



UNIVERSIDADE DO ESTADO DA BAHIA
DEPARTAMENTO DE TECNOLOGIA E CIÊNCIAS SOCIAIS –
DTCS -CAMPUS III - JUAZEIRO
COLEGIADO DE ENGENHARIA DE BIOPROCESSOS E
BIOTECNOLOGIA

PROSPECÇÃO E CARACTERIZAÇÃO *in silico* DE
PEPTÍDEOS DO TIPO TAUMATINA NA MANIÇOBA
(*Manihot pseudoglaziovii*) PARA APLICAÇÕES
BIOTECNOLÓGICAS

Discente: GABRIEL ANTUNES LIMA

Orientador: Dr. IGOR COSTA DE AMORIM

JUAZEIRO-BA

JULHO, 2025

GABRIEL ANTUNES LIMA

**PROSPECÇÃO E CARACTERIZAÇÃO *in silico* DE PEPTÍDEOS
DO TIPO TAUMATINA NA MANIÇOBA (*Manihot pseudoglaziovii*)
PARA APLICAÇÕES BIOTECNOLÓGICAS**

Trabalho de conclusão de curso apresentado ao Colegiado de Engenharia de Bioprocessos e Biotecnologia da Universidade do Estado da Bahia – UNEB Campus III, como requisito parcial para avaliação da disciplina de Trabalho de Conclusão do Curso de Engenharia de Bioprocessos e Biotecnologia.

Orientador: Prof. Dr. Igor Costa de Amorim

JUAZEIRO-BA

JULHO, 2025


GABRIEL ANTUNES LIMA

PROSPECÇÃO E CARACTERIZAÇÃO *in silico* DE PEPTÍDEOS DO TIPO TAUMATINA NA MANIÇOBA (*Manihot pseudoglaziovii*) PARA APLICAÇÕES BIOTECNOLÓGICAS


Trabalho de conclusão de curso aprovado como requisito parcial para a obtenção do título de Bacharel em Engenharia de Bioprocessos e Biotecnologia, pelo Departamento de Tecnologia e Ciências Sociais, Campus III, da Universidade do Estado da Bahia.

Aprovado em 10/07/2025


BANCA EXAMINADORA

Documento assinado digitalmente
 **IGOR COSTA DE AMORIM**
Data: 01/09/2025 16:20:24-0300
Verifique em <https://validar.iti.gov.br>

Dr. Igor Costa de Amorim
Universidade do Estado da Bahia – UNEB

Documento assinado digitalmente
 **LUANA PEREIRA GONCALVES**
Data: 01/09/2025 16:28:10-0300
Verifique em <https://validar.iti.gov.br>

Dra. Luana Pereira Gonçalves
Universidade do Estado da Bahia – UNEB

Documento assinado digitalmente
 **JOAO PACIFICO BEZERRA NETO**
Data: 01/09/2025 19:05:39-0300
Verifique em <https://validar.iti.gov.br>

Dr. João Pacifico Bezerra Neto
Universidade de Pernambuco – UPE

JUAZEIRO-BA

AGOSTO, 2025

AGRADECIMENTOS

Gostaria de dedicar um agradecimento a todos que fizeram parte da minha trajetória nessa etapa da minha vida. Em especial, agradeço a minha mãe, por servir de fonte de inspiração para mim e por sempre ter feito o melhor por mim, apesar de todas as dificuldades que encontramos. Agradeço também os meus irmãos e ao meu pai, por sempre estarem ao meu lado em tudo na minha vida e por terem moldado a pessoa que sou hoje.

Sou igualmente grato a todos os amigos e colegas da faculdade, por me ajudarem a ter uma rotina mais leve e pelo apoio em momentos que precisei.

Também registro minha profunda gratidão ao meu orientador, prof. Igor Amorim, cuja dedicação, paciência e ajuda foram fundamentais para a realização desse trabalho.

Por fim, dedico um agradecimento a todos que de alguma forma contribuíram para a realização desse trabalho e para a minha formação pessoal e profissional.

SUMÁRIO

1. INTRODUÇÃO	6
2. OBJETIVOS	7
2.1 Geral	7
2.2 Específicos	7
3. REFERENCIAL TEÓRICO	7
3.1 Maniçoba (<i>Manihot pseudoglaziovii</i> Pax & K. Hoffm.)	7
3.2 Proteínas relacionadas a patogênese em plantas	8
3.3 Potencial biotecnológico das TLPs	9
3.4 Papel da bioinformática na prospecção de TLPs	10
4. MATERIAL E MÉTODOS	10
5. RESULTADOS E DISCUSSÃO	13
6. CONSIDERAÇÕES FINAIS	21
REFERÊNCIAS BIBLIOGRÁFICAS	22

RESUMO

A caatinga é um bioma brasileiro que abriga plantas nativas com potencial biotecnológico pouco explorado, incluindo a maniçoba (*Manihot pseudoglaziovii* Pax & K. Hoffm.), sendo esta frequentemente utilizada como forrageira. As proteínas semelhantes à taumatina (TLPs) são proteínas expressas por plantas em casos de estresse biótico que vêm sendo alvo de interesse biotecnológico por suas funções antimicrobianas. Este trabalho tem o objetivo de prospectar e caracterizar TLPs a partir de dados de sequenciamento da maniçoba. Para identificar as sequências de potenciais TLPs, foi construída uma base de dados contendo sequências de TLPs a partir da base de dados do UNIPROT. Em seguida, o tblastn foi utilizado para obter as sequências codificantes de proteínas similares às da base de dados construída. As fases de leitura aberta foram obtidas a partir da ferramenta getorf. As potenciais TLPs foram caracterizadas por meio de diferentes ferramentas de bioinformática para a identificação de domínios conservados, previsão de peptídeo sinal, localização subcelular, predição da atividade antimicrobiana, predição de toxicidade, análise de propriedades físico-químicas e modelagem tridimensional. Os resultados demonstraram que 30 sequências apresentaram pelo menos 15 cisteínas em sua composição, com 10 delas apresentando oito ligações dissulfeto. Dessas TLPs, todas apresentaram atividade antimicrobiana, 23 proteínas apresentaram natureza ácida e sete natureza básica. Além disso, 13 proteínas foram classificadas como estáveis, oito apresentaram probabilidade de presença de peptídeo sinal do tipo Sec/SPI e 12 TLPs foram classificadas como extracelulares. A busca de domínios conservados revelou que 28 proteínas apresentaram domínios específicos, incluindo taumatina, TLP-PA e THN, duas sequências apresentaram domínio da superfamília GH64-TLP-SF. Em relação à toxicidade, seis TLPs não apresentaram toxicidade. A modelagem tridimensional das TLPs revelou a presença de estruturas características nas TLPs avaliadas. No final, considerando todos os parâmetros, foram identificadas três possíveis TLPs candidatas para uso como antimicrobiano. Dessa maneira, os resultados obtidos reforçam o potencial da maniçoba como fonte de moléculas bioativas com potencial biotecnológico, mas que necessitam de estudos posteriores para a validação de sua atividade.

Palavras-chave: Bioprospecção; Bioinformática; Peptídeos antimicrobianos

ABSTRACT

The Caatinga is a Brazilian biome that is home to native plants with underexplored biotechnological potential, including maniçoba (*Manihot Pseudoglaziovii* Pax & K. Hoffm.), which is often used as fodder. Thaumatin-like proteins (TLPs) are proteins expressed by plants in case of biotic stress that have been the subject of biotechnological interest due to their antimicrobial functions. This study aims to prospect and characterize TLPs from maniçoba sequencing data. To identify the sequences of potential TLPs, a database containing TLP sequences was constructed from the UNIPROT database. Next, tblastn was used to obtain the protein-coding sequences similar to those in the constructed database. The open reading frames were obtained using the getorf tool. Potential TLPs were characterized using different bioinformatics tools to identify conserved domains, predict signal peptides, subcellular localization, antimicrobial activity, toxicity, analyze physicochemical properties, and perform three-dimensional modeling. The results showed that 30 sequences had at least 15 cysteines in their composition, with 10 of them having eight disulfide bonds. Of these TLPs, all showed antimicrobial activity, 23 proteins were acidic in nature, and seven were basic in nature. In addition, 13 proteins were classified as stable, eight showed a probability of Sec/SPI-type signal peptide presence, and 12 TLPs were classified as extracellular. The search for conserved domains revealed that 28 proteins had specific domains, including thaumatin, TLP-PA, and THN, and two sequences had the GH64-TLP-SF superfamily domain. Regarding toxicity, six TLPs showed no toxicity. Three-dimensional modeling of the TLPs revealed the presence of characteristic structures in the TLPs evaluated. In the end, considering all parameters, three possible TLP candidates were identified for use as antimicrobials. Thus, the results obtained reinforce the potential of maniçoba as a source of bioactive molecules with biotechnological potential, but further studies are needed to validate their activity.

Keywords: Bioprospecting; Bioinformatics; Antimicrobial Peptides

LISTA DE TABELAS

Tabela 1 – Códigos de acesso de sequências obtidas no UNIPROT para montagem de banco de dado local.	11
Tabela 2 – caracterização físico-química das sequências potenciais de TLPs. Sigla das abreviações dos parâmetros: PM = peso molecular; PI = ponto isoelétrico; INS = índice de instabilidade; CIS = número de cisteínas.....	13
Tabela 3 – Resultados obtidos no Batch CD-Search para a identificação de domínios conservados nas proteínas analisadas.	15
Tabela 4 – Predição de atividade antimicrobiana das sequências de potenciais TLPs identificadas no genoma de <i>Manihot Pseudoglaziovii</i> . Sigla das abreviações dos algoritmos: SVM = Support Vector Machine; RFC = Random Forest Classifier; ANN = Artificial Neural Network; DAC = Discriminant Analysis Classifier.....	15
Tabela 5 – Análise das sequências de proteínas utilizando o SignalP 6.0 para predição de peptídeos sinal.	16
Tabela 6 – Predição de localização subcelular das potenciais TLPs utilizando a ferramenta BUSCA.	17
Tabela 7 – Predição de toxicidade das TLPs utilizando a ferramenta ToxinPred 3.0...	18
Tabela 8 – Modelos estruturais utilizados na modelagem de TLPs de <i>Manihot pseudoglaziovii</i> e os percentuais de identidade de sequência entre as TLPs e o modelo.	18

LISTA DE FIGURAS

Figura 1 – Modelagem 3D das TLPs utilizando o SWISS-MODEL. As α -hélices estão representadas em azul, as folhas beta em azul e as ligações dissulfeto em vermelho.... 20

Figura 2 – Comparação das TLPs candidatas, classificadas em ordem decrescente de acordo com o número de características desejadas apresentadas (destacadas em verde).
..... 21

1. INTRODUÇÃO

A Caatinga é um domínio fitogeográfico exclusivo do território brasileiro, abrangendo uma área de 912.000 km², sendo localizado quase que totalmente na região nordeste do país (Fernandes; Queiroz, 2018). Esse bioma é caracterizado pelos baixos índices pluviométricos, variando de 250 a 900 mm anuais, além de apresentar, em geral, solos com pouca capacidade de retenção de água, pedregosos e ricos em minerais (Alves; Araújo; Nascimento, 2009). A sua vegetação é composta por plantas xerófilas, predominantemente caracterizadas por árvores baixas, arbustos e cactáceas (Alves; Araújo; Nascimento, 2009).

Dentre as espécies nativas da Caatinga, destaca-se a maniçoba (*Manihot pseudoglaziovii* Pax & K. Hoffm., 1924) que é amplamente utilizada como forrageira, apesar de conter ácido cianídrico, capaz de causar toxicidade nos animais dependendo da quantidade ingerida (Costa et al., 2007).

Diante das condições adversas desse bioma, plantas como a maniçoba desenvolveram mecanismos de defesas eficientes contra estresses bióticos e abióticos, como a produção de proteínas relacionadas à patogênese (PRs), incluindo as proteínas similares à taumatina (TLPs), pertencentes ao grupo PR-5 (Jesús-Pires et al., 2020). Essas proteínas apresentam domínio conservado da taumatina (Pfam: PF00314) em aproximadamente 95% do seu tamanho (Petre et al., 2011). As TLPs são amplamente distribuídas nas plantas, sendo encontradas nas algas verdes, briófitas, gimnospermas e angiospermas e, apresentam funções fisiológicas e antimicrobianas que tem atraído o interesse biotecnológico (Faillace et al., 2019).

Nesse contexto, diversos estudos têm buscado caracterizar essas moléculas, especialmente em espécies adaptadas a ambientes extremos, utilizando diferentes abordagens de análise. A bioinformática tem se consolidado como uma área essencial para a análise de dados biológicos, permitindo a identificação e caracterização de diferentes moléculas, incluindo as TLPs, a partir de dados de sequenciamento genômico. Embora as TLPs já tenham sido estudadas em outras espécies vegetais, ainda existe uma escassez de estudos em plantas nativas da Caatinga, como a maniçoba. Essa abordagem pode revelar novas proteínas com potencial de aplicação e valoriza a exploração da biodiversidade regional. Dessa forma, o presente trabalho busca identificar e caracterizar TLPs através de dados genômicos da maniçoba utilizando ferramentas de bioinformática.

2. OBJETIVOS

2.1 Geral

Identificar e caracterizar TLPs a partir de ferramentas de bioinformática na maniçoba (*Manihot pseudoglaziovii* Pax & K. Hoffm., 1924) com foco em identificação, caracterização estrutural e avaliação do potencial para aplicações biotecnológicas.

2.2 Específicos

- Analisar sequências codificantes para TLPs na espécie *Manihot pseudoglaziovii* Pax & K. Hoffm., 1924;
- Caracterizar *in silico* as propriedades físico-químicas, toxicidade, atividade antimicrobiana e domínios conservados das sequências identificadas;
- Realizar a predição da estrutura tridimensional das sequências identificadas para comparar a similaridade com TLPs já descritos.
- Selecionar sequências com maior potencial para uso como agentes antimicrobianos em aplicações biotecnológicas.

3. REFERENCIAL TEÓRICO

3.1 Maniçoba (*Manihot pseudoglaziovii* Pax & K. Hoffm., 1924)

A família de planta Euphorbiaceae é amplamente distribuída no território brasileiro, com diversas espécies economicamente importantes, que podem ser utilizadas na medicina e na alimentação humana (Nina et al., 2021). Dentre os seus gêneros, *Manihot* se destaca pela capacidade de produzir glicosídeos cianogênicos, que em determinadas condições, pode causar intoxicação se consumida em grandes quantidades (Moraes et al., 2007). Apesar disso, o gênero inclui espécies economicamente importantes, como a *Manihot esculenta* Crantz, popularmente conhecida como mandioca, macaxeira ou aipim, frequentemente utilizada na alimentação humana (Ciarini, 2019). Outra espécie é a *Manihot pseudoglaziovii* Pax & K. Hoffm., 1924, conhecida popularmente como maniçoba, que é utilizada como forragem para alimentação de rebanhos (Ciarini, 2019).

A maniçoba é uma planta nativa da Caatinga, sendo encontrada no Semiárido do nordeste do Brasil (Beltrão et al., 2006). Segundo Beltrão et al. (2006), a maniçoba é uma planta heliófila capaz de crescer em diversos tipos de solos, incluindo solos

calcários, drenados ou pedregosos. A maniçoba, quando submetida ao processo de silagem, apresenta características químicas e bromatológicas para a alimentação de ruminantes, sendo uma alternativa importante em épocas de seca (Silva et al., 2024). Apesar de sua importância ecológica e econômica, especialmente em regiões que apresentam condições adversas, ainda são escassos os estudos moleculares envolvendo a maniçoba, limitando o conhecimento acerca dos mecanismos genéticos envolvidos na resistência aos estresses desse ambiente.

3.2 Proteínas relacionadas a patogênese em plantas

As plantas que vivem em ambientes com condições adversas, como é o caso da maniçoba, desenvolvem mecanismos para a sua sobrevivência. Um deles, as proteínas relacionadas a patogênese (PR), são proteínas produzidas pelas plantas em resposta a ataques causados por patógenos, nematoides, insetos ou herbívoros (Sudisha et al., 2012). Segundo Loon, Rep e Pieterse (2006), as PRs são classificadas em 17 famílias. Dentre esses grupos, as proteínas PR-1 são as mais abundantes, apresentando um percentual de 1% a 2% do total de proteínas da folha, sendo detectadas no tomate, tabaco, milho, cevada e muitas outras plantas (Sudisha et al., 2012). As características das outras famílias incluem as proteínas que apresentam atividade de β -1,3 glucanase (PR-2), as quitinases (PR-3, PR-8 e PR-11), as proteínas Win e heveínas (PR-4), as semelhantes à taumatina (PR-5), os inibidores de proteinases (PR-6), as com atividade endoproteinase (PR-7), as com atividade semelhante à da peroxidase (PR-9), as com atividade e estrutura semelhantes à ribonuclease (PR-10), as defensinas (PR-12), as tioninas (PR-13), as proteínas transportadoras de lipídeos (PR-14), as semelhantes à oxalato-oxidase (PR-15 e PR-16) e as proteínas identificadas na cevada que apresentam domínios semelhantes aos de metaloproteinases (PR-17) (Sudisha et al., 2012).

As moléculas do grupo PR-5 são conhecidas como proteínas semelhantes à taumatina, devido a sua similaridade estrutural à taumatina, proteína com sabor doce e que foi purificada pela primeira vez em *Thaumatococcus danielli* Benth, planta originada na África Ocidental (Jesús-Pires et al., 2020). Apesar de possuírem estrutura semelhante, as TLPs não possuem sabor doce, mas desempenham diferentes funções nas plantas, como a inibição do crescimento de micélios e redução do crescimento dos esporos dos fungos (Nawrot et al., 2021). Em geral, as TLPs apresentam peso molecular variando de 10 a 34 kDa e são compostas por cerca de 200 resíduos de aminoácidos (Smole et al., 2007). Dentre esses, 16 são resíduos de cisteínas que formam oito

ligações dissulfeto as quais conferem alta estabilidade estrutural, tornando-as resistentes à degradação por proteases (Smole et al., 2007), à altas temperaturas e à mudanças no pH (Rydzyńska et al., 2025). Por outro lado, algumas TLPs são pequenas e apresentam apenas 10 resíduos de cisteínas, devido à deleção interna de seu peptídeo, sendo encontradas principalmente em coníferas e cereais (Liu; Sturrock; Ekramoddoullah, 2010).

3.3 Potencial biotecnológico das TLPs

Do ponto de vista biotecnológico, as TLPs podem ser aplicadas na agricultura, devido à atividade antifúngica e à estabilidade que essas proteínas apresentam, sendo promissoras para obter plantas melhoradas capazes de ter uma melhor produção sob estresse biótico ou abiótico (Feng; Wei; Li, 2024). No trabalho de Archarya et al. (2012), plantas transgênicas de batata expressando o gene de uma TLP da planta *Camellia sinensis* apresentaram maior resistência aos fungos *Macrophomina phaseolina* e *Phytophthora infestans*, quando comparadas às plantas selvagens. Por outro lado, as TLPs também podem atuar em situações de estresse abiótico, como demonstrado no trabalho de Li et al (2020), onde plantas que tiveram uma superexpressão de genes da GhTLP19 apresentaram maior resistência à estresse hídrico.

Na indústria de alimentos e na área da saúde, o uso peptídeos antimicrobianos têm despertado interesse como alternativa para o controle de microrganismos. Nesse contexto Viktorova et al. (2017), demonstraram o potencial da osmotina, proteína pertencente à família PR-5, em fungos patogênicos e contaminantes de alimentos. Essa proteína apresentou atividade inibitória contra *Candida parapsilosis*, *Candida tropicalis*, *Debaryomyces hansenii* e *Saccharomyces ludwigii*, sendo os dois primeiros patogênicos e os dois últimos contaminantes de alimentos. Na saúde, também foi observado que uma TLP extraída da maçã *Annurca* apresentou propriedades neuroprotetivas, ao inibir as enzimas acetilcolinesterase e monoamina oxidase A, além de inibir a agregação de peptídeos beta-amiloide em ensaios *in vitro*, indicando que essas proteínas também podem ser aplicadas à saúde humana (D'Errico et al., 2024).

Dessa maneira, as TLPs podem ser moléculas potenciais para atuarem como novos antimicrobianos. Seu uso no combate a microrganismos resistentes é particularmente interessante, uma vez que sua atuação ocorre principalmente na membrana celular o que dificulta o desenvolvimento de resistência (Lima et al., 2021).

Apesar do seu potencial para o desenvolvimento de drogas, o custo de produção e a dificuldade de extração são fatores limitantes para a produção em larga escala (Boto; Lastra; González, 2017). Como alternativa, o desenvolvimento de peptídeos sintéticos por meio de modificações químicas ou através do *design* racional de peptídeos antimicrobianos (AMPs) naturais têm se mostrado uma alternativa promissora. Essas mudanças são realizadas para obter moléculas que são mais fáceis de produzir, que apresentam maior estabilidade, biodisponibilidade, atividade antimicrobiana ou menor toxicidade (Tossi, 2011). Para que essas alternativas sejam desenvolvidas, é fundamental uma compreensão inicial dos peptídeos de interesse. Dessa forma, a obtenção e caracterização das sequências primárias das TLPs é um passo importante para definir sequências com propriedades desejadas, além de permitir a definição de possíveis alvos para modificações. Apesar de demonstrarem potencial biotecnológico, algumas TLPs de plantas já foram descritas como alergênicos, embora seu efeito ainda não seja muito bem compreendido (Rydzyńska et al., 2025).

3.4 Papel da bioinformática na prospecção de TLPs

O avanço nas tecnologias de sequenciamento permitiu a geração de um volume muito grande de dados biológicos, especialmente nas áreas das ciências ômicas. Em função disso, a bioinformática apresenta um papel fundamental para organizar, analisar, manipular e interpretar esses dados, aplicados para solucionar problemas na área da biologia e biotecnologia (Altman; Dugan, 2003).

As técnicas de bioinformática também permitem a prospecção e *design* de TLPs e de outros peptídeos antimicrobianos. As técnicas de triagem *in silico* utilizando dados de genômica, transcriptômica, proteômica e metabolômica tem permitido a identificação e caracterização de genes codificadores de TLPs de maneira mais rápida e precisa, apesar de necessitar validação experimental posterior (Pestana-Calsa; Ribeiro; Calsa, 2010). Essa triagem descarta a necessidade de análises robustas iniciais e torna o processo economicamente mais viável. Por causa disso, diversos estudos têm sido realizados para identificar peptídeos de interesse biotecnológico em espécies adaptadas a ambientes extremos. Entretanto, ainda são escassos os estudos moleculares em plantas da Caatinga, que enfrentam severas condições bióticas e abióticas.

4. MATERIAL E MÉTODOS

Inicialmente foi montado um banco de dado local, com sequências de peptídeos antimicrobianos do tipo taumatina de plantas pertencentes à mesma família taxonômica da maniçoba (tabela 1). Essas sequências foram obtidas no UNIPROT (<https://www.uniprot.org/>). Por sua vez, os dados brutos de sequenciamento genômico de *Manihot pseudoglaziovii* Pax & K. Hoffmn (1924) foram obtidos na base de dados do SRA (*Sequence Read Archive*) do NCBI (*National Center for Biotechnology Information*), utilizando o código de acesso SRR2847457.

Tabela 1 – Códigos de acesso de sequências obtidas no UNIPROT para montagem de banco de dado local.

Espécie	Código de acesso
<i>Hevea brasiliensis</i>	A0A6A6LC04, A0A6A6NHR4, A0A6A6K6R3, A0A6A6LIP4, A0A6A6MRU8, A0A6A6KC45, A0A6A6MUK8, A0A6A6LP26, A0A6A6LCN4, A0A6A6LPM9, A0A0M3TCA9, A0A6A6MLL7, A0A6A6M1R2, A0A6A6M8W8, A0A6A6N5U5, A0A6A6LCE9, A0A6A6LIW0, A0A6A6KP52, A0A6A6NHT0, A0A6A6LPH7, A0A6A6N487, A0A6A6M0V7, A0A6A6M4M3, A0A6A6L5P2, A0A6A6N561, A0A6A6KKS7, A0A6A6N398, A0A6A6KXL3, A0A6A6KXN1, A0A6A6MW98, A0A6A6N6U5, A0A6A6M664, A0A6A6N386, A0A6A6L659, A0A6A6LDP9, A0A6A6NH82, A0A6A6KM49, A0A6A6L9W1, A0A6A6KX94, A0A6A6L4M8, A0A6A6NH86, A0A6A6M8R0, A0A6A6KTZ8, A0A6A6NDK2, A0A6A6MWC0, A0A6A6L4P7, A0A6A6MBU2 e A0A6A6MJ05
<i>Jatropha curca</i>	A0A067KWH8, A0A067KAB5, A0A067K080, A0A067KXA9, A0A067KCR4, A0A067KYH3, A0A067LE89, A0A067JPQ1, A0A067LPG9, A0A067KRY0, A0A067KLV1, A0A067LL33, A0A067LJS7, A0A067LAK9, A0A067JWI9, A0A067KBM3, A0A067KK59, A0A067KNP2, A0A067KQP6, A0A067KLF3, A0A067LKF6, A0A067JNB0, A0A067L4U0, A0A067JSE7, A0A067LH15, A0A067JRM3, A0A067KXG4, A0A067KKG0, A0A067KK54.
<i>Manihot esculenta</i>	A0A2C9VW80, A0A2C9UZW3, A0A2C9VYH8, A0A2C9TZV0, A0A2C9VHM4, A0A2C9VYR0, A0A2C9WM54, A0A2C9TZQ1, A0A2C9WD37, A0A2C9WI05, A0A2C9WCS1, A0A2C9V7X0, A0A2C9WMA8, A0A2C9WED2, A0A2C9WV7, A0A2C9UHY3, A0A2C9UMU2, A0A2C9V239, A0A2C9W2L5, A0A2C9UMU9, A0A2C9VM36, A0A2C9UUA1, A0A2C9V999, A0A2C9VLT0, A0A2C9WIE5, A0A2C9TZV2, A0A2C9WH09, A0A2C9VCH4, A0A2C9UTF2, A0A199UB31, A0A2C9WBS9, A0A2C9WBU3, A0A2C9WIE6 e A0A2C9WAB5.
<i>Ricinus communis</i>	B9RJ48, B9T7D5, B9R8A6, B9S0V1, B9R8A5, B9SUI9, B9T6D4, B9SRI3, B9T382, B9SDY3, B9RUJ5, B9R7P7, B9SZE5, B9SZY6, B9SLL4, B9S0C7, B9SLM3, B9RHG5, B9S0C6, B9RC64, B9RC66, B9RM54, B9SLM1 e B9RHJ7.
<i>Vernicia fordii</i>	A0A1Z1G639

Fonte: elaborado pelo autor (2025).

Ao longo do trabalho, muitas análises foram realizadas na plataforma Galaxy Europe (Afgan et al., 2018), tais como visualização da qualidade do sequenciamento no *FastQC* (Andrews, 2010) e tratamento dos dados com qualidade inferior a 20 (Q20) no *Trimmomatic* (Bolger; Lohse; Usadel, 2014). A ferramenta *Fastq interlacer* foi utilizada para intercalar as *reads* pareadas, unindo as sequências das extremidades opostas dos diferentes fragmentos da biblioteca. O arquivo resultante foi convertido no formato FASTA, com auxílio da ferramenta FASTQ to FASTA converter. Esse arquivo foi analisado no tblastn, utilizando como referência (*seed*) o banco de dados local previamente obtido, para identificar potenciais genes de TLPs na maniçoba. Com base no resultado, a ferramenta *Filter Sequence by ID* foi aplicada para filtrar e selecionar as *reads* com potenciais genes de TLPs do genoma. Em seguida, essas *reads* foram montadas em *contigs* com o auxílio do CAP3. As análises posteriores foram realizadas fora do ambiente do Galaxy Europe. A ferramenta getorf, do pacote EMBOSS (*European Molecular Biology Open Software Suite*) (Rice; Longden; Bleasby, 2000), foi utilizada para obter as matrizes de leitura aberta (ORF) nos *contigs* obtidos. As sequências foram nomeadas no formato TLPX_Y, em que X corresponde ao número do *contig* e Y ao número da ORF identificada pelo getorf.

A caracterização físico-química dos peptídeos potenciais foi realizada na ferramenta ProtParam (Gasteiger et al., 2005). Com base nos resultados, foi realizada uma triagem inicial das potenciais TLPs, considerando como critério o número mínimo de 15 resíduos de cisteínas, em referência ao padrão conservado de 16 cisteínas das TLPs (Rydzynska et al., 2025). A predição das ligações dissulfeto entre as cisteínas das TLPs potenciais foi realizada através da ferramenta CYS_REC (http://www.softberry.com/berry.phtml?topic=cys_rec&group=programs&subgroup=pr opt). Em seguida, o ToxinPred 3.0 (Rathore et al., 2024) e o Batch CD-Search (Wang et al., 2023) foram utilizados para predizer toxicidade e realizar busca de domínios proteicos conservados, respectivamente.

A atividade antimicrobiana foi predita utilizando o CAMPr3 (Waghu et al., 2016), a partir de quatro algoritmos: *Support Vector Machine* (SVM), *Random Forest Classifier* (RFC), *Artificial Neural Network* (ANN) e *Discriminant Analysis Classifier* (DAC). As sequências candidatas a TLP foram caracterizadas utilizando o SignalP 6.0 (Teufel et al., 2022) para verificar a presença e localização de peptídeo sinal, juntamente

com a ferramenta BUSCA (<https://busca.biocomp.unibo.it/>) (Savojardo et al., 2018) para localização subcelular do possível peptídeo. A modelagem dos peptídeos candidatos foi realizada através do SWISS-MODEL (Waterhouse et al., 2018) e os modelos foram visualizados através do software Pymol (Delano, 2002).

5. RESULTADOS E DISCUSSÃO

O *tblastn* indicou que 67.513 *reads* dos dados genômicos brutos de *M. pseudoglaziovii* apresentaram similaridade com genes codificadores de TLPs, com *e-value* menor que 0,001. A montagem dessas sequências no CAP3 resultou em 151 *contigs*, os quais possuem 1.178 ORFs identificadas pelo *getorf*. Após a caracterização físico-química dos peptídeos no ProtParam e triagem pelo número de 15 resíduos de cisteínas, foram selecionadas 30 possíveis sequências de TLPs (tabela 2). O peso molecular dessas sequências apresentou valores compatíveis com aos descrito na literatura (10 a 34 kDa) (Rydzyńska et al., 2025).

Com base no ponto isoelétrico teórico, 23 sequências apresentaram natureza ácida e 7 com natureza básica (tabela 2). A presença de uma superfície carregada negativamente nas TLPs ácidas pode facilitar a atividade antifúngica através da interação da superfície da TLP com cátions inorgânicos (Tachi et al., 2009), os quais são capazes de aumentar a afinidade das TLPs com a parede celular dos fungos (Salzman et al., 2004). Além disso, 13 proteínas demonstraram índice de instabilidade abaixo de 40, sendo classificadas como estáveis (tabela 2). A presença de TLPs estáveis é importante, pois um dos principais desafios do desenvolvimento de produtos à base de peptídeos antimicrobianos é sua instabilidade (Marr; Gooderham; Hancock, 2006). De acordo com os resultados do CYS_REC (tabela 2), apenas 10 proteínas apresentaram o número característico das TLPs de oito ligações dissulfeto.

Tabela 2 – caracterização físico-química das sequências potenciais de TLPs. Sigla das abreviações dos parâmetros: PM = peso molecular; PI = ponto isoelétrico; INS = índice de instabilidade; CIS = número de cisteínas.

Proteína	Comprimento	PM (kDa)	PI	INS	CIS	Ligações dissulfeto
TLP1_16	248	26,41	7,48	33,02	16	7
TLP3_17	242	26,50	4,82	50,03	15	6
TLP4_7	253	26,65	4,45	51,76	16	7
TLP5_11	238	25,20	4,88	39,90	19	8
TLP21_9	231	24,36	7,52	40,08	17	8
TLP30_7	242	26,73	4,89	36,74	16	7
TLP31_18	264	27,83	4,58	46,19	16	7

TLP35_6	227	23,75	5,40	37,96	16	7
TLP37_12	240	25,99	5,40	22,12	16	7
TLP40_7	237	24,51	4,53	43,68	17	8
TLP41_17	230	24,53	8,92	48,47	16	7
TLP43_6	245	25,56	4,48	39,82	17	8
TLP51_17	254	27,50	5,19	45,72	17	6
TLP55_6	241	26,04	8,73	50,10	17	8
TLP78_16	242	25,70	8,90	43,62	16	7
TLP80_5	223	23,35	4,94	42,59	16	8
TLP85_17	247	26,17	6,10	39,23	16	7
TLP90_14	225	24,53	6,36	31,67	16	7
TLP97_13	244	25,45	4,37	46,99	17	8
TLP98_5	236	25,83	6,33	33,51	16	7
TLP101_6	242	25,48	4,91	46,54	17	7
TLP103_6	230	24,78	7,47	30,00	16	7
TLP106_5	247	25,77	4,40	42,55	16	8
TLP109_15	242	26,08	4,60	43,28	16	7
TLP112_14	223	23,39	4,73	39,54	16	7
TLP119_7	236	24,97	4,69	40,47	16	8
TLP120_19	263	28,89	6,31	59,49	17	7
TLP126_7	273	29,68	7,95	51,65	17	7
TLP136_15	237	25,76	6,45	36,42	17	7
TLP143_20	244	25,43	4,69	39,70	16	8

Fonte: elaborado pelo autor (2025).

A busca de domínios conservados, realizada com CD-Search, revelou que 28 proteínas apresentaram domínios específicos (tabela 3), como taumatina (pfam00314), TLP-PA (cd09218) e THN (smart00205). A presença do domínio da taumatina está relacionado com a atividade antifúngica, enquanto o domínio TLP-PA, que é um domínio semelhante ao domínio da taumatina, está presente em proteínas alergênicas e antifúngicas (Alves; Souza; Silva, 2024). Além disso, observou-se que duas sequências apresentaram domínios da superfamília GH64-TLP-SF (tabela 3), que incluem os glicosídeos hidrolases (Alves; Souza; Silva, 2024). De acordo com os critérios do CD-Search, os domínios classificados como *specific hits* indicam uma alta confiança na anotação funcional baseado em modelos curados pelo NCBI, enquanto os domínios de superfamília se referem a um agrupamento amplo de modelos que representam fragmentos proteicos homólogos, apresentando menor especificidade (Marchler-Bauer et al., 2010).

Tabela 3 – Resultados obtidos no Batch CD-Search para a identificação de domínios conservados nas proteínas analisadas.

Domínio	Tipo	Código de acesso	Proteínas
Thaumatococcus	Específico	pfam00314	TLP1_16, TLP3_17, TLP4_7, TLP5_11, TLP21_9, TLP30_7, TLP31_18, TLP43_6, TLP51_17, TLP97_13, TLP106_5, e TLP109_15
TLP-PA	Específico	cd09218	TLP35_6, TLP40_7, TLP41_17, TLP55_6, TLP78_16, TLP80_5, TLP85_17, TLP101_6, TLP103_6, TLP112_14, TLP119_7, TLP120_19, TLP126_7, TLP136_15 e TLP143_20
THN	Específico	smart00205	TLP90_14
GH64-TLP-SF superfamily	Superfamília	cl02511	TLP37_12, TLP98_5

Fonte: elaborado pelo autor (2025).

A partir dos dados do CAMPr3, todas as proteínas foram classificadas como peptídeos antimicrobianos pelos quatro algoritmos avaliados (Tabela 4). Esses resultados sugerem que as TLPs avaliadas podem atuar como peptídeos antimicrobianos. Apesar de não existirem estudos investigando propriedades antifúngicas de *M. pseudoglaziovii*, outras plantas do gênero *Manihot* apresentam propriedades antifúngicas descritas, como *M. esculenta* (Ehiobu; Ogu, 2018), *M. glaziovii* (Pereira et al., 1999) e *M. multifida* (Fabri et al., 2015). Esses resultados demonstram a necessidade de estudos experimentais para confirmar as propriedades antifúngicas das TLPs da maniçoba.

Tabela 4 – Predição de atividade antimicrobiana das sequências de potenciais TLPs identificadas no genoma de *Manihot Pseudoglaziovii*. Sigla das abreviações dos algoritmos: SVM = Support Vector Machine; RFC = Random Forest Classifier; ANN = Artificial Neural Network; DAC = Discriminant Analysis Classifier.

Proteína	SVM	RFC	ANN	DAC
TLP1_16	1,000	0,9315	AMP	1,000
TLP3_17	1,000	0,8810	AMP	1,000
TLP4_7	1,000	0,9320	AMP	1,000
TLP5_11	1,000	0,8400	AMP	1,000
TLP21_9	1,000	0,9255	AMP	1,000
TLP30_7	1,000	0,9185	AMP	1,000
TLP31_18	1,000	0,9370	AMP	1,000
TLP35_6	1,000	0,9415	AMP	1,000
TLP37_12	1,000	0,9485	AMP	1,000
TLP40_7	1,000	0,9335	AMP	1,000
TLP41_17	1,000	0,9250	AMP	1,000

TLP43_6	1,000	0,9495	AMP	1,000
TLP51_17	1,000	0,9210	AMP	1,000
TLP55_6	1,000	0,9005	AMP	1,000
TLP78_16	1,000	0,9325	AMP	1,000
TLP80_5	1,000	0,9700	AMP	1,000
TLP85_17	1,000	0,9390	AMP	1,000
TLP90_14	1,000	0,9360	AMP	1,000
TLP97_13	1,000	0,9375	AMP	1,000
TLP98_5	1,000	0,9565	AMP	1,000
TLP101_6	1,000	0,9285	AMP	1,000
TLP103_6	1,000	0,9420	AMP	1,000
TLP106_5	1,000	0,9285	AMP	1,000
TLP109_15	1,000	0,9515	AMP	1,000
TLP112_14	1,000	0,9475	AMP	1,000
TLP119_7	1,000	0,9605	AMP	1,000
TLP120_19	1,000	0,8795	AMP	1,000
TLP126_7	1,000	0,9395	AMP	1,000
TLP136_15	1,000	0,8810	AMP	1,000
TLP143_20	1,000	0,9415	AMP	1,000

Fonte: elaborado pelo autor (2025).

A predição de peptídeo sinal nas possíveis TLPs, verificada no SignalP 6.0, revelou que oito sequências apresentaram alta probabilidade de conter peptídeo sinal ($\geq 0,90$) do tipo Sec/SPI (tabela 5). Adicionalmente, outra sequência (TLP3_17) apresentou probabilidade de 0,6024, indicando a possível presença do peptídeo sinal Sec/SPI. A presença desse peptídeo sinal sugere que essas proteínas são secretadas por meio da via Sec, um mecanismo de translocação pós-traducional conservado em todas as formas de vida, em que as proteínas são mantidas desdobradas para serem reconhecidas e translocadas através da maquinaria celular responsável pela secreção (Owji et al., 2018).

A presença desse peptídeo sinal compatível com a via Sec pode representar uma vantagem para a produção biotecnológica de TLPs de *M. pseudoglaziovii*, uma vez que essa via reduz tempo de contato da cadeia proteica com proteases intracelulares (Mordkovich; Okorokova; Veiko, 2015).

Tabela 5 – Análise das sequências de proteínas utilizando o SignalP 6.0 para predição de peptídeos sinal.

Proteína	Probabilidade
TLP3_17	0,6024
TLP4_7	0,9755
TLP37_12	0,9957
TLP51_17	0,9998
TLP90_14	0,9997

TLP98_5	0,9995
TLP101_6	0,9995
TLP126_7	0,9124

Fonte: elaborado pelo autor (2025).

A predição da localização subcelular por homologia das TLPs avaliadas indicou que 12 delas são extracelulares, 14 estão localizadas na membrana plasmática e quatro no núcleo (tabela 6). Esses resultados, em conjunto com os obtidos na predição de peptídeo sinal, reforçam a natureza extracelular das proteínas TLP 4_7, TLP37_12 e TLP51_17. Nesse sentido, as TLPs extracelulares são importantes na defesa contra os fungos fitopatogênicos, que infectam através da penetração pela membrana ou parede celular das células vegetais suprimindo o sistema de defesa da planta (Munis et al., 2010). Por outro lado, a presença das TLPs na membrana plasmática e no meio intercelular pode estar relacionada à resistência à seca em plantas sob condições de estresse hídrico, conforme descrito na literatura para outras TLPs (Muoki et al., 2021).

Uma vantagem das TLPs extracelulares para produção biotecnológica é a simplificação de etapas de purificação e a não acumulação das proteínas no interior das células, o que pode evitar mudanças negativas no crescimento do microrganismo utilizado para produção da TLP (Zhang et al., 2021).

Tabela 6 – Predição de localização subcelular das potenciais TLPs utilizando a ferramenta BUSCA.

Localização subcelular	Proteínas
Extracelular	TLP4_7, TLP5_11, TLP31_18, TLP35_6, TLP37_12, TLP41_17, TLP51_17, TLP85_17, TLP103_6, TLP109_15, TLP120_19, TLP143_20
Membrana plasmática	TLP1_16, TLP3_17, TLP21_9, TLP40_7, TLP43_6, TLP55_6, TLP78_16, TLP90_14, TLP97_13, TLP98_5, TLP101_6, TLP106_5, TLP119_7, TLP126_7
Núcleo	TLP30_7, TLP80_5, TLP112_14, TLP136_15

Fonte: elaborado pelo autor (2025).

A avaliação de toxicidade no ToxinPred 3.0 (tabela 6) demonstrou que 6 das TLPs avaliadas não apresentaram toxicidade, sendo elas a TLP4_7, TLP35_6, TLP78_16, TLP103_6, TLP112_14 e TLP136_15. Para aplicação como agentes antimicrobianos, é importante que os peptídeos antimicrobianos (AMPs) apresentem atividade contra o patógeno de interesse e apresentem baixa toxicidade na dose

terapêutica (Fjell et al., 2012), o que torna as TLPs não tóxicas alvos interessantes para desenvolvimento de novos antimicrobianos. Apesar de 24 TLPs terem sido classificadas como tóxicas, o uso delas como AMPs ainda não está descartado. Para lidar com essas TLPs tóxicas, análises voltadas para a identificação/modificação de resíduos de aminoácidos responsáveis pela atividade tóxica podem ser utilizadas visando obter sequências menos tóxicas (Fjell et al., 2012). Além disso, o encapsulamento dos peptídeos com nanoestruturas é uma técnica que pode ser utilizada para a diminuição da toxicidade dessas proteínas em aplicações, além de permitirem maior estabilidade, direcionamento ao tecido ou às células-alvo e melhorar a eficiência terapêutica (Imperlini; Massaro; Buonocore, 2023).

Tabela 7 – Predição de toxicidade das TLPs utilizando a ferramenta ToxinPred 3.0.

Toxicidade	Proteínas
Não tóxica	TLP4_7, TLP35_6, TLP78_16, TLP103_6, TLP112_14 e TLP136_15 TLP1_16, TLP3_7, TLP 5_11, TLP21_9, TLP30_7, TLP31_18, TLP37_12, TLP40_7, TLP 41_17, TLP43_6, TLP51_17, TLP55_6,
Tóxica	TLP78_16, TLP85_17, TLP90_14, TLP97_13, TLP98_5, TLP101_6, TLP106_5, TLP109_15, TLP119_7, TLP120_19, TLP126_7 e TLP143_20

Fonte: elaborado pelo autor (2025).

Para a modelagem 3D das TLPs potenciais, foram utilizados os templates demonstrados na tabela 8. Dentre os modelos, a 3ZS3 é a Mal d 2, uma TLP presente em *Malus domestica* (Krebitz et al., 2003); a 2AHN é referente à TLP Pru av 2 presente em *Prunus avium* (Dall’Antonia et al., 2005); a 4L2J é uma osmotina (proteína pertencente à família PR-5) contida no látex de *Calotropis procera* (Bruno-Moreno et al., 2013); e a 7P20 é uma TLP de *Juniperus ashei* (Eder; Hofer; Keller, 2021). Segundo Baker e Sali (2001), modelos de alta precisão estão relacionados à uma identidade superior a 50% da proteína a ser modelada com o modelo utilizado. Nesse sentido, das TLPs avaliadas, 19 apresentaram identidade superior a 50% e 11 delas apresentaram valor inferior a 50%.

Tabela 8 – Modelos estruturais utilizados na modelagem de TLPs de *Manihot pseudoglaziovii* e os percentuais de identidade de sequência entre as TLPs e o modelo.

Código PDB do modelo	Proteína	Identidade (%)
-----------------------------	-----------------	-----------------------

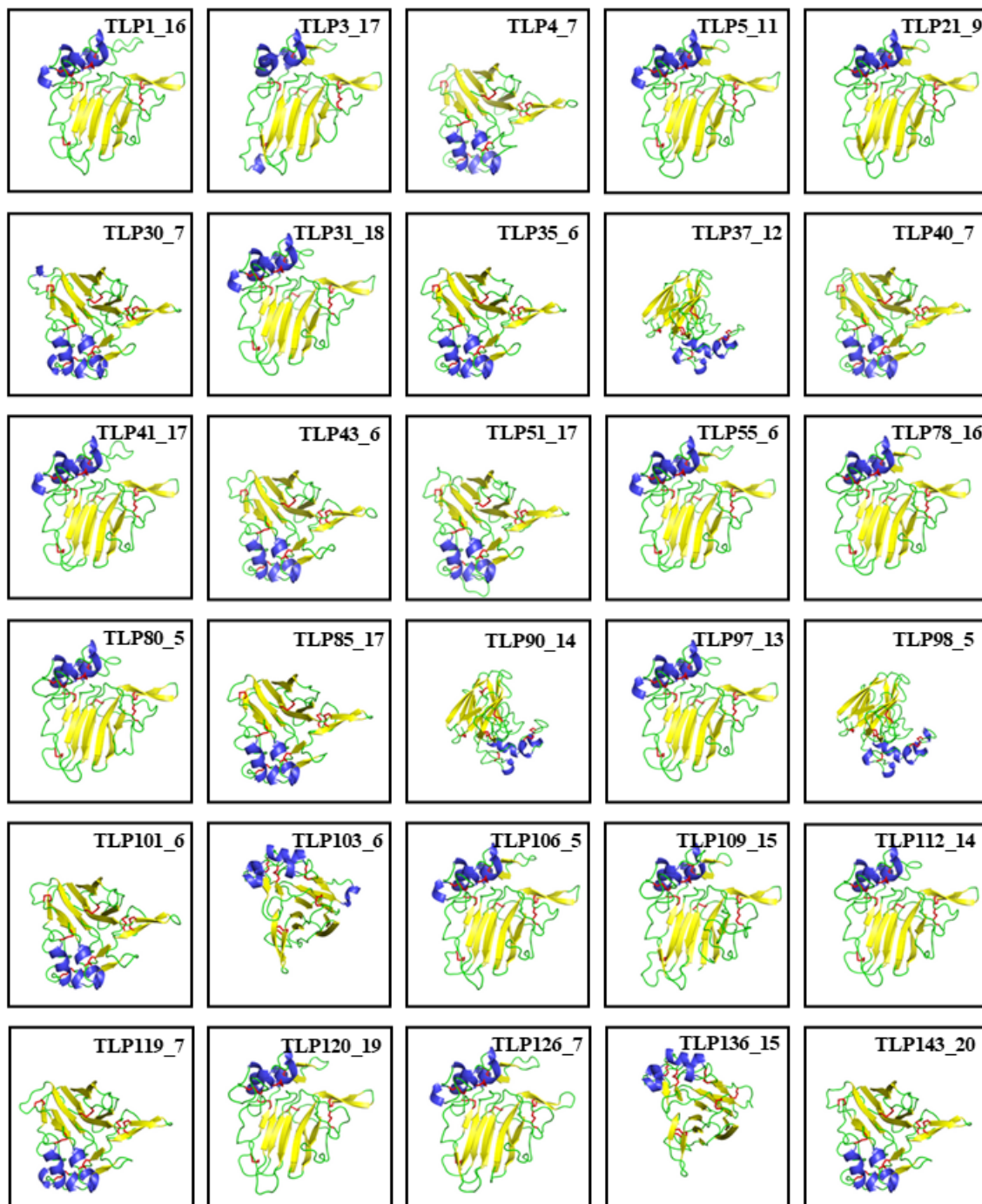
3ZS3	TLP1_16	52,94
	TLP3_17	35,29
	TLP5_11	46,12
	TLP21_9	48,86
	TLP31_18	55,56
	TLP41_17	50,46
	TLP55_6	49,54
	TLP78_16	55,56
	TLP80_5	54,93
	TLP97_13	57,41
	TLP106_5	55,20
	TLP109_15	35,19
	TLP112_14	54,46
	TLP120_19	42,73
TLP126_7	41,36	
2AHN	TLP4_7	53,70
	TLP30_7	32,72
	TLP35_6	43,19
	TLP40_7	51,83
	TLP43_6	57,47
	TLP51_17	36,28
	TLP85_17	43,19
	TLP101_6	59,46
	TLP119_7	54,75
	TLP143_20	54,75
7P20	TLP103_6	54,27
	TLP136_15	51,01
4L2J	TLP37_12	70,65
	TLP90_14	78,92
	TLP98_5	72,64

Fonte: elaborado pelo autor (2025).

Os modelos 3D de TLPs gerados a partir dos templates (tabela 8) estão demonstrados na figura 1. A modelagem das proteínas revelou que as TLPs modeladas a partir de um mesmo molde apresentaram características tridimensionais semelhantes. Os resultados são consistentes aos obtidos por Freitas et al. (2023), em que os autores destacam a presença de três domínios característicos das TLPs. O domínio I apresenta 11 folhas betas, onde seis folhas beta antiparalelas interagem com cinco folhas beta (Freitas et al., 2023). O domínio II é composto por α -hélices e o domínio III apresenta um laço longo e uma folha beta antiparalela (Freitas et al., 2023). Nos modelos de *M. pseudoglaziovii*, foi possível observar a presença das α -hélices características do domínio II, representadas em azul na figura 1. Além disso, a partir das folhas betas, representadas na cor amarelo (figura 1), é possível observar o grupo de folhas beta

antiparalela do domínio I, bem como a folha beta antiparalela do domínio III. Essas características estruturais evidenciam a presença da arquitetura típica das TLPs nos modelos de *M. pseudoglaziovii*, o que sugere que as proteínas modeladas podem conservar a atividade funcional relacionada à família, embora estudos experimentais sejam necessários para a validação definitiva.

Figura 1 – Modelagem 3D das TLPs utilizando o SWISS-MODEL. As α -hélices estão representadas em azul, as folhas beta em amarelo e as ligações dissulfeto em vermelho.



Fonte: elaborado pelo autor (2025).

A partir das análises realizadas, foi realizada uma comparação entre as propriedades apresentadas pelas TLPs, exceto a atividade antimicrobiana, já que todas a apresentaram, com o objetivo de selecionar as mais promissoras (figura 2). Com isso, as TLPs candidatas que apresentaram maior potencial foram a TLP143_20, TLP5_11, TLP35_6 e TLP4_7, por apresentarem cinco das sete propriedades desejadas comparadas. Em contrapartida, as proteínas TLP1_16, TLP41_17, TLP55_6, TLP78_16, TLP109_15 e TLP126_7 apresentaram somente duas das sete propriedades desejadas utilizadas na comparação, indicando menor potencial.

Figura 2 – Comparação das TLPs candidatas, classificadas em ordem decrescente de acordo com o número de características desejadas apresentadas (destacadas em verde).

TLP	Domínio conservado específico	Estável	Oito pontes dissulfeto	Natureza básica	Peptídeo Sinal	Não tóxica	Extracelular
TLP143_20	Sim	Sim	Sim	Sim	Não	Não	Sim
TLP5_11	Sim	Sim	Sim	Sim	Não	Não	Sim
TLP35_6	Sim	Sim	Não	Sim	Não	Sim	Sim
TLP4_7	Sim	Não	Não	Sim	Sim	Sim	Sim
TLP136_15	Sim	Sim	Não	Sim	Não	Sim	Não
TLP43_6	Sim	Sim	Sim	Sim	Não	Não	Não
TLP51_17	Sim	Não	Não	Sim	Sim	Não	Sim
TLP85_17	Sim	Sim	Não	Sim	Não	Não	Sim
TLP90_14	Sim	Sim	Não	Sim	Sim	Não	Não
TLP103_6	Sim	Sim	Não	Não	Não	Sim	Sim
TLP112_14	Sim	Sim	Não	Sim	Não	Sim	Não
TLP37_12	Não	Sim	Não	Sim	Sim	Não	Sim
TLP3_17	Sim	Não	Não	Sim	Sim	Não	Não
TLP21_9	Sim	Não	Sim	Sim	Não	Não	Não
TLP30_7	Sim	Sim	Não	Sim	Não	Não	Não
TLP31_18	Sim	Não	Não	Sim	Não	Não	Sim
TLP40_7	Sim	Não	Sim	Sim	Não	Não	Não
TLP80_5	Sim	Não	Sim	Sim	Não	Não	Não
TLP97_13	Sim	Não	Sim	Sim	Não	Não	Não
TLP98_5	Não	Sim	Não	Sim	Sim	Não	Não
TLP101_6	Sim	Não	Não	Sim	Sim	Não	Não
TLP106_5	Sim	Não	Sim	Sim	Não	Não	Não
TLP119_7	Sim	Não	Sim	Sim	Não	Não	Não
TLP120_19	Sim	Não	Não	Sim	Não	Não	Sim
TLP1_16	Sim	Sim	Não	Não	Não	Não	Não
TLP41_17	Sim	Não	Não	Não	Não	Não	Sim
TLP55_6	Sim	Não	Sim	Não	Não	Não	Não
TLP78_16	Sim	Não	Não	Não	Não	Sim	Não
TLP109_15	Não	Não	Não	Sim	Não	Não	Sim
TLP126_7	Sim	Não	Não	Não	Sim	Não	Não

Fonte: elaborado pelo autor (2025).

6. CONSIDERAÇÕES FINAIS

Foram identificadas 30 possíveis sequências de TLPs candidatas para uso como antimicrobiano considerando todos os parâmetros analisados. Destas, as proteínas TLP5_11, TLP43_6 e TLP143_20 são candidatas promissoras para desenvolvimento de novas aplicações biotecnológicas, mas necessitam de validação experimental para confirmação da atividade antimicrobiana e investigação da toxicidade em sistemas vivos. Além disso, essas TLPs são alvos interessantes para estudos posteriores envolvendo modificações nas suas sequências, visando otimização das propriedades

apresentadas ou o desenvolvimento de variantes com características desejadas aprimoradas.

REFERÊNCIAS BIBLIOGRÁFICAS

ACHARYA, Karan *et al.* Overexpression of *Camellia sinensis* thaumatin-like protein, CsTLP in potato confers enhanced resistance to *Macrophomina phaseolina* and *Phytophthora infestans* infection. **Molecular Biotechnology**, v. 54, p. 609-622, 2013.

AFGAN, Enis *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. **Nucleic acids research**, v. 46, p. 537-544, 2018. Disponível em: <https://academic.oup.com/nar/article/46/W1/W537/5001157>. Acesso em: 15 abr. 2025.

ALTMAN, Russ B.; DUGAN, Jonathan M. Defining bioinformatics and structural bioinformatics. *In*: BOURNE, P. E.; WEISSIG, H. **Structural Bioinformatics**, v. 44, p. 1-14, 2003.

ALVES, J. J. A.; ARAÚJO, M. A. N.; NASCIMENTO, S. S. Degradação da Caatinga: uma investigação ecogeográfica. **Revista Caatinga**, v. 22, n. 3, p. 126-135, 2009.

ALVES, M. C. S.; SOUZA, R. S.; SILVA, R. C. C. Genome-Wide Identification and Characterization Thaumatin-Like Protein (TLP) Genes in Wild Olive (*Olea europaea* var. *sylvestris*). **Scientia Plena**, v. 20, n. 9, 2024. Disponível em: <https://doi.org/10.14808/sci.plena.2024.095201>. Acesso em 10 jun. 2025.

ANDREWS, Simon. FastQC: a quality control tool for high throughput sequence data. 2010. Disponível em: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Acesso em: 09 abr. 2025.

BAILEY, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. **Nucleic acids research**, v. 37, p. 202-208, 2009. Disponível em: <https://doi.org/10.1093/nar/gkp335>. Acesso em: 17 maio. 2025.

BAKER, D.; SALI, A. Protein Structure Prediction and Structural Genomics. **Science**, v. 294, n. 5540, p. 93-96, 2001.

BELTRÃO, F. A. S. *et al.* Morfometria de acessos de maniçoba (*Manihot pseudoglaziovii* Pax & Hoffm.) e de duas espécies afins de interesse forrageiro. **Revista Caatinga**, v. 19, n. 2, p. 103-111, 2006.

BOLGER, Anthony M.; LOHSE, Marc; USADEL, Bjoern. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-2120, 2014. Disponível em: <https://doi.org/10.1093/bioinformatics/btu170>. Acesso em: 06 abr. 2025.

BOTO, A.; LASTRA, J. M. P.; GONZÁLEZ, C. C. The road from host-defense peptides to a new generation of antimicrobial drugs. **Molecules**, v. 23, n. 2, p. 311, 2018. Disponível em: <https://doi.org/10.3390/molecules23020311>. Acesso em: 21 abr. 2025.

BRUNO-MORENO, F. *et al.* Crystallization and X-ray diffraction analysis of an antifungal laticifer protein. **Structural Biology and Crystallization Communications**, v. 69, n. 6, p. 646-649, 2013.

COSTA, Fernando Guilherme Perazzo *et al.* Avaliação do feno de maniçoba (*Manihot pseudoglaziovii* Paz & Hoffman) na alimentação de aves caipiras. **Revista Caatinga**, v. 20, n. 3, p. 42-48, 2007.

DALL'ANTONIA, Y. *et al.* Crystallization and preliminary structure determination of the plant food allergen Pru av 2. **Structural Biology and Crystallization Communications**, v. 61, n. 2, p.186-188, 2005.

DELANO, W. L. Pymol: An open-source molecular graphics tool. **CCP4 Newsletter on Protein Crystallography**, v. 40, 2002.

D'ERRICO, Antonio *et al.* Identification and Characterization of Neuroprotective Properties of Thaumatin-like Protein 1a from Annurca Apple Flesh Polyphenol Extract. **Nutrients**, v. 16, n. 2, p. 307, 2024. Disponível em: <https://doi.org/10.3390/nu16020307>. Acesso em: 14 abr. 2025.

EDER, M.; HOFER, G.; KELLER, W. High resolution structure of the *Juniperus ashei* allergen - Jun a 3. Disponível em: <https://doi.org/10.2210/pdb7p20/pdb>. Acesso em: 23 jun. 2025.

EHIObU, J. M.; OGU, G. I. Phytochemical Content and In Vitro Antimycelial Efficacy of *Colocasia esculenta* (L), *Manihot esculenta* (Crantz) and *Dioscorea rotundata* (Poir) Leaf Extracts on *Aspergillus niger* and *Botryodiplodia theobromae*. **Journal of Horticulture and Plant Research**, v. 1, p. 9-18, 2018.

FABRI, R. L. *et al.* Antimicrobial, antioxidant and cytotoxic potential of *Manihot multifida* (L.) Crantz (Euphorbiaceae). **Anais da Academia Brasileira de Ciências**, v. 87, n. 1, p. 303-311, 2015.

FAILLACE, Giulia Ramos *et al.* Genome-wide analysis and evolution of plant thaumatin-like proteins: a focus on the origin and diversification of osmotins. **Molecular Genetics and Genomics**, v. 294, p. 1137-1157, 2019.

FERNANDES, M. F.; QUEIROZ, L. P. Vegetação e flora da Caatinga. **Ciência e cultura**, v. 70, n. 4, p. 51-56, 2018. Disponível em: <http://dx.doi.org/10.21800/2317-66602018000400014>. Acesso em: 14 abr. 2025.

FENG, L.; WEI, S.; LI, Y. Thaumatin-like Proteins in Legumes: Functions and Potential Applications—A Review. **Plants**, v. 13, n. 8, 2024. Disponível em: <https://doi.org/10.3390/plants13081124>. Acesso em: 14 abr. 2025.

FJELL, C. D. *et al.* Designing antimicrobial peptides: form follows function. **Nature Reviews Drug Discovery**, v. 11, n. 1, p- 37-51, 2012.

- FREITAS, C. D. T. *et al.* Proteomic analysis of *Cryptostegia grandiflora* latex, purification, characterization, and biological activity of two osmotin isoforms. **International Journal of Biological Macromolecules**, v. 252, 2023.
- GASTEIGER *et al.* Protein Identification and Analysis Tools on the ExPASy Server. In: WALKER, J. M. **The Proteomics Protocols Handbook**. v.1, p. 571-607, 2005.
- GUPTA *et al.* *In Silico* Approach for Predicting Toxicity of Peptides and Proteins. **PLOS One**, v. 8, n. 9, 2013. Disponível em: <https://doi.org/10.1371/journal.pone.0073957>. Acesso em: 17 maio. 2025.
- IMPERLINI, E.; MASSARO, F.; BUONOCORE, F. Antimicrobial Peptides Against Bacterial Pathogens: Innovative Delivery Nanosystems for Pharmaceutical Applications. **Antibiotics**, v. 12, n. 1, 2023. Disponível em: <https://doi.org/10.3390/antibiotics12010184>. Acesso em: 08 jun. 2025.
- JESÚS-PIRES, C. *et al.* Plant thaumatin-like proteins: function, evolution and biotechnological applications. **Current Protein and Peptide Science**, v. 21, n. 1, p. 36-51, 2020.
- KREBITZ, M. *et al.* Plant-based Heterologous Expression of Mal d 2, a Thaumatin-like Protein and Allergen of Apple (*Malus domestica*), and its Characterization as an Antifungal Protein. **Journal of Molecular Biology**, v. 329, n. 4, p. 721-730, 2003.
- LIMA, P. G. *et al.* Synthetic antimicrobial peptides: Characteristics, design, and potential as alternative molecules to overcome microbial resistance. **Life Sciences**, v. 278, 2021. Disponível em: <https://doi.org/10.1016/j.lfs.2021.119647>. Acesso em: 21 abr. 2025.
- LI, Z. *et al.* Comprehensive genome-wide analysis of thaumatin-like gene family in four cotton species and functional identification of GhTLP19 involved in regulating tolerance to *Verticillium dahlia* and drought. **Frontiers in Plant Science**, v. 11, 2020.
- LIU, J.; STURROCK, R.; EKRAMODDOULLAH, A. K. M. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. **Plant cell reports**, v. 29, p. 419-436, 2010.
- MARCHLER-BAUER, A. *et al.* CDD: A Conserved Domain Database for the functional annotation of proteins. **Nucleic Acids Research**, v. 39, 2010. Disponível em: <https://doi.org/10.1093/nar/gkq1189>. Acesso em: 07 jun. 2025.
- MADEIRA, F. *et al.* The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. **Nucleic Acids Research**, v. 52, p. 521-525, 2024. Disponível em: <https://doi.org/10.1093/nar/gkae241>. Acesso em: 18 maio. 2025.
- MARR, A. K.; GOODERHAM, W. J.; HANCOCK, R. E. Antibacterial peptides for therapeutic use: obstacles and realistic outlook. **Current Opinion in Pharmacology**, v. 6, n. 5, p. 468-472, 2006.

MASUDA, Tetsuya *et al.* A hypersweet protein: removal of the specific negative charge at Asp21 enhances thaumatin sweetness. **Scientific reports**, v. 6, n. 1, 2016. Disponível em: <https://doi.org/10.1038/srep20255>. Acesso em: 13 abr. 2025.

MORAES, Jadson Patrick Santana *et al.* Avaliação do crescimento vegetativo em plantas de Maniçoba (*Manihot pseudoglaziovii*) encontradas no bioma Caatinga-Região do vale do São Francisco. **Revista Brasileira de Biociências**, v. 5, n. 2, p. 1071-1073, 2007.

MORDKOVICH, N. N.; OKOROKOVA, N. A.; VEIKO, V. P. Structural and Functional Organization of the Signal Peptide of Pro-Enterotoxin B from *Staphylococcus aureus*. **Applied Biochemistry and Microbiology**, v. 51, n. 6, p. 641-648, 2015.

MUNIS, H. F. H. *et al.* A thaumatin-like protein gene involved in cotton fiber secondary cell wall development enhances resistance against *Verticillium dahlia* and other stresses in transgenic tobacco. **Biochemical and Biophysical Research Communications**, v. 393, n. 1, p. 38-44, 2010.

MUOKI, R. C. *et al.* Membrane localized *thaumatin-like protein* from tea (CsTLP) enhanced seed yield and the plant survival under drought stress in *Arabidopsis thaliana*. **Plant Physiology and Biochemistry**, v. 163, p. 36-44, 2021.

NAWROT, R. *et al.* Characterization and expression of a novel thaumatin-like protein (CcTLP1) from papaveraceous plant *Corydalis cava*. **International Journal of Biological Macromolecules**, v. 189, p. 678-689, 2021. Disponível em: <https://doi.org/10.1016/j.ijbiomac.2021.08.067>. Acesso em: 20 abr. 2025.

NEGRÃO, B. L. P. C. 2019. **Maniçobeira (*Manihot cf. pseudoglaziovii* Pax & K. Hoffm.): utilização e importância para a agricultura familiar em Santo Antonio do Tauá, Pará**. 2019. 87 p. Dissertação (Mestrado em Gestão de Empreendimentos Agroalimentares) – Instituto Federal de Educação, Ciência e Tecnologia do Pará, Castanhal, 2019.

NINA, Matheus Mendes *et al.* Potencialidade de *Manihot esculenta* Crantz (Euphorbiaceae) na Floresta Amazônica, Brasil. **Diversitas Journal**, v. 6, n. 2, p. 2247-2260, 2021. Disponível em: <https://doi.org/10.17648/diversitas-journal-v6i2-1253>. Acesso em: 11 abr. 2025.

OWJI, H. *et al.* A comprehensive review of signal peptides: Structure, roles, and applications. **European journal of cell biology**, v. 97, n. 6, p. 422-441, 2018.

PEREIRA, L. S. *et al.* Insecticidal and antifungic proteins of the latex of *Manihot glaziovii* Muell. Arg. **Brazilian Journal of Botany**, v. 22, p. 27-30, 1999.

PESTANA-CALSA C. M. *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.

PETRE, Benjamin *et al.* Genome-wide analysis of eukaryote thaumatin-like proteins (TLPs) with an emphasis on poplar. **BMC plant biology**, v. 11, p. 1-16, 2011. Disponível em: <https://doi.org/10.1186/1471-2229-11-33>. Acesso em: 14 abr. 2025.

RATHORE, A. S. *et al.* ToxinPred 3.0: An improved method for predicting the toxicity of peptides. **Computers in Biology and Medicine**, v. 179, 2024. Disponível em: <https://doi.org/10.1016/j.compbiomed.2024.108926>. Acesso em: 31 maio. 2025.

RICE, P; LONGDEN, I.; BLEASBY A. EMBOSS: The European Molecular Biology Open Software Suite. **Trends in genetics**, v. 16, n. 6, p. 276-277, 2000.

RYDZYŃSKA, Magdalena *et al.* Allergy to Thaumatin-like Proteins—What Do We Know?. **Foods**, v. 14, n. 4, p. 543, 2025. Disponível em: <https://doi.org/10.3390/foods14040543>. Acesso em: 13 abr. 2025.

SALZMAN, R. A. *et al.* Inorganic Cations Mediate Plant PR5 Protein Antifungal Activity through Fungal Mnn1- and Mnn4-Regulated Cell Surface Glycans. **Molecular Plant-Microbe Interactions**, v. 17, n. 0, p. 780-788, 2004. Disponível em: <https://doi.org/10.1094/MPMI.2004.17.7.780>. Acesso em: 16 jun. 2025.

SAVOJARDO, C. *et al.* BUSCA: an integrative web server to predict subcellular localization of proteins. **Nucleic acids research**, v. 46, p. 459-466, 2018. Disponível em: <https://doi.org/10.1093/nar/gky320>. Acesso em: 17 maio. 2025.

SILVA, Pedro Henrique Ferreira da *et al.* Níveis crescentes de fubá de milho melhoram parâmetros químico-bromatológicos e fermentativos da silagem de maniçoba. **Ciência Animal Brasileira**, v. 25, 2024. Disponível em: <https://doi.org/10.1590/1809-6891v25e-75861P>. Acesso em: 11 abr. 2025.

SMOLE, U. *et al.* Mal d 2, the thaumatin-like allergen from apple, is highly resistant to gastrointestinal digestion and thermal processing. **International Archives of Allergy and Immunology**, v. 147, n. 4, p. 289-298, 2008.

SUDISHA, J. *et al.* Pathogenesis related proteins in plant defense response. *In*: MÉRILLON, J.; RAMAWAT, K. **Plant defence: biological control**, v. 12, p. 379-403, 2012.

SVENSON J. *et al.* Antimicrobial Peptides with Stability toward Tryptic Degradation. **Biochemistry**, v. 47, n. 12, p. 3777-3788, 2008.

TACHI, H. *et al.* Molecular characterization of a novel soybean gene encoding a neutral PR-5 protein induced by high-salt stress. **Plant Physiology and Biochemistry**, v. 47, n. 1, p. 73-79, 2009.

TALAVERA, G.; CASTRESANA, J. Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. **Systematic Biology**, v. 54, n. 4, p. 564-577, 2007.

TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. **Molecular Biology and Evolution**, v. 38, n. 7, p. 3022-

3027, 2021. Disponível em: <https://doi.org/10.1093/molbev/msab120>. Acesso em: 18 maio. 2025.

TEUFEL, F. *et al.* SignalP 6.0 predicts all five types of signal peptides using protein language models. **Nature biotechnology**, v. 40, n. 7, p. 1023-1025, 2022. Disponível em: <https://doi.org/10.1038/s41587-021-01156-3>. Acesso em: 17 maio. 2025.

TOSSI, A. Design and Engineering Strategies for Synthetic Antimicrobial Peptides. *In*: DRIDER, D.; REBUFFAT, S. **Prokaryotic Antimicrobial Peptides**, v. 1, p. 81-98, 2011.

VIKTOROVA, J. *et al.* New findings in potential applications of tobacco osmotin. **Protein Expression and Purification**, v. 129, p. 84-93, 2017.

WAGHU, F. H. *et al.* CAMPR3: a database on sequences, structures and signatures of antimicrobial peptides. **Nucleic acids research**, v. 44, p. 1094-1097, 2016. Disponível em: <https://doi.org/10.1093/nar/gkv1051>. Acesso em: 17 maio. 2025.

WANG, J. *et al.* The conserved domain database in 2023. **Nucleic Acids Research**, v. 51, p. 384-388, 2023. Disponível em: <https://doi.org/10.1093/nar/gkac1096>. Acesso em: 18 maio. 2025.

WATERHOUSE, A. M. *et al.* Jalview Version 2—a multiple sequence alignment editor and analysis workbench. **Bioinformatics**, v. 25, n. 9, p. 1189-1191, 2009. Disponível em: <https://doi.org/10.1093/bioinformatics/btp033>. Acesso em: 28 maio. 2025.

WATERHOUSE, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic acids research**, v. 46, p. 296-303, 2018. Disponível em: <https://doi.org/10.1093/nar/gky427>. Acesso em: 18 maio. 2025.

ZHANG, Y. *et al.* A Method for Rapid Screening, Expression, and Purification of Antimicrobial Peptides. **Microorganisms**, v. 9, n. 9, 2021. Disponível em: <https://doi.org/10.3390/microorganisms9091858>. Acesso em: 29 jul. 2025.