al. RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and

energy. **Nucleic acids research**, v. 47, n. D1, p. D464-D474, 2019. <https://doi.org/10.1093/nar/gky1004>.

CAPRIOTTI, Anna Laura et al. Identification of potential bioactive peptides generated by simulated gastrointestinal digestion of soybean seeds and soy milk proteins. **Journal of Food Composition and Analysis**, v. 44, p. 205-213, 2015. <https://doi.org/10.1016/j.jfca.2015.08.007>.

CHENG, Xinran et al. Identification and analysis of the GASR gene family in common wheat (*Triticum aestivum* L.) and characterization of TaGASR34, a gene associated with seed dormancy and germination. **Frontiers in genetics**, v. 10, p. 980, 2019. <https://doi.org/10.3389/fgene.2019.00980>.

CHUNG, Chia-Ru et al. Characterization and identification of antimicrobial peptides with different functional activities. **Briefings in bioinformatics**, v. 21, n. 3, p. 1098-1114, 2020. <https://doi.org/10.1093/bib/bbz043>.

CLARA PESTANA-CALSA, Maria et al. Bioinformatics-coupled molecular approaches for unravelling potential antimicrobial peptides coding genes in Brazilian native and crop plant species. **Current Protein and Peptide Science**, v. 11, n. 3, p. 199-209, 2010. <https://doi.org/10.2174/138920310791112138>.

COCK, Peter JA et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. **Bioinformatics**, v. 25, n. 11, p. 1422, 2009. <https://doi.org/10.1093/bioinformatics/btp163>.

CRAIK, David J. et al. Cyclotides as a basis for drug design. **Expert opinion on drug discovery**, v. 7, n. 3, p. 179-194, 2012. <https://doi.org/10.1517/17460441.2012.661554>.

CROOKS, Gavin E. et al. WebLogo: a sequence logo generator. **Genome research**, v. 14, n. 6, p. 1188-1190, 2004. <https://doi.org/10.1101/gr.849004>.

DAS, Durdam et al. PlantPepDB: A manually curated plant peptide database. **Scientific reports**, v. 10, n. 1, p. 2194, 2020. <https://doi.org/10.1038/s41598-020-59165-2>.

DE CALEYA, R. Fernandez et al. Susceptibility of phytopathogenic bacteria to wheat purothionins in vitro. **Applied microbiology**, v. 23, n. 5, p. 998-1000, 1972. <https://doi.org/10.1128/am.23.5.998-1000.1972>.

DE SANTANA LOPES, Amanda et al. The *Linum usitatissimum* L. plastome reveals atypical structural evolution, new editing sites, and the phylogenetic position of Linaceae within Malpighiales. **Plant cell reports**, v. 37, n. 2, p. 307-328, 2018. <https://doi.org/10.1007/s00299-017-2231-z>.

DENARDI, Laura Bedin et al. Activity of MSI-78, h-Lfl-11 and cecropin B antimicrobial peptides alone and in combination with voriconazole and amphotericin B against clinical isolates of *Fusarium solani*. **Journal of Medical Mycology**, v. 31, n. 2, p. 101119, 2021. <https://doi.org/10.1016/j.mycmed.2021.101119>.

DE OLIVEIRA CARVALHO, André; MOREIRA GOMES, Valdirene. Plant defensins and

defensin-like peptides-biological activities and biotechnological applications. **Current pharmaceutical design**, v. 17, n. 38, p. 4270-4293, 2011. <https://doi.org/10.2174/138161211798999447>.

DUPASQUIER, Chantal MC et al. Effects of dietary flaxseed on vascular contractile function and atherosclerosis during prolonged hypercholesterolemia in rabbits. **American Journal of Physiology-Heart and Circulatory Physiology**, v. 291, n. 6, p. H2987-H2996, 2006. <https://doi.org/10.1152/ajpheart.01179.2005>.

EDDY, Sean. HMMER user's guide. **Department of Genetics, Washington University School of Medicine**, v. 2, n. 1, p. 13, 1992.

EDSTAM, Monika M. et al. Evolutionary history of the non-specific lipid transfer proteins. **Molecular plant**, v. 4, n. 6, p. 947-964, 2011.

EKINS, Sean; MESTRES, Jordi; TESTA, Bernard. In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. **British journal of pharmacology**, v. 152, n. 1, p. 9-20, 2007. <https://doi.org/10.1038/sj.bjp.0707305>.

EKMEKCI, Berk; MCANANY, Charles E.; MURA, Cameron. An introduction to programming for bioscientists: a Python-based primer. **PLoS computational biology**, v. 12, n. 6, p. e1004867, 2016. <https://doi.org/10.1371/journal.pcbi.1004867>.

FORSLUND, Kristoffer; PEKKARI, Isabella; SONNHAMMER, Erik LL. Domain architecture conservation in orthologs. **BMC bioinformatics**, v. 12, p. 1-14, 2011. <https://doi.org/10.1186/1471-2105-12-326>.

GOYAL, Ravinder K.; MATTOO, Autar K. Multitasking antimicrobial peptides in plant development and host defense against biotic/abiotic stress. **Plant Science**, v. 228, p. 135-149, 2014. <https://doi.org/10.1016/j.plantsci.2014.05.012>.

GURUPRASAD, Kunchur; REDDY, BV Bhasker; PANDIT, Madhusudan W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. **Protein Engineering, Design and Selection**, v. 4, n. 2, p. 155-161, 1990.

GAWDE, Ulka et al. CAMPR4: a database of natural and synthetic antimicrobial peptides. **Nucleic Acids Research**, v. 51, n. D1, p. D377-D383, 2023. <https://doi.org/10.1093/nar/gkac933>.

GUZMÁN-RODRÍGUEZ, Jaquelina Julia et al. Plant antimicrobial peptides as potential anticancer agents. **BioMed research international**, v. 2015, 2015. <https://doi.org/10.1155/2015/735087>.

HAMMAMI, Riadh et al. PhytAMP: a database dedicated to antimicrobial plant peptides. **Nucleic acids research**, v. 37, n. suppl_1, p. D963-D968, 2009. <https://doi.org/10.1093/nar/gkn655>.

HAN, Xiaohong et al. Discovery and Characterization of MaK: A Novel Knottin Antimicrobial Peptide from *Monochamus alternatus*. **International Journal of Molecular**

Sciences, v. 24, n. 24, p. 17565, 2023. <https://doi.org/10.3390/ijms242417565>.

HERBEL, Vera; SIEBER-FRANK, Julia; WINK, Michael. The antimicrobial peptide snakin-2 is upregulated in the defense response of tomatoes (*Solanum lycopersicum*) as part of the jasmonate-dependent signaling pathway. **Journal of plant physiology**, v. 208, p. 1-6, 2017. <https://doi.org/10.1016/j.jplph.2016.10.006>.

HUAN, Yuchen et al. Antimicrobial peptides: classification, design, application and research progress in multiple fields. **Frontiers in microbiology**, v. 11, p. 582779, 2020. <https://doi.org/10.3389/fmicb.2020.582779>.

HUNTER, John D. Matplotlib: A 2D graphics environment. **Computing in science & engineering**, v. 9, n. 03, p. 90-95, 2007. <https://doi.org/10.1109/MCSE.2007.55>.

IQBAL, Aneela et al. Antimicrobial peptides as effective tools for enhanced disease resistance in plants. **Plant Cell, Tissue and Organ Culture (PCTOC)**, v. 139, p. 1-15, 2019. <https://doi.org/10.1007/s11240-019-01668-6>.

IRELAND, David C. et al. Isolation, sequencing, and structure-activity relationships of cyclotides. **Journal of natural products**, v. 73, n. 9, p. 1610-1622, 2010. <https://doi.org/10.1021/np1000413>.

JOHNSON, Mark et al. NCBI BLAST: a better web interface. **Nucleic acids research**, v. 36, n. suppl_2, p. W5-W9, 2008. <https://doi.org/10.1093/nar/gkn201>.

KANAPIN, Alexander et al. Genomic regions associated with Fusarium wilt resistance in flax. **International Journal of Molecular Sciences**, v. 22, n. 22, p. 12383, 2021. <https://doi.org/10.3390/ijms222212383>.

KINTZING, James R.; COCHRAN, Jennifer R. Engineered knottin peptides as diagnostics, therapeutics, and drug delivery vehicles. **Current opinion in chemical biology**, v. 34, p. 143-150, 2016. <https://doi.org/10.1016/j.cbpa.2016.08.022>.

KUMAR, Prashant; KIZHAKKEDATHU, Jayachandran N.; STRAUS, Suzana K. Antimicrobial peptides: diversity, mechanism of action and strategies to improve the activity and biocompatibility in vivo. **Biomolecules**, v. 8, n. 1, p. 4, 2018. <https://doi.org/10.3390/biom8010004>.

LAY, F. T.; ANDERSON, M. A. Defensins-components of the innate immune system in plants. **Current protein and peptide science**, v. 6, n. 1, p. 85-101, 2005. <https://doi.org/10.2174/1389203053027575>.

LEI, Jun et al. The antimicrobial peptides and their potential clinical applications. **American journal of translational research**, v. 11, n. 7, p. 3919, 2019. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6684887/>.

LEPHUTHING, Mantshiuwa C. et al. Progress and Challenges in Improving Nutritional Quality in Wheat. **Wheat Improvement, Management and Utilization**, v. 345, 2017.

- LI, Heng et al. The sequence alignment/map format and SAMtools. **bioinformatics**, v. 25, n. 16, p. 2078-2079, 2009. <https://doi.org/10.1093/bioinformatics/btp352>.
- LI, Jianguo et al. Membrane active antimicrobial peptides: translating mechanistic insights to design. **Frontiers in neuroscience**, v. 11, p. 244777, 2017. <https://doi.org/10.3389/fnins.2017.00073>.
- LI, Junpeng et al. Plant antimicrobial peptides: structures, functions, and applications. **Botanical Studies**, v. 62, n. 1, p. 1-15, 2021. <https://doi.org/10.1186/s40529-021-00312-x>.
- LIMA, Adrienne M. et al. Plant antimicrobial peptides: An overview about classification, toxicity and clinical applications. **International Journal of Biological Macromolecules**, v. 214, p. 10-21, 2022. <https://doi.org/10.1016/j.ijbiomac.2022.06.043>.
- LIU, Fang et al. Non-specific lipid transfer proteins in plants: presenting new advances and an integrated functional analysis. **Journal of experimental botany**, v. 66, n. 19, p. 5663-5681, 2015. <https://doi.org/10.1093/jxb/erv313>.
- LI, Pai et al. The lifecycle of the plant immune system. **Critical reviews in plant sciences**, v. 39, n. 1, p. 72-100, 2020. <https://doi.org/10.1080/07352689.2020.1757829>.
- LÓPEZ-GARCÍA, Guadalupe et al. Antioxidant and antimicrobial peptides derived from food proteins. **Molecules**, v. 27, n. 4, p. 1343, 2022. <https://doi.org/10.3390/molecules27041343>.
- MACHADO, J. da C. et al. Tratamento de sementes no controle de fitopatógenos e pragas. 2006.
- MARCHLER-BAUER, Aron et al. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. **Nucleic acids research**, v. 30, n. 1, p. 281-283, 2002. <https://doi.org/10.1093/nar/30.1.281>.
- MCDOWALL, Jennifer; HUNTER, Sarah. InterPro protein classification. **Bioinformatics for Comparative Proteomics**, p. 37-47, 2011. https://doi.org/10.1007/978-1-60761-977-2_3.
- MCKINNEY, Wes. **Python for data analysis: Data wrangling with Pandas, NumPy, and IPython**. " O'Reilly Media, Inc.", 2012.
- MIRDITA, Milot et al. ColabFold: making protein folding accessible to all. *Nature Methods*, p. 1-4, 2022. <https://doi.org/10.1038/s41592-022-01488-1>.
- MOLESINI, Barbara et al. Plant cystine-knot peptides: pharmacological perspectives. **British journal of clinical pharmacology**, v. 83, n. 1, p. 63-70, 2017. <https://doi.org/10.1111/bcp.12932>.
- MOYER, Tessa B.; BRECHBILL, Amanda M.; HICKS, Leslie M. Mass Spectrometric Identification of Antimicrobial Peptides from Medicinal Seeds. **Molecules**, v. 26, n. 23, p. 7304, 2021. <https://doi.org/10.3390/molecules26237304>.
- MOYER, Tessa B. et al. Multiple classes of antimicrobial peptides in *Amaranthus tricolor* revealed by prediction, proteomics, and mass spectrometric characterization. **Journal of**

natural products, v. 84, n. 2, p. 444-452, 2021. <https://doi.org/10.1021/acs.jnatprod.0c01203>.

NAWROT, Robert et al. Plant antimicrobial peptides. **Folia microbiologica**, v. 59, n. 3, p. 181-196, 2014. <https://doi.org/10.1007/s12223-013-0280-4>.

ODINTSOVA, Tatyana I.; SLEZINA, Marina P.; ISTOMINA, Ekaterina A. Defensins of grasses: a systematic review. **Biomolecules**, v. 10, n. 7, p. 1029, 2020. <https://doi.org/10.3390/biom10071029>.

OLGA, Kulaeva et al. The role of plant antimicrobial peptides (AMPs) in response to biotic and abiotic environmental factors. **Biological Communications**, v. 65, n. 2, p. 187-199, 2020.

O'NEIL, Shawn T. **A Primer for Computational Biology**. Oregon State University Press, 2017.

PARISI, Kathy et al. The evolution, function and mechanisms of action for plant defensins. In: Seminars in cell & developmental biology. **Academic Press**, 2019. p. 107-118. <https://doi.org/10.1016/j.semcdb.2018.02.004>.

PARK, Sungkyu et al. Cyclotide evolution: insights from the analyses of their precursor sequences, structures and distribution in violets (*Viola*). **Frontiers in plant science**, v. 8, p. 2058, 2017. <https://doi.org/10.3389/fpls.2017.02058>.

PELEGRINI, Patrícia B.; FRANCO, Octávio L. Plant γ -thionins: novel insights on the mechanism of action of a multi-functional class of defense proteins. **The international journal of biochemistry & cell biology**, v. 37, n. 11, p. 2239-2253, 2005. <https://doi.org/10.1016/j.biocel.2005.06.011>.

PETTERSEN, Eric F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. **Protein science**, v. 30, n. 1, p. 70-82, 2021. <https://doi.org/10.1002/pro.3943>.

PESTANA-CALSA, Maria Clara; CALSA JR, Tercilio. In silico identification of plant-derived antimicrobial peptides. **Systems and computational biology—molecular and cellular experimental systems**. Rijeka, Croatia: InTech, p. 249-272, 2011.

PEVZNER, Pavel; SHAMIR, Ron (Ed.). Bioinformatics for biologists. **Cambridge University Press**, 2011.

PIRTSKHALAVA, Malak et al. Physicochemical Features and Peculiarities of Interaction of AMP with the Membrane. **Pharmaceuticals**, v. 14, n. 5, p. 471, 2021. <https://doi.org/10.3390/ph14050471>.

PORTO, William F.; FRANCO, Octavio L. Theoretical structural insights into the snakain/GASA family. **Peptides**, v. 44, p. 163-167, 2013. <https://doi.org/10.1016/j.peptides.2013.03.014>.

PORTO, William F. et al. In silico identification of novel hevein-like peptide precursors. **Peptides**, v. 38, n. 1, p. 127-136, 2012. <https://doi.org/10.1016/j.peptides.2012.07.025>.

QUINTANS, Isadora Louise Alves da Costa Ribeiro et al. An Overview of databases and bioinformatics tools for plant antimicrobial peptides. **Current Protein and Peptide Science**, v. 23, n. 1, p. 6-19, 2022. <https://doi.org/10.2174/1389203723666211222170342>.

QUINTANS, I. L. A. C. R. et al. **PlantAMP - A database for plant antimicrobial peptides**. In: CONGRESSO BRASILEIRO DE GENÉTICA, 68., 2023, Ouro Preto. Anais [...]. Ouro Preto, 2023. p. 585. Disponível em: https://www.sbg.org.br/admin/files/book/book_GGbabDgnlbu3.pdf. Acesso em: 26 fev. 2024.

QUITANS, Ribeiro; DA COSTA, Isadora Louise Alves. **Molecular responses of plants to symbiotic fungi**. 2019. Tese de Doutorado. University of British Columbia. <https://doi.org/10.14288/1.0384823>.

RAMAZI, Shahin et al. A review on antimicrobial peptides databases and the computational tools. **Database**, v. 2022, p. baac011, 2022. <https://doi.org/10.1093/database/baac011>.

RANGANATHAN, Shoba; NAKAI, Kenta; SCHONBACH, Christian. Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics. **Elsevier**, 2018.

RHEE, Seung Yon; DICKERSON, Julie; XU, Dong. Bioinformatics and its applications in plant biology. **Annual review of plant biology**, v. 57, n. 1, p. 335-360, 2006. <https://doi.org/10.1146/annurev.arplant.56.032604.144103>.

RODRIGUEZ-LEYVA, Delfin et al. Potent antihypertensive action of dietary flaxseed in hypertensive patients. **Hypertension**, v. 62, n. 6, p. 1081-1089, 2013. <https://doi.org/10.1161/HYPERTENSIONAHA.113.02094>.

SALEEM, Muhammad Hamzah et al. Flax (*Linum usitatissimum* L.): A potential candidate for phytoremediation? Biological and economical points of view. **Plants**, v. 9, n. 4, p. 496, 2020. <https://doi.org/10.3390/plants9040496>.

SALMINEN, Tiina A.; BLOMQVIST, Kristina; EDQVIST, Johan. Lipid transfer proteins: classification, nomenclature, structure, and function. **Planta**, v. 244, p. 971-997, 2016. <https://doi.org/10.1007/s00425-016-2585-4>.

SANTOS-SILVA, Carlos André et al. Plant antimicrobial peptides: state of the art, in silico prediction and perspectives in the omics era. **Bioinformatics and Biology Insights**, v. 14, 2020. <https://doi.org/10.1177/1177932220952739>.

SHELENKOV, Andrey; SLAVOKHOTOVA, Anna; ODINTSOVA, Tatyana. Predicting antimicrobial and other cysteine-rich peptides in 1267 plant transcriptomes. **Antibiotics**, v. 9, n. 2, p. 60, 2020. <https://doi.org/10.3390/antibiotics9020060>.

SLAVOKHOTOVA, A. A. et al. Hevein-like antimicrobial peptides of plants. **Biochemistry (Moscow)**, v. 82, n. 13, p. 1659-1674, 2017. <https://doi.org/10.1134/S0006297917130065>.

SLAVOKHOTOVA, Anna A.; ROGOZHIN, Eugene A. Defense peptides from the α -hairpinin family are components of plant innate immunity. **Frontiers in Plant Science**, v. 11, p. 465, 2020. <https://doi.org/10.3389/fpls.2020.00465>.

SOLTANIAN, Nouredin; JANGHORBANI, Mohsen. A randomized trial of the effects of flaxseed to manage constipation, weight, glycemia, and lipids in constipated patients with type 2 diabetes. **Nutrition & metabolism**, v. 15, n. 1, p. 1-9, 2018. <https://doi.org/10.1186/s12986-018-0273-z>.

STAJICH, Jason E. et al. The Bioperl toolkit: Perl modules for the life sciences. **Genome research**, v. 12, n. 10, p. 1611-1618, 2002. <https://doi.org/10.1101/gr.361602>.

SU, Qiang; WANG, Ke; ZHANG, Zengyan. Ecotopic expression of the antimicrobial peptide DmAMP1W improves resistance of transgenic wheat to two diseases: Sharp eyespot and common root rot. **International journal of molecular sciences**, v. 21, n. 2, p. 647, 2020. <https://doi.org/10.3390/ijms21020647>.

TAM, James P. et al. Antimicrobial peptides from plants. **Pharmaceuticals**, v. 8, n. 4, p. 711-757, 2015. <https://doi.org/10.3390/ph8040711>.

TANTONG, Supaluk et al. Two novel antimicrobial defensins from rice identified by gene coexpression network analyses. **Peptides**, v. 84, p. 7-16, 2016. <https://doi.org/10.1016/j.peptides.2016.07.005>.

TERRAS, F. R et al. Small cysteine-rich antifungal proteins from radish: their role in host defense. **The plant cell**, v. 7, n. 5, p. 573-588, 1995. <https://doi.org/10.1105/tpc.7.5.573>.

TORRENT, Marc; VICTORIA NOGUES, M.; BOIX, Ester. Discovering new in silico tools for antimicrobial peptide prediction. **Current drug targets**, v. 13, n. 9, p. 1148-1157, 2012. <https://doi.org/10.2174/138945012802002311>.

TOKMAKOV, Alexander A.; KUROTANI, Atsushi; SATO, Ken-Ichi. Protein pI and intracellular localization. **Frontiers in Molecular Biosciences**, v. 8, p. 775736, 2021. <https://doi.org/10.3389/fmolb.2021.775736>.

VAN ROSSUM, Guido et al. **Python reference manual**. Amsterdam: Centrum voor Wiskunde en Informatica, 1995.

VERMA, Satish K. et al. Fungal disease prevention in seedlings of rice (*Oryza sativa*) and other grasses by growth-promoting seed-associated endophytic bacteria from invasive *Phragmites australis*. **Microorganisms**, v. 6, n. 1, p. 21, 2018. <https://doi.org/10.3390/microorganisms6010021>.

WAGHU, Faiza Hanif et al. CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides. **Nucleic acids research**, v. 44, n. D1, p. D1094-D1097, 2016. <https://doi.org/10.1093/nar/gkv1051>.

WANG, Jiyao et al. The conserved domain database in 2023. **Nucleic Acids Research**, v. 51, n. D1, p. D384-D388, 2023. <https://doi.org/10.1093/nar/gkac1096>.

WANG, Weixuan et al. Rice secondary metabolites: structures, roles, biosynthesis, and metabolic regulation. **Molecules**, v. 23, n. 12, p. 3098, 2018. <https://doi.org/10.3390/molecules23123098>.

WAR, Abdul Rashid et al. Mechanisms of plant defense against insect herbivores. **Plant signaling & behavior**, v. 7, n. 10, p. 1306-1320, 2012. <https://doi.org/10.4161/psb.21663>.

ZHAO, Mingchao; LIN, Yongjun; CHEN, Hao. Improving nutritional quality of rice for human health. **Theoretical and Applied Genetics**, v. 133, n. 5, p. 1397-1413, 2020. <https://doi.org/10.1007/s00122-019-03530-x>.