

**UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE  
DEPARTAMENTO DE GENÉTICA E EVOLUÇÃO**

**YAN EIJI MATUHARA**

**EXPRESSÃO RECOMBINANTE E ENSAIOS DE ATIVIDADE DE DUAS  
*THAUMATIN-LIKE PROTEINS* DO MAQUI (*Aristotelia chilensis*)**

**SÃO CARLOS - SP  
2022**

YAN EIJI MATUHARA

**EXPRESSÃO RECOMBINANTE E ENSAIOS DE ATIVIDADE DE DUAS  
*THAUMATIN-LIKE PROTEINS* DO MAQUI (*Aristolelia chilensis*)**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos como pré-requisito para a obtenção do título de Mestre em Genética Evolutiva e Biologia Molecular

Orientador: Prof. Dr. Flávio Henrique da Silva

**SÃO CARLOS - SP  
2022**



**UNIVERSIDADE FEDERAL DE SÃO CARLOS**

Centro de Ciências Biológicas e da Saúde  
Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

---

**Folha de Aprovação**

---

Defesa de Dissertação de Mestrado do candidato Yan Eiji Matuhara, realizada em 24/10/2022.

**Comissão Julgadora:**

Prof. Dr. Flavio Henrique da Silva (UFSCar)

Prof. Dr. Maurício Bacci Júnior (UNESP)

Prof. Dr. Fernando Segato (USP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular.

Dedico com muito carinho este trabalho à minha mãe, Sonia, por todo apoio, amor, companhia, incentivo e paciência. Sem ela, este trabalho não seria possível.

## AGRADECIMENTOS

Meu primeiro agradecimento é voltado especialmente para minha mãe, Sonia, por seu amor incondicional e apoio antes e durante todo o percurso deste trabalho. Você está sempre comigo, mãe, mesmo quando estamos fisicamente longe.

Agradeço meu pai, Kiyoshi, e meu irmão, Hideki, por toda a confiança, carinho e suporte.

Ao Prof. Dr. Flávio Henrique da Silva, não somente pela orientação e ensinamentos que levarei comigo sempre, mas também pela paciência que teve ao me orientar, por acreditar em mim e me incentivar a ir sempre mais longe e pela amizade que nutrimos ao longo dos anos.

À Profa. Dra. Andrea Soares da Costa Fuentes, por sempre ajudar quando necessário.

Ao Prof. Dr. Ariel D. Arencibia Rodriguez, pela colaboração e disponibilização de material biológico que possibilitou a realização deste trabalho.

À Dra. Milene Ferro, que realizou as análises de bioinformática.

Aos meus colegas do Laboratório de Biologia Molecular, Iara, Eduardo, Thais, Henrique, Amanda, Felipe, Josilene, Bianca, Gabrieli, Karen, Letícia e Priscila, com os quais compartilhei risadas, conversas, ensinamentos, dicas, companhias e ajudas. Um agradecimento especial à Iara, já que compartilhamos mais fortemente entre nós tudo o que foi citado.

Aos meus amigos e amigas que fiz durante a graduação. Alguns já distantes, outros ainda pela cidade. Todos vocês me ajudaram a seguir em frente com esta pesquisa.

Aos meus amigos e amigas virtuais do *Nox* e do *Chat Universal*, por toda companhia, conversas e brincadeiras que vocês sempre mantiveram comigo. Um agradecimento especial ao Pedro, que me ajudou com alguns gráficos, e ao Samuel, que me ensinou e auxiliou na edição de algumas imagens. A conclusão deste trabalho não seria possível sem vocês.

À Universidade Federal de São Carlos, ao Departamento de Genética e Evolução e ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular, pela oportunidade, suporte e infraestrutura oferecidos que possibilitaram a conclusão deste trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa e apoio financeiro.

E a todos que contribuíram direta ou indiretamente para a realização deste trabalho.

## RESUMO

As *Thaumatococcus-like Proteins* (TLPs) são membros da família 5 das proteínas *Pathogenesis-Related* (PR-5), estando presentes em fungos, plantas e animais invertebrados. Em plantas, elas estão envolvidas tanto com a defesa do organismo, frequentemente relacionadas com atividades antifúngicas, e com a regulação de processos fisiológicos, como a floração e maturação dos frutos. O Maqui (*Aristotelia chilensis*) é uma planta nativa da América do Sul e profundamente ligada ao Chile, sendo usada na indústria alimentícia e na medicina popular da cultura chilena. Recentemente, esta planta recebeu o título de “Superfruta”, destacada por causa de sua alta concentração de antocianinas. Como é uma espécie vegetal pouco estudada, a exploração pelo potencial biotecnológico de suas proteínas é ainda limitada. Diante do exposto, este trabalho consistiu na identificação, expressão heteróloga e realização de ensaios de atividade de TLPs do Maqui. A partir do transcriptoma da planta cultivada em biorreator de imersão temporária (TIBs), foram identificadas quatro sequências, das quais duas tiveram suas sequências obtidas por amplificação do cDNA. Estas foram clonadas em vetor de expressão pPICZαA para a produção recombinante em *Pichia pastoris* X-33. As proteínas purificadas foram submetidas a ensaios de B-1,3-Glucanase, de inibição do crescimento de hifas de fungos, de tolerância ao estresse osmótico e de permeabilização da membrana plasmática de levedura. Os resultados indicam que as TLPs de Maqui são capazes de hidrolisar beta-glucanos e também de permeabilizar a membrana plasmática de *Pichia pastoris*. Contudo, elas não apresentaram capacidade de inibir o crescimento de hifas fúngicas ou de conferir tolerância ao estresse osmótico.

**Palavras-chave:** *Aristotelia chilensis*. Biorreator de Imersão Temporária. Glucanase. Permeabilização. *Thaumatococcus-like Proteins*.

## ABSTRACT

*Thaumatococcus*-like Proteins (TLP) are members of the *Pathogenesis-Related Proteins* family 5 (PR-5), and are present in fungi, plants and invertebrates. In plants, these proteins are involved both in the host's defense, often associated with antifungal properties, and in the regulation of physiological processes, such as floration or fruit maturation. Maqui (*Aristotelia chilensis*) is a plant native to South America and intimately related to Chile, being used in the food industry and traditional medicine of the Chilean culture. Lately, this plant has received the title of "Superfruit" because of its high anthocyanin content. Since this is a species that has recently gained the spotlight, the prospect for the biotechnological potential of its proteins is still limited. Therefore, this work consisted in the identification, heterologous expression and activity assays of TLP from Maqui. From the transcriptome of the plant cultivated in Temporary Immersion Bioreactors, four sequences were identified, of which two were obtained by cDNA amplification. These were cloned in the expression vector pPICZαA for the recombinant production in *Pichia pastoris* expression system. The purified proteins were submitted to assays of β-1,3-Glucanase, inhibition of fungal hyphae growth, osmotic stress tolerance and permeabilization of yeast plasmatic membrane. The results suggest that the TLP from Maqui are capable of hydrolyzing β-glucans and also able to permeabilize *Pichia pastoris* plasmatic membrane, although they did not show any inhibition on the growth of fungal hyphae or increased the host's osmotic stress tolerance.

Keywords: *Aristotelia chilensis*. Glucanase. Permeabilization. Temporary Immersion Bioreactor. *Thaumatococcus*-like Proteins.

## LISTA DE FIGURAS

**Figura 1. Vista geral do fruto Katemfe**, de *Thaumatococcus danielli*, do qual a Taumatina original foi primeiramente isolada (Fonte: Wikimedia Commons).

**Figura 2. Vista geral da árvore *Aristotelia chilensis***, popularmente conhecida como Maqui (Fonte: Wikimedia Commons).

**Figura 3. Mudanças de *Aristotelia chilensis* micropropagadas em biorreatores de imersão temporária.** A: Panorama geral dos TIBs. B: Vista ampliada dos TIBs. C: Vista mais ampliada dos TIBs. D: Muda de *Aristotelia chilensis* micropropagada em TIB com tratamento com ABA, sendo evidente coloração escura de material vegetal e meio de cultivo, decorrente da produção e liberação de composto fenólicos pela planta.

**Figura 4. Sequência de nucleotídeos e de aminoácidos da AcTLP1.** A região codificante da proteína é mostrada a partir do codon de iniciação ATG e se encerra no codon de terminação TGA, indicado por um asterisco (\*). Destacados em amarelo estão os resíduos do peptídeo sinal putativo da proteína e sua região codificante. Em letra minúscula e destacado azul está a UTR 5'. Em letra minúscula e destacado vermelho está a UTR 3'.

**Figura 5. Sequência de nucleotídeos e de aminoácidos da AcTLP2.** A região codificante da proteína é mostrada a partir do codon de iniciação ATG e se encerra no codon de terminação TAG, indicado por um asterisco (\*). Destacados em amarelo estão os resíduos do peptídeo sinal putativo da proteína e sua região codificante. Em letra minúscula e destacado azul está a UTR 5'. Em letra minúscula e destacado vermelho está a UTR 3'.

**Figura 6. Sequência de nucleotídeos e de aminoácidos da AcTLP3.** A região codificante da proteína é mostrada a partir do codon de iniciação ATG e se encerra no codon de terminação TGA, indicado por um asterisco (\*). Destacados em amarelo estão os resíduos do peptídeo sinal putativo da proteína e sua região codificante. Em letra minúscula e destacado azul está a UTR 5'. Em letra minúscula e destacado vermelho está a UTR 3'.

**Figura 7. Sequência de nucleotídeos e de aminoácidos da AcTLP4.** A região codificante da proteína é mostrada a partir do codon de iniciação ATG e se encerra no codon de terminação TGA, indicado por um asterisco (\*). Destacados em amarelo estão os resíduos do peptídeo sinal putativo da proteína e sua região codificante. Em letra minúscula e destacado azul está a UTR 5'. Em letra minúscula e destacado vermelho está a UTR 3'.

**Figura 8. Alinhamento global das AcTLPs.** As caixas em vermelho indicam as regiões dos prováveis peptídeos sinais identificadas pelo SignalP 6.0. Os colchetes indicam a assinatura familiar das taumatinas (G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C). As setas em vermelho mostram os dezesseis resíduos de cisteínas conservados, responsáveis pelas oito ligações dissulfeto que estabilizam a estrutura da proteína. Os asteriscos representam as posições conservadas dos cinco aminoácidos do motivo REDDD, responsável pela fenda ácida entre os domínios I e II da proteína. Os resíduos considerados altamente similares estão destacados em negrito e dentro de uma caixa vazia de contorno preto. Nos resíduos com identidade estrita, a coluna é colorida com fundo preto.

**Figura 9. Árvore filogenética das AcTLPs.** A inferência filogenética foi feita pelo método da máxima verossimilhança utilizando o modelo Dayhoff+G(4). Os valores de *bootstrap* estão indicados nos nós como valores de porcentagem.

**Figura 10. Alinhamento global das AcTLPs com TLPs de outras espécies.** As caixas em vermelho indicam as regiões dos prováveis peptídeos sinais identificadas pelo SignalP 6.0. Os colchetes indicam a assinatura familiar das taumatinas (G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C). As setas em vermelho mostram os dezesseis resíduos de cisteínas conservados. Os asteriscos representam as posições conservadas dos cinco aminoácidos do motivo REDDD. Os resíduos considerados altamente similares estão destacados em negrito e dentro de uma caixa vazia de contorno preto. Nos resíduos com identidade estrita, a coluna é colorida com fundo preto. As sequências utilizadas estão disponíveis no Anexo II.

**Figura 11. Árvore filogenética das AcTLPs com TLPs de outras plantas.** A inferência filogenética foi feita pelo método de *Neighbour-Joining* com o modelo de Poisson. Os valores de *bootstrap* estão indicados nos nós como valores de porcentagem. Os asteriscos vermelhos indicam as AcTLPs. Os códigos de acesso das sequências proteicas utilizados estão apresentados nos respectivos táxons. Art: *Arabidopsis thaliana*. Pim: *Pinus monticola*. Php: *Physcomitrella patens*. Ors: *Oryza sativa* subsp. *japonica*.

**Figura 12. Predição da estrutura terciária das AcTLPs de *Aristolelia chilensis*.** As circunferências vermelhas indicam os três domínios de cada proteína. A: AcTLP1. B: AcTLP2. C: AcTLP3. D: AcTLP4.

**Figura 13. Amplificação das ORFs das AcTLPs e clivagem do pPICZA.** Gel de agarose 1% contendo em M: marcador de peso molecular (GeneRuler 1kb ThermoScientific). 1: produto de PCR da AcTLP1. 2: produto de PCR da AcTLP2. 3: pPICZA digerido com *EcoRI* e *SalI*. A seta indica as bandas referentes à amplificação das TLPs (aproximadamente 660 pb) e a cabeça de seta indica a banda referente ao pPICZA clivado (aproximadamente 3.200 pb).

**Figura 15. Amplificação da ORF da GFP e clivagem do pPICZaA para clonagem.** Gel de agarose 1% contendo em M: marcador de peso molecular (GeneRuler 1kb ThermoScientific). 1: produto de PCR da GFP. 2: pPICZaA digerido com *EcoRI* e *SalI*. A: Amplificação da ORF da GFP. B: Clivagem do pPICZaA para clonagem. A seta indica a banda referente à amplificação da GFP (aproximadamente 720 pb). A cabeça de seta indica a banda referente ao pPICZaA clivado (aproximadamente 3.500 pb).

**Figura 15. Confirmação da clonagem por clivagem dos plasmídeos pPICZaA-AcTLP1 e pPICZaA-AcTLP2.** Gel de agarose 1% contendo em M: marcador de peso molecular (GeneRuler 1kb ThermoScientific). 1: pPICZaA-AcTLP1. 2: pPICZaA-AcTLP2. A seta indica as bandas referentes ao inserto clivado (aproximadamente 660 pb). A cabeça de seta indica o vetor pPICZaA clivado (aproximadamente 3.500 pb).

**Figura 16. Confirmação da clonagem por clivagem dos plasmídeos pPICZA-AcTLP1 e pPICZA-AcTLP2.** Gel de agarose 1% contendo em M: marcador de peso molecular (GeneRuler 1kb ThermoScientific). 1: pPICZA-AcTLP1. 2: pPICZA-AcTLP2. A seta indica as bandas referentes ao inserto clivado (aproximadamente 660 pb). A cabeça de seta indica o vetor pPICZA clivado (aproximadamente 3.200 pb).

**Figura 17. Análise da indução das AcTLPs em *Pichia pastoris*.** SDS-PAGE 12% contendo em M: BlueClassic Prestained Protein Marker (Jena Bioscience). 1 e 5: Aliquotas não-induzida (0 h). 2 e 6: Aliquotas de 24 h de indução. 3 e 7: Aliquotas de 48 h de indução. 4 e 8: Aliquotas de 72 h de indução. As cabeças de seta indicam as

várias bandas de proteínas incomuns à expressão secretada em *Pichia pastoris*. A seta aponta as possíveis bandas referentes às AcTLPs.

**Figura 18. Análise das purificações das AcTLPs.** SDS-PAGE 12% mostrando em M: BlueClassic Prestained Protein Marker (Jena Bioscience). 1 e 2: Gradientes de 75 mM de Imidazol. 3 e 4: Gradientes de 100 mM de Imidazol. As setas indicam as bandas referentes às AcTLP1 ou AcTLP2. A: AcTLP1. B: AcTLP2.

**Figura 19. Gráficos de colunas representando a absorbância a 540 nm dos ensaios de atividade de  $\beta$ -1,3-glucanase e seus respectivos controles.** A: Ensaios com AcTLP1. B: Ensaios com AcTLP2. Um asterisco (\*) indica que  $p \leq 0,05$ , três (\*\*\*) indicam que  $p \leq 0,001$ , quatro (\*\*\*\*) indicam que  $p \leq 0,0001$ .

**Figura 20. Alinhamento entre a NP24-I de tomate e as AcTLPs.** Os asteriscos em vermelho representam os resíduos do motivo REDDD conservado. As setas em vermelho indicam os resíduos conservados de Ácido Glutâmico e Ácido Aspártico, os quais possivelmente atuam na hidrólise de  $\beta$ -glucanos. Os resíduos considerados altamente similares estão destacados em negrito e dentro de uma caixa vazia de contorno preto. Nos resíduos com identidade estrita, a coluna é colorida com fundo preto. NP24-I PDB: 2I0W.

**Figura 21. Análise da permeabilização de membrana no co-cultivo com GFP.** SDS-PAGE 12% mostrando em 1, 4 e 7: Amostras controle X-33 + GFP. 2, 5 e 8: Amostras AcTLP1 + GFP. 3, 6 e 9: Amostras AcTLP2 + GFP. As cabeças de seta indicam as possíveis bandas de proteínas oriundas das permeabilização da membrana de *Pichia pastoris*. As setas se referem às prováveis bandas de GFP (~28) e AcTLPs (~24). A: Amostras de 0 a 48 h. B: Amostras de 72 a 120 h.

**Figura 22. Gráfico representando a emissão a 503 nm das alíquotas do co-cultivo ao longo de 144 horas.**

**Figura 23. Ensaio de tolerância ao estresse osmótico de *Pichia pastoris* em placa.** A figura mostra células não transformadas (X-33), células transformadas expressando AcTLPs internamente (pPICZA) ou células transformadas expressando e secretando AcTLPs (pPICZaA). O painel esquerdo não foi induzido com metanol [(-) MeOH], enquanto que o direito foi induzido com metanol [(+) MeOH].

**Figura 24. Alinhamento global entre as AcTLPs e TLPs de outras espécies vegetais.** Os resíduos considerados altamente similares estão destacados em negrito e dentro de uma caixa vazia de contorno preto. Nos resíduos com identidade estrita, a coluna é colorida com fundo preto. As sequências utilizadas estão no Anexo II.

**Figura 25. Mapa geral do vetor pPICZaA.** 5' AOX1: promotor do gene da álcool oxidase 1.  $\alpha$ -factor: peptídeo sinal de secreção fator alfa de *Saccharomyces cerevisiae*. 6xHis: cauda de polihistidina. AOX1 TT: terminador de transcrição do gene da álcool oxidase 1. Zeocin: gene de resistência ao antibiótico Zeocin™ (Fonte: EasySelect Pichia Expression Kit, 2001).

**Figura 26. Sequência do sítio múltiplo de clonagem (MCS) do vetor pPICZaA.** O códon iniciador (ATG) está indicado em negrito. As caixas representam as regiões das enzimas de restrição utilizadas (*EcoRI* e *SalI*) (Fonte: EasySelect Pichia Expression Kit, 2001).

**Figura 27. Mapa geral do vetor pPICZA.** 5' AOX1: promotor do gene da álcool oxidase 1. 6xHis: cauda de polihistidina. AOX1 TT: terminador de transcrição do gene da álcool oxidase 1. Zeocin: gene de resistência ao antibiótico Zeocin™ (Fonte: EasySelect Pichia Expression Kit, 2001).

**Figura 28. Sequência do sítio múltiplo de clonagem (MCS) do vetor pPICZA.** As caixas representam as regiões das enzimas de restrição utilizadas (*EcoRI* e *SalI*). O códon iniciador foi inserido após o sítio da *EcoRI*, usando os oligonucleotídeos indicados nas tabelas 1 e 2 (Fonte: EasySelect Pichia Expression Kit, 2001).

## LISTA DE TABELAS

**Tabela 1. Oligonucleotídeos para PCR das sequências de AcTLP1 e AcTLP2.** Em sublinhado está indicado o sítio de restrição das enzimas correspondentes. Em negrito está indicado o códon iniciador ATG adicionado.

**Tabela 2. Oligonucleotídeos para PCR da sequência da GFP.** Em sublinhado está indicado o sítio de restrição das enzimas correspondentes. Em negrito está indicado o códon iniciador ATG adicionado.

**Tabela 3.** Características das AcTLPs.

**Tabela 4. Identidade e similaridade da estrutura primária entre as AcTLPs e TLPs de outras espécies de plantas.** Os códigos de acesso do NCBI estão ao lado do nome da espécie.

**Tabela 5. Características das AcTLPs recombinantes expressas em *Pichia pastoris*,** possuindo aminoácidos oriundos do plasmídeo de expressão, incluindo *his-tag*.

## LISTA DE ABREVIATURAS E SIGLAS

ABA: ácido abscísico  
BLAST: Basic Local Alignment Search Tool  
BMGY: meio glicerol complexo tamponado (*buffered glycerol-complex medium*)  
BMMY: meio metanol complexo tamponado (*buffered methanol-complex medium*)  
Br: bromo  
CaCl<sub>2</sub>: cloreto de cálcio  
cDNA: DNA Complementar  
Co: cobalto  
Cr: cromo  
Cys (C): cisteína  
DNA: ácido desoxirribonucleico  
dNTP: desoxirribonucleotídeos trifosfatos  
D.O.<sub>600</sub>: densidade ótica a seiscentos nanômetros  
DTT: ditionitrito  
Fe: ferro  
g: gramas  
h: horas  
HEPES: ácido 4-(2-hidroxietil)-1-piperazina etano sulfônico  
HCl: ácido clorídrico  
kDA: quilodalton  
kg: quilogramas  
L: longo (*long*)  
M: molar  
mm: milímetro  
mM: milimolar  
mL: mililitro  
min.: minutos  
Mo: molibdênio  
ORAC: capacidade de absorbância de radical de oxigênio (*oxygen radical absorbance capacity*)  
ORF: fase aberta de leitura (*open reading frame*)  
PBS: tampão fosfato salino (*phosphate-buffered saline*)  
PCR: reação em cadeia da polimerase (*polymerase chain reaction*)  
pH: potencial hidrogeniônico  
pI: ponto isoelétrico  
PR: relacionadas à patogênese (*pathogenesis related*)  
RNA: ácido ribonucleico  
rpm: rotações por minuto  
S: pequeno (*small*)  
s: segundos  
SDS-PAGE: eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio

TIB: biorreator de imersão temporária (*temporary immersion bioreactor*)

TLP: proteína do tipo taumatina (*thaumatin-like protein*)

U: unidade enzimática

UV: ultravioleta

UTR: região não traduzida (*untranslated region*)

YEPD: extrato de levedura, peptona e dextrose (*yeast extract peptone dextrose*)

YEPDS: extrato de levedura, peptona, dextrose e sorbitol (*yeast extract peptone dextrose sorbitol*)

YNB: base nitrogenada para leveduras (*yeast nitrogen base*)

° C: graus celsius

µg: micrograma

µL: microlitro

µm: micrômetro

## SUMÁRIO

<b>1. INTRODUÇÃO</b>	<b>16</b>
1.1. Thaumatin-like Proteins	16
1.2. O Maqui	19
<b>2. OBJETIVOS</b>	<b>22</b>
2.1. Objetivo Geral	22
2.2. Objetivos Específicos	23
<b>3. PERCURSO METODOLÓGICO</b>	<b>23</b>
3.1. Identificação e caracterização das sequências de TLPs	23
3.1.1. Identificação das Taumatinas de Maqui	23
3.1.2. Alinhamentos, filogenia, análises in silico e estatísticas	24
3.2. Construção dos cassetes de expressão	24
3.2.1. Síntese do cDNA	24
3.2.2. Amplificação das ORFs das AcTLPs	24
3.2.3. Clonagem das ORFs das AcTLPs	26
3.2.4. Amplificação e clonagem da GFP	27
3.3. Produção recombinante das AcTLPs	27
3.3.1. Transformação de <i>Pichia pastoris</i>	27
3.3.2. Screening de clones produtores	28
3.3.3. Indução em maior escala e purificação	29
3.8. Ensaio de atividades das TLPs	30
3.8.1. Ensaio de glucanase	30
3.8.2. Permeabilização de Membrana Fúngica	30
3.8.3. Resistência ao Estresse Osmótico	31
<b>4. RESULTADOS E DISCUSSÃO</b>	<b>31</b>
4.1. Alinhamentos e análises filogenéticas	31
4.2. Análises in silico das sequências das AcTLPs	41

4.3. Construção dos cassetes de expressão	44
4.4. Expressão e purificação das TLPs recombinantes	46
4.5. Caracterização das AcTLPs	48
4.5.1. Atividade de $\beta$ -1,3-Glucanase	48
4.5.2. Permeabilização de membrana plasmática de levedura	51
4.5.3. Tolerância ao estresse osmótico	53
<b>5. CONSIDERAÇÕES FINAIS</b>	<b>55</b>
<b>6. PERSPECTIVAS FUTURAS</b>	<b>55</b>
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>56</b>
<b>APÊNDICE I</b>	<b>59</b>
<b>APÊNDICE II</b>	<b>61</b>
<b>APÊNDICE III</b>	<b>65</b>
<b>ANEXO I</b>	<b>67</b>
<b>ANEXO II</b>	<b>70</b>

## 1. INTRODUÇÃO

### 1.1. *Thaumatin-like Proteins*

As plantas possuem tanto mecanismos constitutivos como mecanismos induzíveis para se defenderem contra fitopatógenos. Barreiras morfológicas, metabólitos secundários e peptídeos antimicrobianos são exemplos de mecanismos constitutivos. Entretanto, uma vez que não são suficientes e o contato com o invasor é estabelecido, os mecanismos induzíveis entram em ação, como o acúmulo de fitoalexinas, a resposta hipersensível e a síntese de *Pathogenesis Related (PR) Proteins*. As proteínas PR são classificadas em 17 famílias baseando-se em sequência, estrutura e modo de ação (SINGH *et al.*, 2013; VAN LOON *et al.*, 2006).

A família PR-5 também é conhecida como *Thaumatin-like Proteins (TLP)*, devido à sua similaridade em sequência e estrutural com a Taumatina, uma proteína de aproximadamente 23 kDa e de sabor intensamente doce, a qual foi primeiramente isolada do fruto do arbusto africano *Thaumatococcus danielli* Benth. (Maranthaceae) (Figura 1). Apesar disto, a grande maioria das TLPs posteriormente reportadas não possuem sabor doce; em vez disso, estão envolvidas com respostas a estresses bióticos, abióticos e processos de desenvolvimento (DE JESÚS-PIRES *et al.*, 2020; NAWROT *et al.*, 2021). Diversos sinais induzem a expressão dos genes PR-5, dentre eles os hormônios vegetais Ácido Salicílico, Ácido Abscísico (ABA), Auxina e Etileno; os fatores abióticos seca, frio, luz UV, condições salinas, dissecação e ferimentos; e os fatores bióticos de infecções fúngica, bacteriana e viral (ULLAH *et al.*, 2018).



**Figura 1.** Vista geral do fruto **Katemfe**, de *Thaumatococcus danielli*, do qual a Taumatina original foi primeiramente isolada (Fonte: Wikimedia Commons).

As TLPs possuem em sua sequência a assinatura familiar das taumatinas (G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C; PROSITE: PS00316) e o motivo REDDD conservado, este último estando envolvido na ligação do receptor de sua ação antifúngica. Em relação ao seu peso molecular, elas são agrupadas em tipo L (*long type*) e tipo S (*small type*), tendo pesos moleculares de aproximadamente 21 a 26 kDa e 16 a 17 kDa para os tipos L e S, respectivamente. No total, as TLPs dos tipos L e S possuem, nesta ordem, 16 ou 10 resíduos de cisteína conservados, os quais formam 8 ou 5 pontes dissulfeto. Estas ligações fornecem resistência ao calor, pH extremo e degradação por proteases. As do tipo L estão presentes em todos os grupos de plantas, enquanto que as de tipo S são restritas à gimnospermas e monocotiledôneas (SHARMA *et al.*, 2022; LIU; STURROCK; EKRAMODDOULLAH., 2010).

Já em relação à sua estrutura tridimensional, as TLPs apresentam três domínios (I, II e III) e uma fenda entre os domínios I e II. O domínio I é um barril beta que forma o núcleo da proteína. De um lado do domínio I, há o domínio II, formado por várias voltas; do outro lado, há o domínio III, formado por uma pequena volta. Cada domínio é estabilizado por pelo menos uma ponte dissulfeto ligadas por até 16 resíduos de cisteína com espaçamento conservado por toda a proteína. O motivo REDDD torna a fenda entre os domínios I e II acídica. Assume-se que esta fenda acídica seja relevante na atividade antifúngica de TLPs,

pois está relacionada com a ligação ao receptor para esta atividade (LIU; STURROCK; EKRAMMODDOULLAH, 2010).

As TLPs são abundantes em plantas, estando presentes desde algas até angiospermas, mas não estão confinadas neste grupo de organismos. Ao longo do tempo, foram reportadas TLPs também em nematóides, fungos e insetos, embora estejam ausentes de vertebrados (DE JESÚS-PIRES *et al.*, 2020). Evolutivamente, as TLPs são agrupadas em nove diferentes famílias (I até IX), cada uma contendo proteínas com relativamente alta identidade. Dentre elas, 5 grupos possuem somente TLPs de plantas (IV, VI, VII, VIII e IX), um grupo contém as TLPs de fungos ascomycota (I), outro contém as de fungos basidiomycota (II), um outro as de nematóides (III) e, por fim, um outro possui as TLPs de insetos (V) (LIU; STURROCK; EKRAMMODDOULLAH, 2010). O conhecimento da presença destas proteínas em animais marinhos ainda é limitado, no entanto existem sequências preditas de TLPs em alguns moluscos bivalves (NIU *et al.*, 2018).

TLPs são frequentemente relacionadas com atividades antifúngicas. Apesar de seu mecanismo de ação na resistência contra fungos não ser bem elucidado, existem algumas hipóteses. Presume-se que ele funcione permeabilizando a membrana plasmática do fungo, culminando no vazamento do conteúdo intracelular (SHARMA *et al.*, 2022). Além disso, sua atividade antifúngica também pode estar relacionada com a atividade de quitinase (NIU *et al.*, 2018); com a hidrólise de (1,3)- $\beta$ -D-glucanos, que são componentes da parede celular de fungos; por atuar como inibidor de algumas enzimas fúngicas, como xilanases; ou por serem indutores de apoptose via receptor de membrana celular (DE JESÚS-PIRES *et al.*, 2020).

Outras funções podem incluir: ação contra infecção bacteriana (LIU; STURROCK; EKRAMMODDOULLAH, 2010), atividade antiviral (WANG; NG, 2002), ação antiinflamatória (ARAUJO SOUSA *et al.*, 2021), aumento da tolerância à seca e aos estresses salino e oxidativo (ULLAH *et al.*, 2018; SINGH *et al.*, 2013) e ação *antifreezing* (YASMIN *et al.*, 2017). *Osmotin* ou *osmotin-like* são nomes dados a membros da família PR-5 que foram especificamente isoladas de plantas sob estresse osmótico ou outros estresses abióticos. Também há o caso da *zeamatin*, cujo nome foi dado em homenagem à espécie em que foi identificada (*Zea mays*) (DE JESÚS-PIRES *et al.*, 2020).

No que diz respeito às aplicações biotecnológicas desta família de proteínas, têm-se em primeiro lugar o uso da Taumatina e Taumatina II como adoçantes e realçadores de sabor, por causa de seu sabor intensamente doce, sendo 100.000 vezes mais doce que a sacarose numa base molar e de 300 a 3000 vezes mais doce que a mesma numa base de massa-para-massa (DE JESÚS-PIRES *et al.*, 2020). Ela também pode ser usada para

modificar o gosto de frutos comestíveis; por exemplo, um pepino transgênico que acumulava a Taumatina II exibiu um fenótipo mais doce e aroma mais aceitável quando comparado com o controle (LIU; STURROCK; EKRAMODDOULLAH, 2010). Salienta-se que a Taumatina é aprovada no Brasil pela ANVISA desde 2008 para o uso como edulcorante em alimentos e bebidas (ANVISA, 2008).

Em segundo aspecto, tem-se o uso dos genes de TLPs na engenharia genética de plantas para torná-las mais resistentes a estresses bióticos e/ou abióticos. Por exemplo, plantas de tabaco expressando uma TLP de amendoim (AdTLP) se mostraram mais resistentes ao ataque de *Rhizoctania solani* que as tipo selvagem; o controle pereceu diante do ataque do patógeno, enquanto que as transformadas ou permaneceram intactas ou atrasaram o desenvolvimento da doença. Similarmente, plântulas deste mesmo tabaco transgênico foram mais tolerantes ao estresse salino e oxidativo que as controle, pois foram capazes de se recuperar e manter uma boa aparência após o período do estresse salino/oxidativo, enquanto que as do tipo selvagem não conseguiram apresentar folhas e um sistema de raízes bem desenvolvido (SINGH *et al.*, 2013).

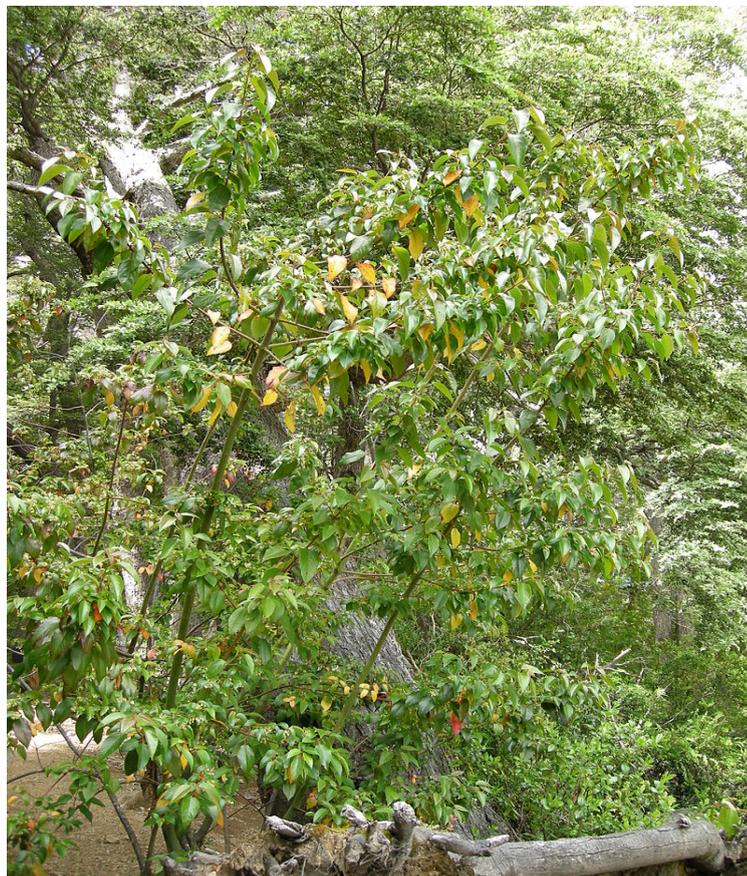
Em terceiro lugar, tem-se a aplicação na indústria de alimentos congelados por causa de sua atividade *antifreezing*. TLPs e outras proteínas *antifreezing* se adsorvem na superfície de cristais de gelo, inibindo a ligação de moléculas de água adicionais. Como um agente *antifreezing*, as TLPs ajudam a controlar a formação de cristais de gelo em uma variedade de alimentos congelados (sorvetes, *frozen yogurts*, *milk-shakes*, dentre outros), desta forma mantendo a consistência macia e cremosa de um produto congelado de alta qualidade (YASMIN *et al.*, 2017).

Por fim, o uso das TLPs também pode ser extrapolado para o de um agente terapêutico, já que o barril beta em sua estrutura parece com a estrutura da adiponectina. Assim, este domínio barril beta das TLPs poderia ser usado como um substituto da adiponectina e atuar como novos agentes terapêuticos no tratamento de desordens relacionadas com a via mediada pelo receptor de adiponectina. Tais desordens incluem doenças cardíacas, diabetes e arteriosclerose, dentre outras (LIU; STURROCK; EKRAMODDOULLAH, 2010).

## 1.2. O Maqui

Nativo do Chile e pertencente à família Elaeocarpaceae, o Maqui (*Aristotelia chilensis* (Mol.) Stuntz) é uma árvore perene com casca macia e muitas ramificações finas e flexíveis,

podendo crescer a até 5 metros de altura (Figura 2). A planta produz uma pequena *berry* comestível, de 5 mm de diâmetro e de cor azul-brilhante a preto-arroxeadado, cuja composição química é predominada por antocianinas, alcalóides indol e flavonóides, mas também possui vitamina C e outros minerais como Br, Zn, Co, Cr, Fe e Mo (TURCHETTI; PAZ, 2019; THOLE *et al.*, 2019; MASOODI; VILLANO; ZAFRILLA, 2019).

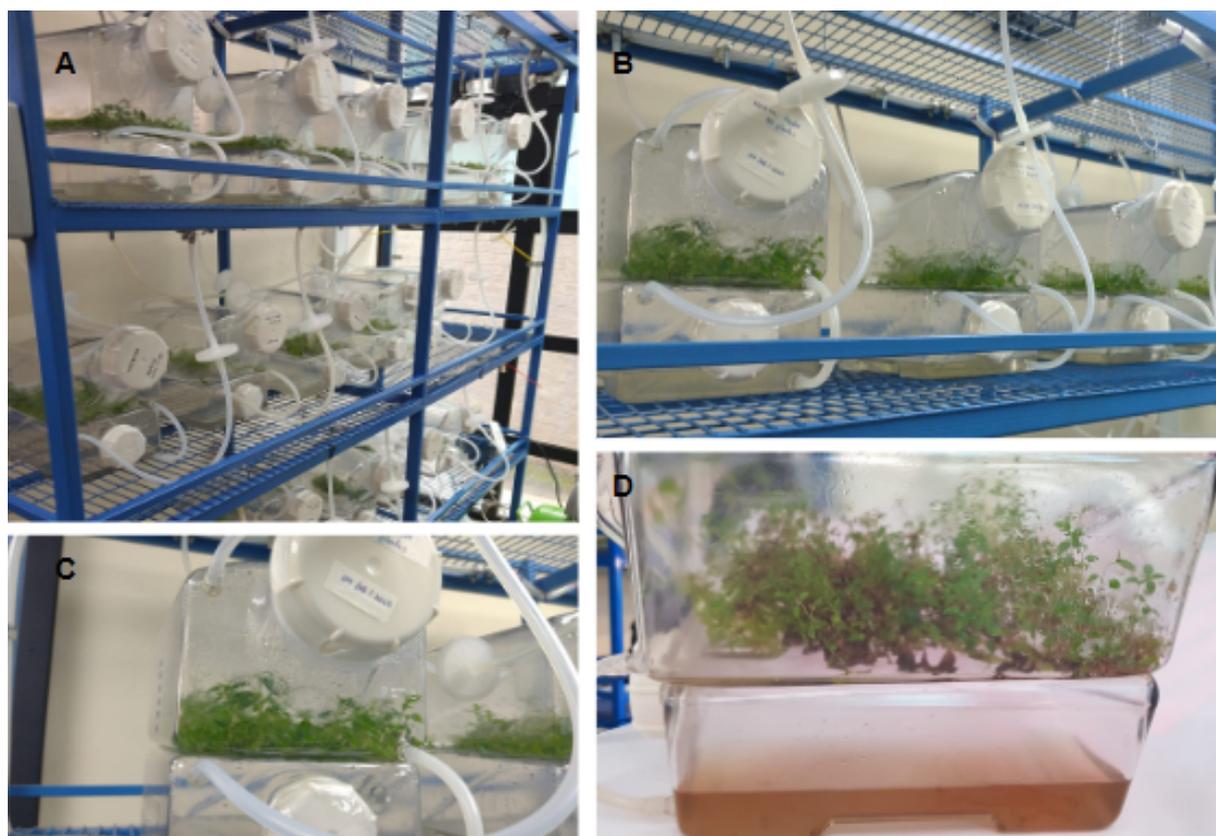


**Figura 2.** Vista geral da árvore *Aristotelia chilensis*, popularmente conhecida como Maqui (Fonte: Wikimedia Commons).

Esta espécie está fortemente associada à cultura chilena. Seus frutos são comestíveis e usados na confecção de doces, geléias, sucos e bebidas alcoólicas, podendo também ser usado como pigmento alimentício orgânico. Além disso, o Chile é um país que utiliza há muito tempo plantas medicinais no conhecimento baseado em experiências pessoais e o Maqui é usado também neste âmbito. Suas folhas podem ser maceradas para fazer cataplasmas em caso de tumores e para efeitos analgêicos e antipiréticos; podem ser usadas como pó para fazer infusões que atuam contra tumores intestinais, úlceras bucais e doenças da garganta; e possuem ação antiinflamatória, antiespasmódica e de relaxamento de fibras musculares. Seu fruto seco ou fresco pode ser usado em infusões para curar diarreia crônica, disenteria e enterite simples. (MISLE *et al.*, 2011).

Os frutos do Maqui possuem alto conteúdo de antocianinas, as quais foram relatadas como sendo as principais responsáveis pelas suas propriedades antioxidantes (MIRANDA-ROTTMAN *et al.*, 2002). Cabe salientar que o fruto do Maqui demonstra alto nível ORAC (*Oxygen radical absorbance capacity Test*) e possui a maior concentração de antocianinas quando comparado com outras *berries* (TURCHETTI; PAZ, 2019). Estudos epidemiológicos sugerem uma forte correlação entre o consumo de alimentos ricos em fenóis e a prevenção de doenças não transmissíveis, por exemplo diabetes, doenças cardiovasculares e câncer. Estes efeitos preventivos são atribuídos aos compostos antioxidantes, como flavonóides e fenilpropanóides, produzidos pelas plantas. As principais fontes desses compostos são as frutas e os vegetais e, dentre elas, os frutos tipo *berry* são as mais ricas (MISLE *et al.*, 2011; MASOODI; VILLANO; ZAFRILLA, 2019). Desta forma, o uso de plantas como fonte de antioxidantes tem crescido significativamente, tendo em vista que contém compostos que são importantes para a indústria nutracêutica por serem a matéria-prima para a produção de suplementos de dieta, aditivos alimentares, drogas herbais, chás, comidas, pigmentos, fragrâncias e saborizantes. Além disso, há uma tendência global crescente de dieta saudável por pessoas que querem ganhar qualidade de vida, gerando um aumento explosivo na indústria de nutrição natural pelo mundo (MISLE *et al.*, 2011). Diante do exposto, o consumo do fruto do Maqui foi associado com benefícios à saúde e rendeu ao Maqui *berry* o título de “Super-fruta”, possuindo efeitos cardioprotetores, antiinflamatórios e antidiabéticos.

Sendo assim, é interessante desenvolver métodos para se obter os metabólitos secundários desta espécie vegetal. O uso de biorreatores de imersão temporária (TIBs, do inglês *Temporary Immersion Bioreactor*) se demonstrou eficaz neste aspecto (Figura 3). Foi-se estabelecida a micropropagação fotomixotrófica de Maqui em TIB, na qual foi induzida a produção de fenilpropanóides pela adição de ABA. Em ambos, material vegetal e meio de cultura, foi observada uma superprodução destes metabólitos antioxidantes (Figura 3D), com atividade biológica similar a de frutas de plantas crescidas *ex vitro* (TRENTINI, 2021).



**Figura 3. Mudanças de *Aristotelia chilensis* micropropagadas em biorreatores de imersão temporária.** A: Panorama geral dos TIBs. B: Vista ampliada dos TIBs. C: Vista mais ampliada dos TIBs. D: Muda de *Aristotelia chilensis* micropropagada em TIB com tratamento com ABA, sendo evidente coloração escura de material vegetal e meio de cultivo, decorrente da produção e liberação de composto fenólicos pela planta (Fotos cedidas pelo Prof. Dr. Ariel D. Arencibia).

Como observado, muito foi descrito a respeito dos extratos botânicos do Maqui e das aplicações de seus metabólitos secundários. Contudo, relatos sobre seus genes e proteínas são limitados. Recentemente, seis cistatinas de Maqui foram descritas (SOUZA, 2022). As proteínas vegetais são essenciais no que se refere à defesa contra patógenos, pragas e estresses abióticos; para a manutenção da homeostase da planta; entre outras funcionalidades. Diante disso, elas apresentam inúmeras aplicações biotecnológicas, tanto em potencial quanto já em prática, abrangendo diversos setores humanos, como a medicina e a agricultura. Neste aspecto, destaca-se as TLPs de *Aristotelia chilensis* e suas possíveis aplicações.

## 2. OBJETIVOS

### 2.1. Objetivo Geral

Identificar, clonar, expressar e realizar ensaios de atividade de *Thaumatococcus-like Proteins* de *Aristotelia chilensis*.

## 2.2. Objetivos Específicos

- Identificar sequências que codificam *Thaumatin-like Protein* no transcriptoma de *Aristotelia chilensis*;
- Realizar análises *in silico* das sequências proteicas;
- Amplificar e clonar dois genes das TLPs nos vetores pPICZaA e pPICZA;
- Expressar as proteínas em *Pichia pastoris* X-33 e purificá-las;
- Realizar ensaios de atividades das proteínas.

## 3. PERCURSO METODOLÓGICO

### 3.1. Identificação e caracterização das sequências de TLPs

#### 3.1.1. Identificação das Taumatinas de Maqui

Plântulas de *Aristotelia chilensis* foram cultivadas em biorreator de imersão temporária em meio MS (GE; SKOOG; , tanto na presença ou ausência de ABA. Foram coletadas amostras do material da planta em 0, 24 e 48 horas de cultura. Toda a plântula foi utilizada para extração de RNA com o reagente TRIZol, seguindo instruções do fabricante. O cultivo em biorreator foi feito no Laboratório de Cultivo y Tejido Vegetal, da Universidade Católica de Maule (UCM), em Talca (Chile), sob coordenação do Professor Doutor Ariel D. Arencibia.

O RNA foi sequenciado utilizando o TruSeq Small RNA Library Preparation Kit (Illumina) na plataforma Hi-Seq (Illumina). Após pré-processamento das *reads*, foi feita a montagem *de novo* do transcriptoma utilizando o Trinity (GRABHERR *et al.*, 2011). A montagem *de novo* teve sua qualidade analisada utilizando a ferramenta BUSCO (SIMÃO *et al.*, 2015). Os *contigs* foram montados a partir do RNA referente à amostra sem tratamento com ABA no início do cultivo. Com a ferramenta OmicsBox, os *contigs* foram usados na busca por sequências de Taumatinas, utilizando como genes modelo dez sequências de TLPs de outras plantas. Utilizou-se sequências de *Prunus persica* (GenBank: JF694926.1), de *Micania micrantha* (GenBank: KR262891.1 e KR262890.1), de *Prunus avium* (GenBank: U32440.1), de *Artemisia annua* (GenBank: PWA97432.1 e PWA46500.1), de *Morus notabilis* (NCBI Reference Sequence: XP\_024028596.1), de *Fragaria chiloensis* (GenBank: ADL16641.1) e também as duas taumatinas originais de *Thaumatococcus danielli* (GenBank: AB265690.1 e J01209.1). O sequenciamento foi realizado no Centro de Genômica Funcional da ESALQ/USP. O processamento das *reads*, montagem *de novo* do transcriptoma e busca pelas sequências foram realizados pela Professora Doutora Milene Ferro, da UNESP, Rio Claro. Assim sendo, foram identificadas quatro sequências de TLPs.

### 3.1.2. Alinhamentos, filogenia, análises *in silico* e estatísticas

As sequências identificadas foram traduzidas com a ferramenta de tradução disponível no website ExPASy (<https://web.expasy.org/translate/>) e a massa molecular e ponto isoelétrico teóricos foram calculados usando a ferramenta ProtParam (<https://web.expasy.org/protparam/>). A presença de peptídeo sinal foi analisada pelo programa SignalP 6.0 (TEUFEL, 2022). A estrutura terciária das proteínas foi modelada com o servidor SWISS-MODEL (WATERHOUSE *et al.*, 2018). Alinhamentos entre proteínas foram feitos usando o software Muscle (EDGAR, 2004). Os alinhamentos foram editados com o ESPrift (ROUBERT; GOUET, 2014). Árvores filogenéticas foram construídas com a ferramenta MEGA X (KUMAR *et al.*, 2018). Quando necessário, os dados foram analisados no software GraphPrism v9.3.1. Significância estatística foi determinada quando  $p < 0,05$ . Comparações entre três ou mais grupos foram feitas com análise de variância (ANOVA) de uma via com o teste de Turkey.

## 3.2. Construção dos cassetes de expressão

### 3.2.1. Síntese do cDNA

A síntese do cDNA a partir do RNA do Maqui foi feita por meio do Kit High Capacity cDNA Reverse Transcription (Applied Biosystems), o qual utiliza oligonucleotídeos aleatórios para iniciar a síntese, seguindo as instruções do fabricante. Resumidamente, na reação de síntese do cDNA foram utilizados 500 ng de RNA total, dNTPs 2,5 mM, tampão da enzima transcriptase reversa 1X, oligonucleotídeo randômico 1X, transcriptase reversa MultiScribe™ 2,5 U/μL e água autoclavada para um volume final de 20 μL. A reação foi realizada em termociclador Veriti (Applied Biosystems) com o seguinte programa: 25°C por 10 min. para hibridização, 37 °C por 120 min. para síntese do cDNA e 85°C por 5 min. para inativação da enzima transcriptase reversa.

### 3.2.2. Amplificação das ORFs das AcTLPs

As ORFs das TLPs foram amplificadas por PCR (*Polymerase Chain Reaction*) a partir do cDNA obtido no item anterior utilizando os oligonucleotídeos apresentados na Tabela 1. A amplificação foi feita no termociclador Veriti (Applied Biosystems) e ocorreu utilizando 5 ng de cDNA, 2,5 U de Taq High Fidelity Pol (Cellco, São Carlos), 10 pmol de cada oligonucleotídeo, 0,2 mM de cada dNTP (Promega), tampão Taq High Fidelity Pol Reaction Buffer 1X (Cellco, São Carlos) em volume final de 25 μL. O ciclo de PCR utilizado foi (i)

95°C por 5 min. para desnaturação inicial; (ii) 35 ciclos de 95 °C por 1min., 50 °C (AcTLP1) ou 55 °C (AcTLP2) por 1min. e 72 °C por 3min. para hibridização e extensão dos oligonucleotídeos; e (iii) extensão final por 10 minutos a 72 °C. Os *amplicons* foram visualizados em gel de agarose 1% corado com brometo de etídio (0,5 µg/mL) no ChemiDoc sob luz UV. As bandas com o tamanho esperado foram recortadas do gel e purificadas com o kit QIAquick gel extraction (Qiagen) de acordo com as instruções do fabricante. Os oligonucleotídeos AcTLP1\_Fw e AcTLP2\_Fw foram usados como *forward* na amplificação para clonagem em plasmídeo de expressão extracelular. Os oligonucleotídeos ZA\_AcTLP1\_F e ZA\_AcTLP2\_F foram usados como *forward* na amplificação para clonagem em plasmídeo de expressão intracelular, possuindo um codon iniciador (ATG) que foi inserido no começo das sequências, já que o pPICZA não possui um codon de início próprio. Os oligonucleotídeos AcTLP1\_Rv e AcTLP2\_Rv foram usados como *reverse* na amplificação para clonagem em vetores de expressão intra e extracelular.

**Tabela 1. Oligonucleotídeos para PCR das sequências de AcTLP1 e AcTLP2.** Em sublinhado está indicado o sítio de restrição das enzimas correspondentes. Em negrito está indicado o codon iniciador ATG adicionado.

Nome	Sequência 5'-3'	Enzima
AcTLP1_Fw	AAAGA <u>ATTTCGCTACTATCACCATC</u> ACAAAC	<i>EcoRI</i>
AcTLP1_Rv	AAAG <u>TCTGACTCCACAGAAAGTTATAACG</u> TAGTT	<i>SalI</i>
ZA_AcTLP1_F	AAAGA <u>ATTTCGGTATGGCTACTATCACCATC</u> ACAAAC	<i>EcoRI</i>
AcTLP2_Fw	AAAGA <u>ATTTCGCCACCTTTGTAATCCG</u> CAAT	<i>EcoRI</i>
AcTLP2_Rv	AAAG <u>TCTGACCTCCTCATTCTTCCCTTCA</u> ACCAT	<i>SalI</i>
ZA_AcTLP2_F	AAAGA <u>ATTTCGGTATGGCCACCTTTGTAATCCG</u> CAAT	<i>EcoRI</i>

Para cada AcTLP foi montado um vetor de expressão intracelular e um de expressão extracelular. O vetor de expressão extracelular (pPICZaA, Invitrogen) possui a sequência do peptídeo sinal de secreção ( $\alpha$ -factor) N-terminal ao sítio múltiplo de clonagem e a sequência para uma cauda de poli-histidina (*his-tag*) C-terminal, o que direciona a proteína recombinante para a rota de secreção, facilitando sua posterior purificação por cromatografia de afinidade em resina contendo metal imobilizado (*Immobilized Metal Affinity Chromatography* - IMAC). Além disso, este vetor também possui o promotor do gene da álcool oxidase 1 (AOX1), o qual é induzível por metanol, e o gene de resistência à Zeocin™

(Anexo I). O vetor de expressão intracelular (pPICZA, Invitrogen) apresenta todas as características do pPICZaA, exceto o  $\alpha$ -factor e portanto a proteína expressa é mantida no interior celular (Anexo I) (EasySelect Pichia Expression Kit, 2001).

### 3.2.3. Clonagem das ORFs das AcTLPs

Os *amplicons* e os plasmídeos pPICZaA e pPICZA foram clivados com as enzimas de restrição *EcoRI* e *SaII* (Thermo Scientific). Para tal, foram realizadas reações utilizando 2  $\mu$ g de DNA, 10 U de cada enzima de restrição, tampão Orange 1X (Thermo Scientific) em volume final de 20  $\mu$ L. As reações foram incubadas a 37 °C por 16 h e, após, as enzimas de restrição foram inativadas por aquecimento a 65 °C por 20 minutos. As reações foram resolvidas em gel de agarose 1% corado com brometo de etídio (0,5  $\mu$ g/mL) e visualizadas sob luz UV. As bandas referentes ao vetor clivado e gene de interesse foram recuperadas do gel com o kit QIAquick gel extraction (Qiagen). Após, os insertos foram ligados aos plasmídeos com os seguintes reagentes: 2,5 U da enzima T4 DNA Ligase (Invitrogen), tampão da ligase 1X (Invitrogen), inserto e plasmídeo na proporção 1:5 plasmídeo:inserto e água para volume final de 20  $\mu$ L de reação, esta que foi incubada por 16 h a 4°C.

Cada reação de ligação foi usada separadamente para transformar, por choque térmico, células de *Escherichia coli* DH5 $\alpha$  competentes preparadas por tratamento com CaCl<sub>2</sub> previamente preparadas (SAMBROOK; RUSSEL, 2000). As células foram incubadas em meio Luria-Bertani (LB) *low salt* (Apêndice I) suplementado com 25  $\mu$ g/mL de Zeocin<sup>TM</sup>. Após crescimento das colônias, feito uma PCR de colônia em termociclador Veriti utilizando tampão 1X Taq Buffer (Thermo Scientific), 3 mM de MgCl<sub>2</sub>, 0,2 mM de cada dNTP, 10 pmol de cada um dos oligonucleotídeos comerciais AOX5' e AOX3' e 1,25 U da enzima Taq Polimerase (Thermo Scientific). A reação ocorreu em 25  $\mu$ L e como molde foram usadas colônias picadas com palito estéril. O protocolo de amplificação usado foi (i) 95 °C por 5 min.; (ii) 35 ciclos de amplificação de 30 s a 95 °C, 1 min. a 50 °C e 1 min. a 72 °C; e (iii) extensão final de 7 min a 72 °C. Os *amplicons* foram analisados em gel de agarose 1% corado com brometo de etídio (0,5  $\mu$ g/mL).

Uma das colônias positivas foi inoculada em 5 mL de LB *low salt* com 25  $\mu$ g/mL de Zeocin<sup>TM</sup> e mantida *overnight* a 37 °C com agitação constante. O DNA plasmidial foi extraído desta cultura com o kit QIAprep Spin MiniPrep (Qiagen), seguindo instruções do fabricante. Os plasmídeos foram submetidos a uma clivagem com as enzimas de restrição *EcoRI* e *SaII* nas mesmas condições da clivagem anterior. A restrição foi resolvida em gel de agarose 1% corado com brometo de etídio (0,5  $\mu$ g/mL) para confirmação da clonagem.

### 3.2.4. Amplificação e clonagem da GFP

A ORF da GFP (*Green Fluorescent Protein*) também foi amplificada e clonada nos plasmídeos pPICZαA e pPICZA, com a finalidade de serem transformados em *Pichia pastoris*, as quais foram usadas para fins científicos neste estudo e também se pretende usá-las futuramente para fins didáticos. A amplificação ocorreu igual a reação de amplificação do item 3.2.2., somente substituindo-se os oligonucleotídeos da Tabela 1 com os apresentados na Tabela 2 e utilizando como molde o plasmídeo pEGFP-N1 (Clontech). A clonagem da ORF da GFP nos vetores de expressão se deu igual ao descrito no item 3.2.3..

**Tabela 2. Oligonucleotídeos para PCR da sequência da GFP.** Em sublinhado está indicado o sítio de restrição das enzimas correspondentes. Em negrito está indicado o codon iniciador ATG adicionado.

Nome	Sequência 5'-3'	Enzima
pPIC-GFP-F	AAAGA <u>ATTCGGTAT</u> <b>GGT</b> GAGCAAGGGGCGAGGAG	<i>EcoRI</i>
pPIC-GFP-R	AAAG <u>TCGACCTT</u> GTACAGCTCGTCCATGCCG	<i>SaII</i>

Após construção, os plasmídeos pPICZαA-GFP e pPICZA-GFP foram transformados em *Pichia pastoris* X-33 como detalhado no item 3.3.1.. Também foi feito um *screening* de clones produtores de acordo com o descrito no item 3.3.2.. Posteriormente, pretende-se utilizar estes clones de *Pichia pastoris* produzindo GFP para fins didáticos, para ilustrar em aula, dentre tantos temas, o papel de um peptídeo sinal, as diferenças entre expressão intracelular e secretada, a produção e purificação de proteínas recombinantes etc..

## 3.3. Produção recombinante das AcTLPs

### 3.3.1. Transformação de *Pichia pastoris*

O preparo das células competentes foi realizado conforme descrito (CREGG, 2007). Uma colônia de *Pichia pastoris* X-33 foi inoculada em 5 mL de meio YEPD (Apêndice I), onde cresceu a 30 °C e agitação de 250 rpm durante 8 horas. Após, 50 µL desta cultura foi adicionada a 50 mL de YEPD e mantidas nas mesmas condições até atingir a D.O.<sub>600</sub> de 1-1,5 (aproximadamente 16h). A cultura foi dividida em dois frascos de 30 mL e centrifugada a 1500 x g por 5 minutos a 4 °C. O sobrenadante foi descartado e as células ressuspendidas em 5 mL de YEPD e 1 mL de HEPES 1 M, seguido da adição de 0,125 mL de DTT 1 M. A solução então foi incubada a 30 °C por 15 minutos. Em seguida, o volume foi completado

para 25 mL com água autoclavada gelada e as células centrifugadas nas mesmas condições anteriores. O sobrenadante foi descartado e as células ressuspensas em 12,5 mL de água autoclavada gelada, seguido de centrifugação e descarte do sobrenadante. O resultado foi ressuspendido com 1 mL de sorbitol 1M e novamente centrifugado e o sobrenadante descartado. Por fim, o pellet foi então ressuspendido em 50  $\mu$ L de sorbitol 1 M e distribuído em alíquotas de 40  $\mu$ L.

Para a transformação da *P. pastoris*, os plasmídeos pPICZ $\alpha$ A-AcTLP1 e pPICZ $\alpha$ A-AcTLP2 foram linearizados com a enzima *PmeI* (Fermentas) em tampão Blue 1X a 37 °C *overnight*. Aos 40  $\mu$ L de células, foi adicionado 1  $\mu$ g de plasmídeo linearizado e o volume colocado em uma cubeta de eletroporação de 0,2 cm. A cubeta foi incubada em gelo por 5 minutos, enxuta e submetida à eletroporação, sob as condições de 1,25 kV, 25 F e 200  $\Omega$  no equipamento Gene Pulser II (Bio-Rad®). Imediatamente após, foi adicionado 1 mL de sorbitol 1 M e as células foram incubadas, sem agitação, a 30 °C por 2 horas. Por fim, 200  $\mu$ L do conteúdo foi plaqueado em placas contendo meio YPDS (Apêndice I) com diferentes concentrações de Zeocina™ (100, 250 e 500  $\mu$ g/mL), as quais foram incubadas a 30 °C até a formação de colônias (2-3 dias).

Após crescimento, uma pequena quantidade de massa de colônias foi picada com um palito estéril e homogeneizada em 20  $\mu$ L de SDS 0,2%, fervida por 20 minutos e depois centrifugada por 1 minuto a 13.000 xg. Um microlitro do sobrenadante foi utilizado como molde para uma PCR de colônia utilizando tampão 1X Taq Buffer (Thermo Scientific), 3 mM de MgCl<sub>2</sub>, 0,2 mM de cada dNTP, 10 pmol de cada um dos oligonucleotídeos comerciais 5'AOX1 (5'-GACTGGTTCCAATTGACAAGC-3', Invitrogen) e 3'AOX1 (5'-GCAAATGGCATTCTGACATCC-3', Invitrogen), 1,25 U da enzima Taq Polimerase (Thermo Scientific) e 1% de Triton X-100 em 25  $\mu$ L de reação. A PCR ocorreu em termociclador Veriti e o protocolo de amplificação usado foi (i) 95 °C por 5 min.; (ii) 35 ciclos de amplificação de 30 s a 95 °C, 1 min. a 50 °C e 1 min. a 72 °C; e (iii) extensão final de 7 min a 72 °C. Os *amplicons* foram analisados em gel de agarose 1% corado com brometo de etídio (0,5  $\mu$ g/mL).

### 3.3.2. *Screening* de clones produtores

Uma vez confirmadas por PCR, as massas celulares das colônias transformadas com pPICZ $\alpha$ A-AcTLP1 ou pPICZ $\alpha$ A-AcTLP2 foram cultivadas em uma placa de 24 poços (Whatman®). Utilizou-se como controle negativo a cepa de *P. pastoris* X-33 não-transformada e como positivo uma cepa que expressa a CaneCPI-5 previamente

produzida (SANTIAGO et al., 2017). Inicialmente, em cada poço foram colocados 3 mL de meio BGMV (Apêndice I). A placa foi mantida por 48 horas sob agitação de 250 rpm e 28 °C para crescimento de massa celular. Após, ela foi centrifugada a 1500 x g por 10 minutos e o sobrenadante descartado. As células foram ressuspensas em 2 mL do meio BMMV (Apêndice I) e mantidos sob agitação de 250 rpm a 28°C. A cada 24 horas, a placa foi centrifugada (1500 xg por 5 minutos), uma alíquota de 100 µL de sobrenadante foi retirada e 15 µL de metanol 100% foi adicionado ao meio para concentração final de 0,75%. O processo foi monitorado durante 144 horas. Dos sobrenadantes recuperados, 20 µL foram tratados em condições desnaturantes com tampão de amostra Laemmli 3X (Apêndice I) e fervidos por 7 minutos. As proteínas foram separadas em SDS-PAGE 12%, o qual foi corado com Comassie Blue e utilizado para analisar o tempo de expressão e secreção das TLPs.

### 3.3.3. Indução em maior escala e purificação

As colônias previamente selecionadas, expressando as TLPs recombinantes, foram colocadas em 10 mL de meio BMGY para o pré-inóculo, sob agitação de 250 rpm a 30 °C por 24 h para crescimento de massa celular. Após, o pré-inóculo foi transferido para um Erlenmeyer com 500 mL de meio BMGY, novamente a 250 rpm e 30 °C por 24 h para crescimento celular. A cultura foi então centrifugada por 5 minutos a 3000 x g a 4 °C. O sobrenadante foi descartado e o precipitado ressuspensado em 100 mL de meio BMMV. Este foi colocado sob agitação de 250 rpm e 30°C por até 72 h. Alíquotas de 1 mL foram coletadas a cada 24 horas, sendo também adicionado metanol para concentração final de 0,75% diariamente. As alíquotas foram armazenadas a -20 °C até o final da indução.

Após, as amostras foram centrifugadas a 13.000 xg em microcentrífuga por 5 minutos e o seu sobrenadante recuperado. Os mesmos foram tratados com tampão de amostra Laemmli 3X e submetidos a 95 °C por 7 minutos, sendo analisadas em SDS-PAGE 12%. Após o tempo de indução, o meio de cultura foi centrifugado em centrífuga Sorvall® a 3000 xg e 4 °C por 5 minutos, o *pellet* foi descartado e o sobrenadante filtrado a vácuo em filtro PVDF 0,45 µm (Millipore).

As TLPs recombinantes têm em sua região C-terminal uma cauda de histidinas, o que torna possível sua purificação por cromatografia de afinidade em resina de Ni-NTA Superflow (Qiagen). A coluna pré-equilibrada com 5 volumes de tampão de equilíbrio (Apêndice I) foi adicionado o sobrenadante da cultura e o eluato recolhido. Após a ligação da proteína na resina, a coluna foi lavada com 3 volumes de tampão de equilíbrio, seguido da eluição da proteína em tampão de eluição (Apêndice I) com concentrações crescentes de

imidazol (10, 25, 50, 75, 100 e 250 mM), sendo dois volumes das concentrações de 10 a 100 e três volumes da solução com 250 mM de imidazol. As amostras eluídas foram analisadas em SDS-PAGE 12% e coradas com Coomassie Blue. As proteínas purificadas foram dialisadas em membrana SnakeSkin™ contra 2 L de tampão de diálise (pH 8,0) a 4 °C, sendo realizadas três trocas de tampão a cada 1 hora de incubação. As proteínas foram filtradas em filtros Millipore de 0,22 µm e quantificadas com o Kit BCA Protein Assay (Thermo Scientific) seguindo as instruções do fabricante.

### 3.8. Ensaio de atividades das TLPs

#### 3.8.1. Ensaio de glucanase

Cem microgramas de proteína foram misturados com Laminarina (concentração final 1% (w/v); Sigma), tampão acetato de sódio 50 mM (pH 5,5) e água para volume final 100 µL. As reações foram incubadas a 40 °C por 30 minutos. Após, 200 µL de solução DNS foram adicionados, fervidos a 100 °C por 10 minutos e finalmente a reação foi interrompida por resfriamento em gelo durante 2 minutos. Subsequentemente, foi feita a leitura da absorbância da solução a 540 nm. Controles com somente Laminarina ou somente TLP também foram feitos. Todas as reações foram feitas em triplicata. Uma curva padrão de glicose foi construída como referência. Uma unidade (U) de atividade enzimática foi definida como a quantidade de enzima necessária para liberar 1 µmol de glicose por minuto nas condições definidas (LUCENA *et al.*, 2013; LIU *et al.*, 2021).

#### 3.8.2. Permeabilização de Membrana Fúngica

O ensaio de permeabilização de membrana plasmática de fungos foi feito com um co-cultivo de uma *Pichia pastoris* produzindo a GFP internamente e uma outra *P. pastoris* produzindo e secretando as AcTLPs. Resumidamente, foram feitos pré-inóculos em 10 mL de BMGY de *P. pastoris* transformada com o vetor pPICZA-GFP e deixado a 30 °C sob agitação constante por 24 h. Também foram feitos pré-inóculos de *P. pastoris* utilizadas para a expressão das AcTLPs (pPICZαA-AcTLP1 ou pPICZαA-AcTLP2) separadamente. Como controle, uma *P. pastoris* X-33 não transformada também foi inoculada. Depois, as *P. pastoris* X-33, com pPICZαA-AcTLP1 e com pPICZαA-AcTLP2 foram transferidas para 500 mL de BMGY separadamente. Em cada um dos inóculos anteriores, *P. pastoris* com pPICZA-GFP também foi adicionada, desta forma tendo-se três co-cultivos de *P. pastoris* produzindo GFP internamente, um controle (GFP+X33) e dois testes (GFP+AcTLP1 e GFP+AcTLP2). O conteúdo foi deixado a 30 °C, 250 rpm, por 24 h. Por fim, as células foram centrifugadas a

1500 xg por 5 min a 4 °C, o sobrenadante foi descartado e o precipitado foi ressuspenso em 100 mL de BMMY. Metanol 100% foi adicionado para concentração final 0,75%, uma alíquota de 1 mL (não induzida) foi recolhida e as culturas foram deixadas a 30 °C e 250 rpm.

O co-cultivo seguiu de 0 a 144 h, com adição de metanol para final 0,75% e retirada de 1 mL de alíquota a cada 24 h. As alíquotas foram centrifugadas a 13.000 xg por 5 min, o pellet foi descartado e o sobrenadante armazenado. As alíquotas foram usadas para análise em SDS-PAGE 12%. Além disso, também foram usadas para leitura da possível fluorescência da GFP em espectrofluorímetro Hitachi F2500 com excitação a 475 nm e emissão a 503 nm, utilizando o sobrenadante das culturas diluído 1:30 com água destilada.

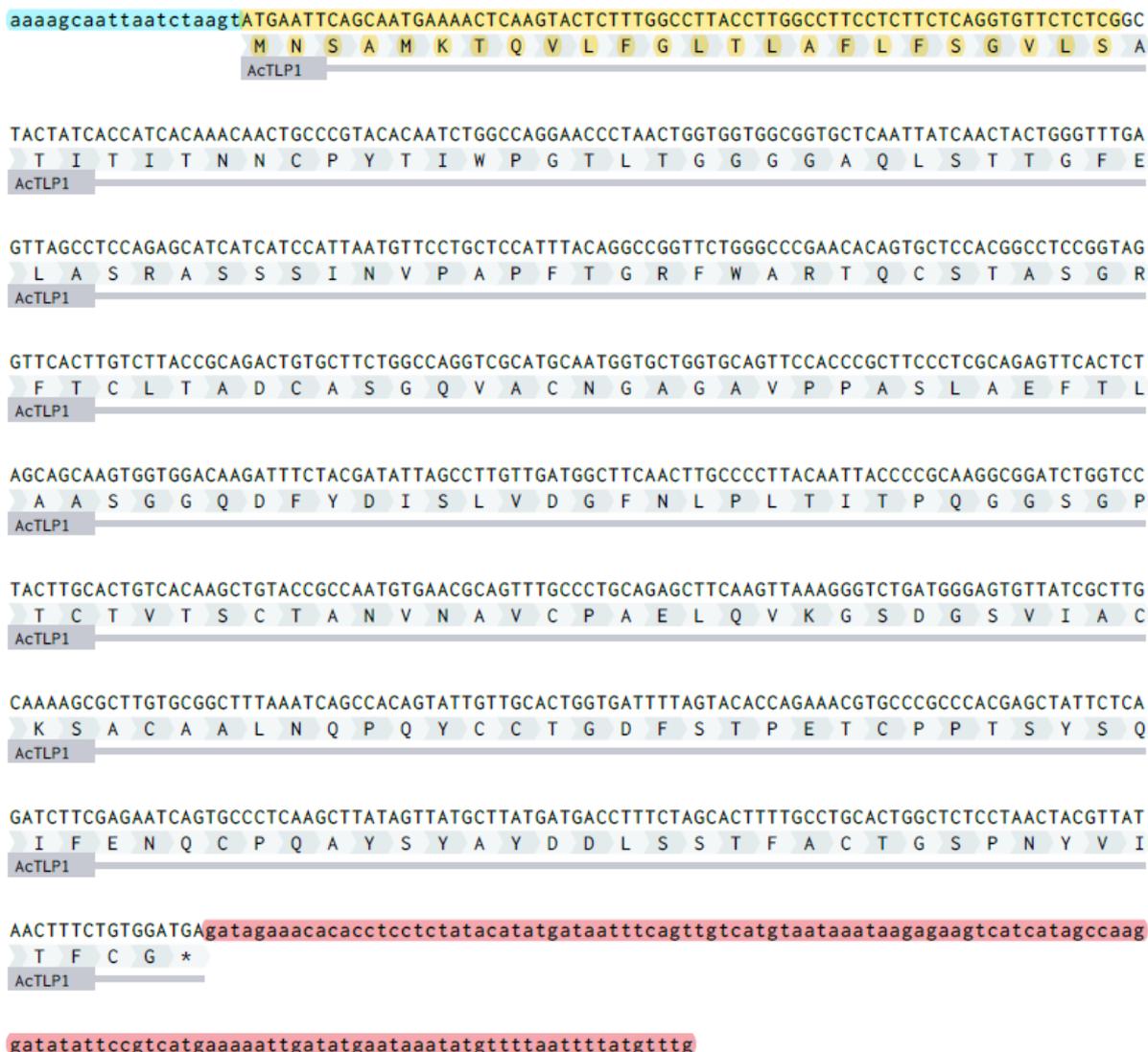
### 3.8.3. Resistência ao Estresse Osmótico

*Pichia pastoris* X-33 contendo os plasmídeos pPICZA-AcTLP1, pPICZA-AcTLP2, pPICZaA-AcTLP1 ou pPICZaA-AcTLP2 foram inoculadas separadamente em 3 mL de YEPD e incubadas a 30 °C, 250 rpm, *overnight*. Um controle de *Pichia pastoris* X-33 não transformada também foi inoculado. Cem microlitros de cada cultura crescida *overnight* foram inoculados em 10 mL de YEPD e deixados a 30 °C e 250 rpm até que a D.O.<sub>600</sub> estivesse entre 0,3 e 0,4. Após, cada cultura foi diluída em 10, 100, 1000, 5000 e 10.000 vezes e 5 uL de cada diluição seriada foi inoculada em meio YEPD ágar contendo 0, 1, 1,5 ou 2 M de Sorbitol. O crescimento foi monitorado por 3 dias a 30 °C. Em metade das placas, 200 uL de metanol 100% foi adicionado todos os dias na tampa da placa invertida para manter a expressão dos genes das taumatinas (JANGPROMMA *et al.*, 2014).

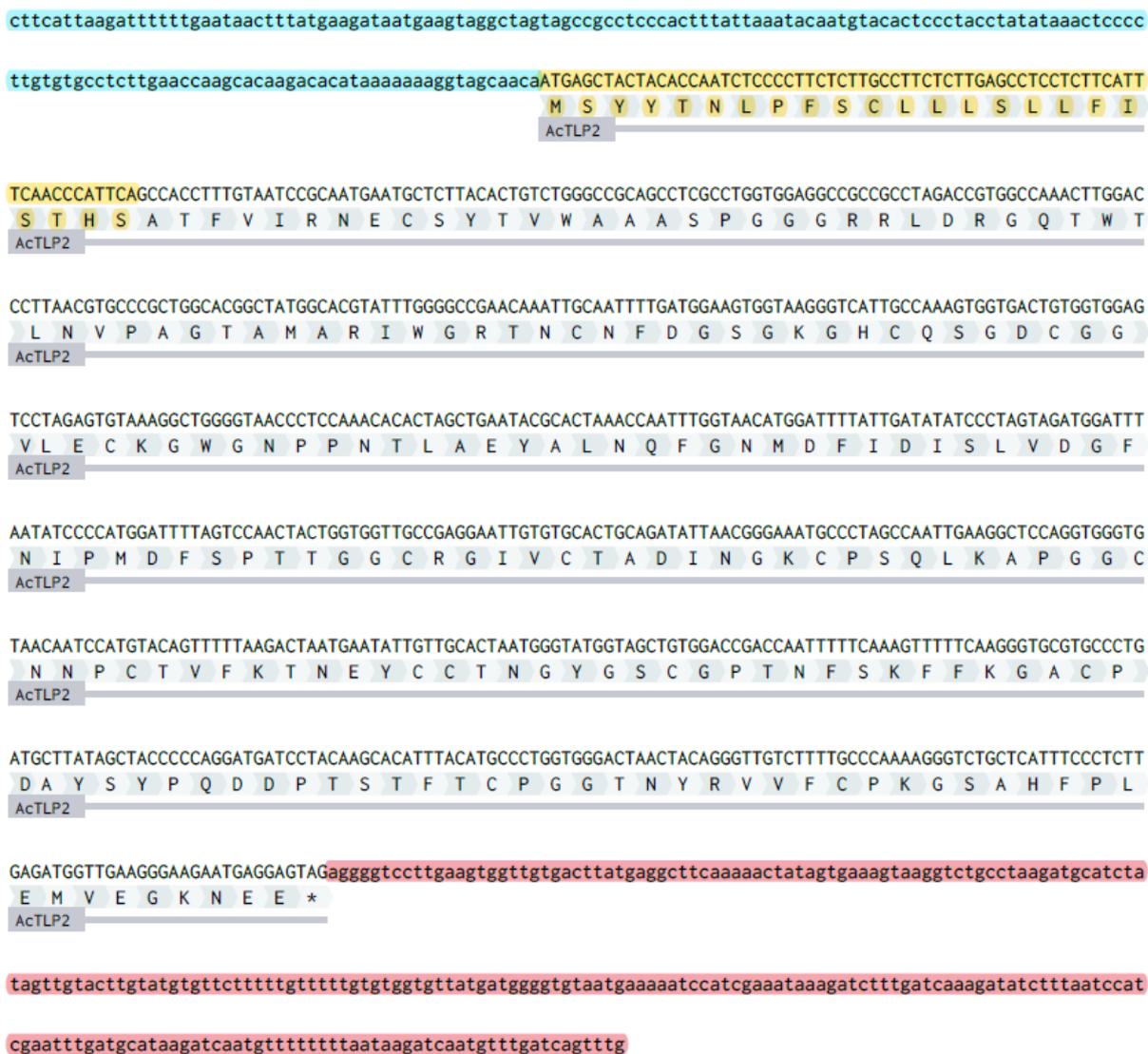
## 4. RESULTADOS E DISCUSSÃO

### 4.1. Alinhamentos e análises filogenéticas

As sequências de nucleotídeos das AcTLPs, incluindo regiões não traduzidas (UTR, do inglês *Untranslated Region*) 3' e 5', e também sua tradução para sequência de aminoácidos, estão disponíveis nas figuras 4 a 7 e também no Apêndice II.



**Figura 4. Sequência de nucleotídeos e de aminoácidos da AcTLP1.** A região codificante da proteína é mostrada a partir do codon de iniciação ATG e se encerra no codon de terminação TGA, indicado por um asterisco (\*). Destacados em amarelo estão os resíduos do peptídeo sinal putativo da proteína e sua região codificante. Em letra minúscula e destacado azul está a UTR 5'. Em letra minúscula e destacado vermelho está a UTR 3'.



**Figura 5. Sequência de nucleotídeos e de aminoácidos da AcTLP2.** A região codificante da proteína é mostrada a partir do codon de iniciação ATG e se encerra no codon de terminação TAG, indicado por um asterisco (\*). Destacados em amarelo estão os resíduos do peptídeo sinal putativo da proteína e sua região codificante. Em letra minúscula e destacado azul está a UTR 5'. Em letra minúscula e destacado vermelho está a UTR 3'.

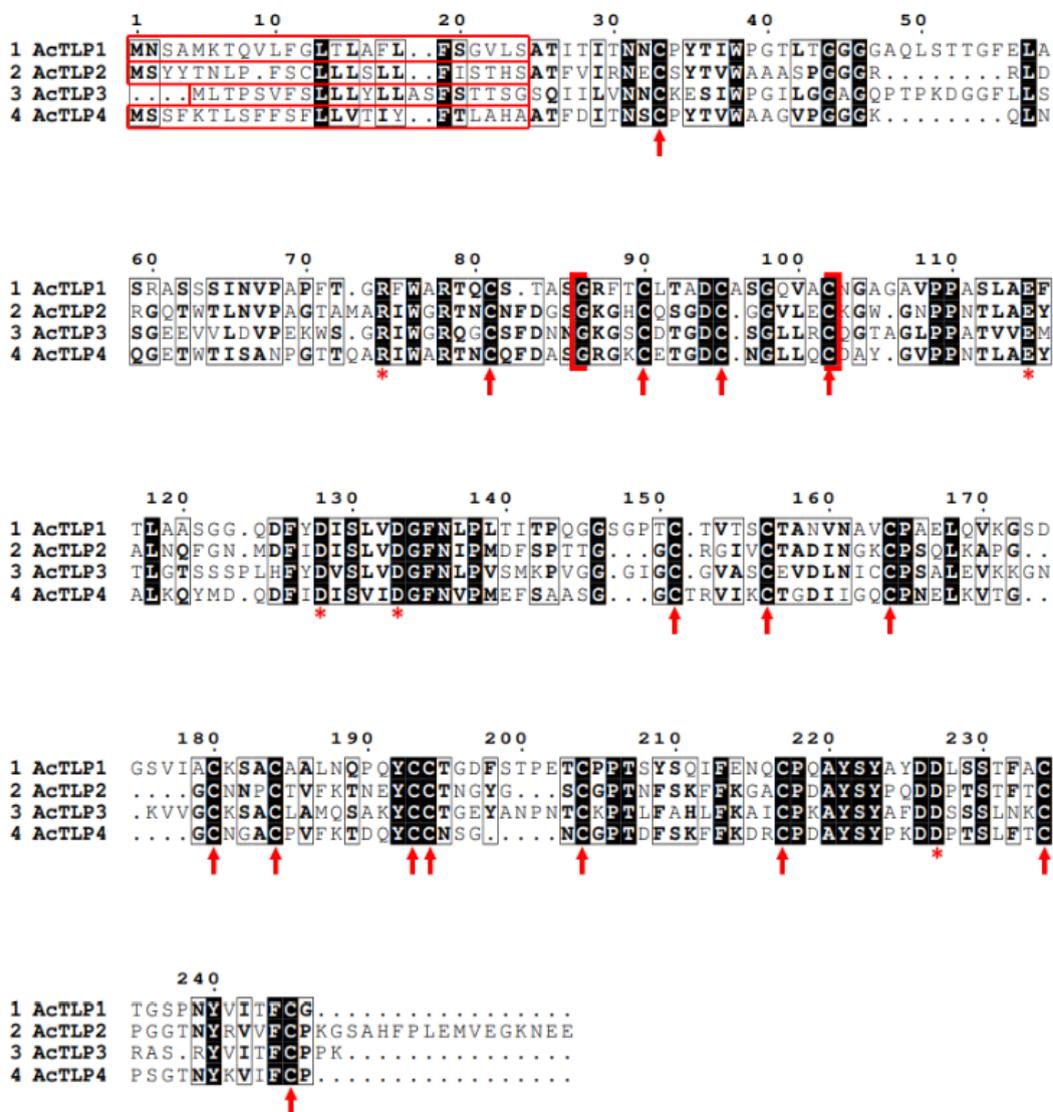




**Figura 7. Sequência de nucleotídeos e de aminoácidos da AcTLP4.** A região codificante da proteína é mostrada a partir do codon de iniciação ATG e se encerra no codon de terminação TGA, indicado por um asterisco (\*). Destacados em amarelo estão os resíduos do peptídeo sinal putativo da proteína e sua região codificante. Em letra minúscula e destacado azul está a UTR 5'. Em letra minúscula e destacado vermelho está a UTR 3'.

As sequências das AcTLPs foram alinhadas usando o software Muscle (EDGAR, 2004) e editadas com o programa ESPript 3 (ROBERT; GOUET, 2014), resultando no alinhamento da Figura 8. No alinhamento estão destacadas as prováveis regiões de peptídeo sinal, as 16 cisteínas conservadas que fazem as 8 ligações dissulfeto das proteínas, o motivo REDDD conservado e a assinatura familiar das taumatinas. Evidencia-se que as AcTLP1, AcTLP3 e AcTLP4 possuem a assinatura familiar como o esperado, G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C, mas na AcTLP2 há uma mudança de uma treonina (T) para uma serina (S), permanecendo com a assinatura da seguinte maneira: G-x-[GF]-x-C-x-S-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C. A mudança é sutil, visto que treonina e

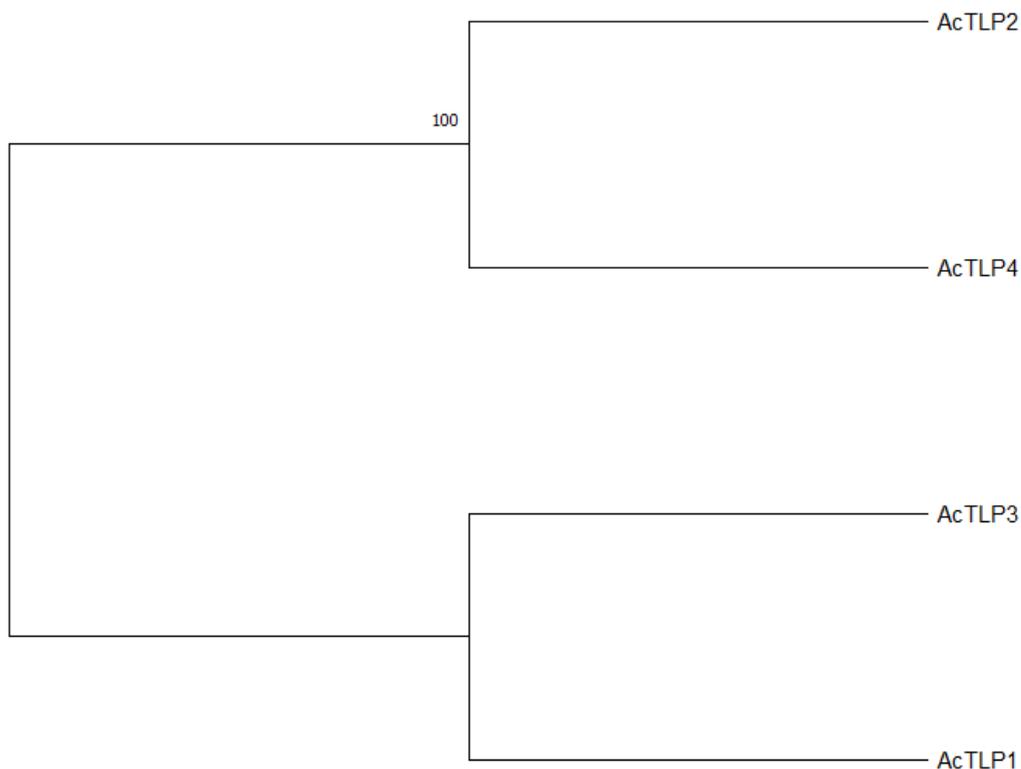
serina são aminoácidos similares, porém não se sabe as implicações desta substituição na função da proteína.



**Figura 8. Alinhamento global das AcTLPs.** As caixas em vermelho indicam as regiões dos prováveis peptídeos sinais identificadas pelo SignalP 6.0. Os colchetes indicam a assinatura familiar das taumatinas (G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C). As setas em vermelho mostram os dezesseis resíduos de cisteínas conservados, responsáveis pelas oito ligações dissulfeto que estabilizam a estrutura da proteína. Os asteriscos representam as posições conservadas dos cinco aminoácidos do motivo REDDD, responsável pela fenda ácida entre os domínios I e II da proteína. Os resíduos considerados altamente similares estão destacados em negrito e dentro de uma caixa vazia de contorno preto. Nos resíduos com identidade estrita, a coluna é colorida com fundo preto.

A análise filogenética das TLPs de Maqui foi feita a partir do alinhamento da Figura 8, gerando a árvore filogenética não enraizada da figura 9. Esta árvore foi construída

utilizando o método de verossimilhança com o modelo Dayhoff+G(4), com a qualidade da construção sendo verificada por *bootstrap* de 1.000 ciclos.



**Figura 9. Árvore filogenética das AcTLPs.** A inferência filogenética foi feita pelo método da máxima verossimilhança utilizando o modelo Dayhoff+G(4). Os valores de *bootstrap* estão indicados nos nós como valores de porcentagem.

Percebe-se que as AcTLPs foram divididas entre dois clados, um composto pelas AcTLP2 e AcTLP4 e outro pelas AcTLP1 e AcTLP3, sendo possivelmente duas famílias. As AcTLPs também foram alinhadas com sequências TLPs de outras espécies, com sequências retiradas do trabalho de Liu, Sturrock e Ekramoddouh (2010), resultando no alinhamento apresentado na figura 10. As sequências utilizadas estão disponíveis no Anexo II. Foram usadas sequências de três plantas dicotiledôneas membros do clado das rosídeas, o qual o Maqui também pertence: (i) uma TLP de *Arabidopsis thaliana* (NP\_173261.1); (ii) uma TLP de *Cephalotus follicularis* (GAV58933.1), espécie da mesma ordem do Maqui (Oxalidales); e (iii) uma TLP de *Manihot esculenta* (XP\_021601423.2). Também foram usadas duas

sequências de plantas monocotiledôneas: (i) a Taumatina original de *T. danielli* (BAF44567.1); e (ii) uma TLP de *Oryza sativa* subsp. *japonica* (NP\_001049473.1). Como representantes de membros vegetais não-angiospermos foi usado uma TLP de uma gimnosperma (*Pinus monticola*; ADB97925.1) e de uma briófito (*Physcomitrella patens*; XP\_001765997.1). Por fim, como representantes de membros fora do reino Plantae, têm-se uma sequência de um fungo Ascomycota (*Aspergillus nidulans*; XP\_660683.1) e de uma de um Nematódeo (*Caenorhabditis elegans*; NP\_502360.1).

Para investigar se as AcTLPs pertencem a diferentes famílias de proteínas, utilizou-se do estudo de Liu, Sturrock e Ekramoddoullah (2010). Estes autores dividiram as TLPs em 9 diferentes famílias (I a IX), sendo que cinco destas (IX, XI, XII, XIII e IX) são de TLPs vegetais, enquanto as outras quatro são uma de insetos, outra de nematóides, uma de fungos ascomycota e mais uma de fungos basidiomycota. As AcTLPs foram alinhadas com representantes das cinco famílias de TLPs vegetais, cujas sequências podem ser encontradas no Anexo II. Este alinhamento foi utilizado para análise filogenética, gerando a árvore não enraizada da figura 11. Esta árvore foi construída pelo método de Neighbour-Joining, com modelo de Poisson, com a qualidade da construção sendo verificada por *bootstrap* de 1.000 ciclos.

	1	10	20	30	40	50
1 AcTLP1	..MNSAMKTQVLFGLTLAF...LF..SGVLSA	TI.TITNNCPYTI..	WPGTLTGGGGAQL			
2 AcTLP2	..MSYYTNLPFSCLLLSL...LF..ISTHSA	TF.VIRNECSYTV..	WAAASPG.....			
3 AcTLP3	.....MLTPSMFSLLLYL...LASFTTSGS	QI.LLVNCKESI..	WPGILGGAGQPTP			
4 AcTLP4	..MSSFKLSEFSSFLLV...TIYFTLAHA	TF.DITNSCPYTV..	WAAAGVPG.....			
5 T_danielli	..MAATTCFFFLFPF...LLLLTLRRAA	TF.EIVNRCPSYTV..	WAAASKG..DAAL			
6 C_follicularis	..MHSAMKTQVIFGLTLAF...LL..AGVPSA	TV.TITNNCPYTI..	WPGTLTGGGQNL			
7 M_esculenta	..MSSFNISFFLVAA...LYYFTFAHAA	TF.DITNKCPYTV..	WGAASPG.....			
8 A_thaliana	..MASINLFLFAF...LLLLSHASAS	TV.IFYNNCKHEV..	WPGIQPSAGQNL			
9 O_sativa	MEFGRTPALQIIVGVLSQ...LQFGABAVGT	VFTLRNNCTYTV..	WPATLSGNTAVAV			
10 P_monticola	MGLILRSRALALLLILFSGGAI	MVICAKTSPLEI.T	IMNSCPTI..	WPGLOQASEGHDVL		
11 P_patens	..MMLVMKIKSYI...LDHALCABGV	TF.FINHCKFTV..	WVGVPQNGGLPLL			
12 A_nidulans	MMFTKALVAATLATALAALPQP	TVVRRREGGDAGV.T	IVNMDSDVYAW	SVTDGVSVMHTL		
13 C_elegans	..MALLIKLTLAV...LLLALGAEAR	TI.TIYNKCPFTI..	WPGILGPGN...P			



	60	70	80	90	100
1 AcTLP1	STIGFELASRASSSINVPAPFT	.GRFWARTQCS.TAS	GRFTCLTADCA	SGQVACNGAGAV	
2 AcTLP2	..GRRRLDRGQTWTLNVPAGTA	MARIWGRTRNCNFDGS	GKGGHCOGSDC	.GGVLECKGWN	
3 AcTLP3	KDGGFLSSGSEVVLVDVPEKWS	.GRIWGRQCQSFDDN	GKGSCTDGGC	.SGLLRCCGTAGL	
4 AcTLP4	..GGKQLNQGETWMTISANPGIT	QARIWARTNCQFDAS	GRGKCEGTGDC	.NGLLQCDAYGV	
5 T_danielli	DAGGRQLNSGSEWNTINVEPGIT	GGKIWARTDCYFDD	SSGICETGDC	.GGLLRCRFRG.R	
6 C_follicularis	STIGFELASGASNSFNVPFFFT	.GRFWARTQCS..TF	GAVTCATADCA	SGQVACNGAGAI	
7 M_esculenta	..GCRELKTGETWMTITANPGIT	QARIWARTNCQFDAS	GKGGKCEGTGDC	.NGLLVCCQYGA	
8 A_thaliana	AGGGFKLPANKAHSLLPLPWS	.GRFWGRHGCTFDRS	GRGHCATGDC	.GGSLSCNGAGGE	
9 O_sativa	GGGGFELSPGANVSFPAPAGWS	.GRLWARTDCAPSGT	ASLACVETGDC	.GGVLSCSLGGGA	
10 P_monticola	EQGGFALLESLSRSFSVANPWN	.GVVWARTCCSFSG	EKGSCTDGGC	.DGGKLCNGTGGR	
11 P_patens	VGGGFELAAGKQDAVTAASASNG	.GRFWGRTGCKFDSA	GKGNCEGTGDC	.GGVLLKCGGAGGN	
12 A_nidulans	SSGG...GSYTENFQANPN	..	..	..GGVLSIKLSTHQ	
13 C_elegans	SGGGFKLNAGQSRDINVDDAWT	AGRWARTGCD...GNFN	CEGTGFC	.RMSERCCNGAGGV	



	110	120	130	140	150	160
1 AcTLP1	PPASLAEFILAASGG.QDFYDISLVDGFNPL	TITPQGG..SGPT	CIVT.SCIANVNAVC			
2 AcTLP2	PPNTLAEYALNQFGN.MDFYDISLVDGFNIP	DFSPITG....GCRGI	VCIADINCKC			
3 AcTLP3	PPATVVEMLGTSSEPLHFYDVSIVDGFNLP	VSMKPVGG...GICGVA	SCVDLNIC			
4 AcTLP4	PPNTLAEYALKQYMD.QDFYDISVIDGFNVPM	EFSAASG....GCTRVIK	CTGDIIGQC			
5 T_danielli	PPITLAEYSLNQYG..KDYDISINIKGFNVPM	DFSPITR....GCRGV	RCAADIVGQ			
6 C_follicularis	PPASLAEFTIAPSGG.QDFYDISLVDGYNPL	LSIAPKGG..SGPT	CTTT.SCSADVNAVC			
7 M_esculenta	APNTLAEYALDQFER.QDFYDISVIDGFNVPM	EFSSASG....SCSRVIK	CIADILIGC			
8 A_thaliana	PPATLAEITLQGE...LDFYDVSIVDGFNPL	AMSIMPVKG..SG	QCSYACVSDLNQMC			
9 O_sativa	PPVTLAEFILGGTDG.KDFYDVSIVDGFNVG	IGVAATGARVNRST	CGYA.GCVGDVNALC			
10 P_monticola	NPVTLAQLSLHHGGNDVSSYTLSLVNGFNPL	TFTPHGG..RG	RCGIP.RCMANMLES			
11 P_patens	PPASLAEITLNGADG.NDFYDISLVDGYNPL	SMAPSGG..TG	KCGAP.GCISNLNDNC			
12 A_nidulans	DQTDVLCQFEMTKSGE.TLFDVMSCIDMDRA	ASTFTKNGF.....	DVSPSCTSGDC			
13 C_elegans	PPASLAEFILKAWGG.QDFYDVSIVDGFNLP	VLDPHGG..SG	CKRAGGCVKDLNABEC			



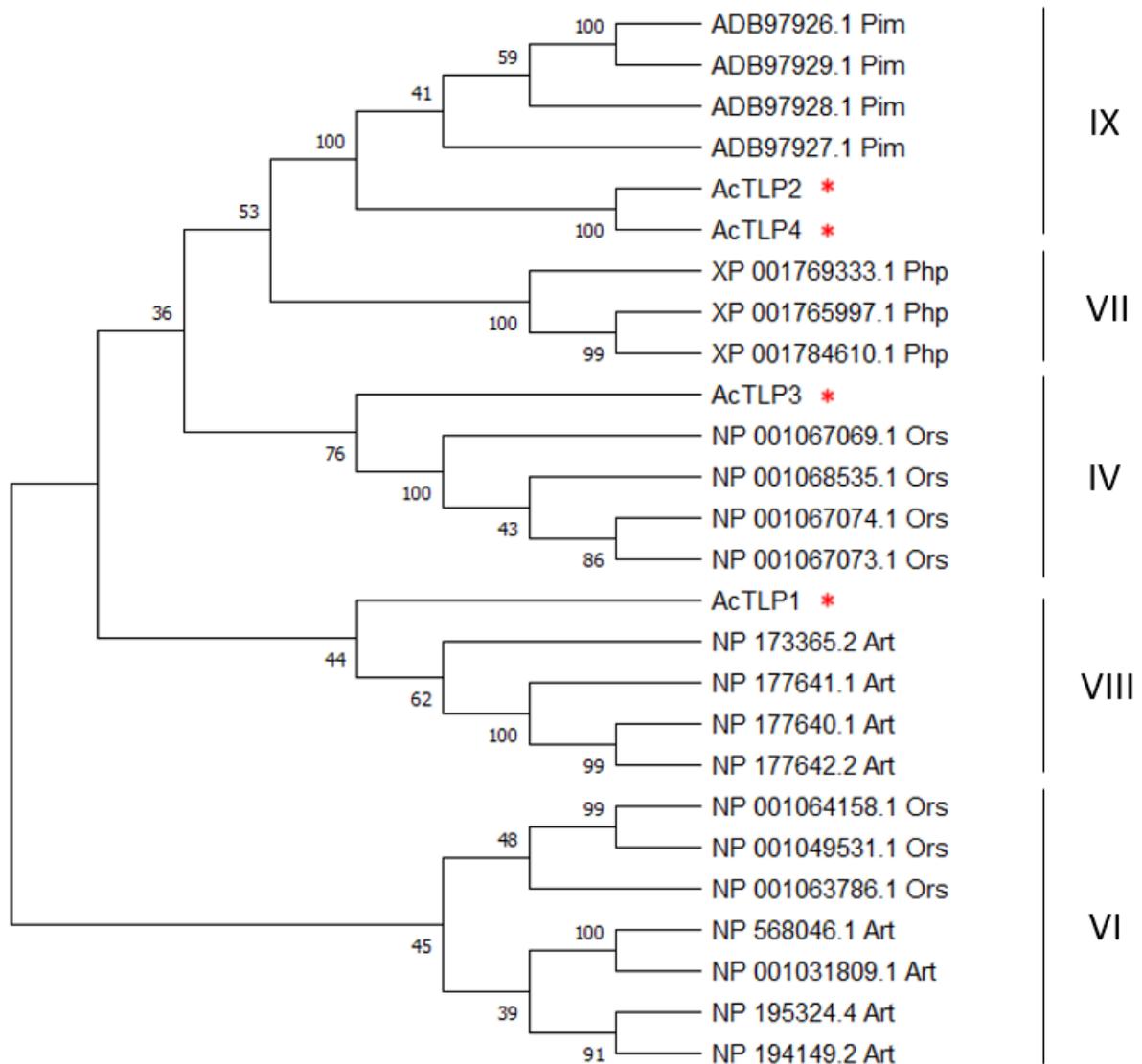
	170	180	190	200	210
1 AcTLP1	PABLQVKGSDGS.....VIA	CKSACAAALNQPOYCCT	GD.PSTPET	CPPTSYSQIFB	
2 AcTLP2	PSQLKAPG.....GCN	PFCTVFKTNEYCCT	NG.YGS..CGPT	NFASKYFK	
3 AcTLP3	PSALEVKKGNK.....VVG	CKSACLAMOSAKYCCT	GE.YANPNT	CKPTLFAHLFK	
4 AcTLP4	PNELKAVTG.....GCNG	ACVFKTDQYCCN	..GN...CGPT	DFSKYFK	
5 T_danielli	PAKLKAAPGG.....GND	ACTVPTSEYCCT	T..GK..CGPT	TSKRFK	
6 C_follicularis	PABLQVKGSDGR.....VIA	CKSACDAPSQPOYCCT	GA.YNTFAT	CPPTSYSNIFK	
7 M_esculenta	PNELKVPG.....GCNG	PCVFKTBEHCNS	..GN...CGPT	NFASKYFK	
8 A_thaliana	PVGLQVRSRNGKR.....VV	ACKSACAFNSPOYCCT	GL.FGNFQS	CKPTAYSKIFK	
9 O_sativa	PABLQVAGKENDQOSGAATITVA	CRSACAFPTAEYCCT	GA.HGGPDS	CGPTYSRIFK	
10 P_monticola	PKELQVKNQDN.....VIA	CKTACAFRTDAYCCT	BSQHYNGSRT	CPPTYSQIFK	
11 P_patens	PAALQFLABGV.....LVG	CNSACNAPNTEYCCT	GA.FGGPFT	CPPTQYSNAFK	
12 A_nidulans	PAVNCBAG.....DTSC	AEAYLQPKDDH.ATHG	CFDTSPTL	GLGAFSEFLF	
13 C_elegans	PAALSIVRGHNGK.....TVA	CKSGCLEYNTDQEC	CRGA.YGTPDK	CHRSAIACQMK	



	220	230	240
1 AcTLP1	NQCPQAYSAYDDLSTFT	CTGSPNYVITFCG	.....
2 AcTLP2	GACPDAYSYPQDDFTSTFT	TEPCGTNYRVVFCP	KGSAHFLEMVEGKNEE
3 AcTLP3	AICPKAYSVAFDSSSLNRC	RA.SRYVITFCPPK	.....
4 AcTLP4	DRCPDAYSYPKDDFTSLE	TCPSGTNYKVI	FCP.....
5 T_danielli	RLCPDAYSYPVLDKPTTV	TCPGSSNYRVVFCP	TALELEDE.....
6 C_follicularis	QQCPQAYSAYYDDTSSTFT	CTGSPNYVITFCP	.....
7 M_esculenta	DRCPDAYSYPKDDFTSLE	TCPTGTNYKVI	FCP.....
8 A_thaliana	VACPKAYSAYYDDFTSI	ATCSK.ANYIVFCP	HRRH.....
9 O_sativa	AACPAYSAYYDDFTSTFT	CGTGAQYVITFCP	AAQQ.....
10 P_monticola	RACPNAFAYPDDNPALVHN	CVTPNEIKLEFCH	.....
11 P_patens	SACPTAYSAYYDDATSTFT	CKG.ANYAITFCP	TGIFIPDLACTMKQ...
12 A_nidulans	AVLIQEYSYLF..PISVI	ICPQSSFLATY	LPLFKSA.....
13 C_elegans	DACPTAYSAYYDDGSSFT	CTCPASYSITVQFC	.....



**Figura 10. Alinhamento global das AcTLPs com TLPs de outras espécies.** As caixas em vermelho indicam as regiões dos prováveis peptídeos sinais identificadas pelo SignalP 6.0. Os colchetes indicam a assinatura familiar das taumatinas (G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C). As setas em vermelho mostram os dezesseis resíduos de cisteínas conservados. Os asteriscos representam as posições conservadas dos cinco aminoácidos do motivo REDDD. Os resíduos considerados altamente similares estão destacados em negrito e dentro de uma caixa vazia de contorno preto. Nos resíduos com identidade estrita, a coluna é colorida com fundo preto. As sequências utilizadas estão disponíveis no Anexo II.



**Figura 11. Árvore filogenética das AcTLPs com TLPs de outras plantas.** A inferência filogenética foi feita pelo método de *Neighbour-Joining* com o modelo de Poisson. Os valores de *bootstrap* estão indicados nos nós como valores de porcentagem. Os asteriscos vermelhos indicam as AcTLPs. Os códigos de acesso das sequências proteicas utilizados estão apresentados nos respectivos táxons. Art: *Arabidopsis thaliana*. Pim: *Pinus monticola*. Php: *Physcomitrella patens*. Ors: *Oryza sativa* subsp. *japonica*.

A partir dos resultados da figura 11, nota-se que as AcTLPs estão divididas entre três famílias: IV, VIII e IX. A AcTLP1 e AcTLP3 estão nas famílias VIII e IV, respectivamente, enquanto que as AcTLPs 2 e 4 estão num mesmo clado dentro da família IX. A partir destes resultados, escolheu-se duas TLPs de diferentes famílias, no caso a AcTLP1 e a AcTLP2, para se prosseguir com amplificação, clonagem, expressão, purificação e ensaios de atividade.

#### 4.2. Análises *in silico* das sequências das AcTLPs

O número de nucleotídeos e de aminoácidos, as massas moleculares e os pontos isoelétricos teóricos das AcTLPs estão representados na Tabela 3. De acordo com SignalP 6.0, todas as AcTLPs possuem um peptídeo sinal que direciona a proteína para a rota de secreção. Este resultado é condizente com o pI das mesmas, tendo em vista que TLPs ácidas tendem a ser extracelulares, enquanto que as básicas tendem a se acumular nos vacúolos ou outras vesículas (VELAZHAHAN; DATTA; MUTHUKRISHNAN, 1999).

**Tabela 3.** Características das AcTLPs.

Taumatina	Nucleotídeos (pb)	Aminoácidos	Massa Molecular Teórica	pI
AcTLP1	741	246	25.278 (Da)	4,29
AcTLP2	729	242	25.997 (Da)	5,89
AcTLP3	735	244	25.657 (Da)	7,81
AcTLP4	678	225	24.328 (Da)	4,92

Para verificar a semelhança das TLPs de Maqui frente a outras TLPs de plantas, foi feito um BLASTP no servidor do NCBI (Tabela 4). Os resultados apresentaram exemplares do clado das rosídeas. Este clado possui 14 ordens, dentre elas Oxalidales, a qual o Maqui pertence. Os exemplares dados pelo BLASTP pertencem às ordens Oxalidales (*C. follicularis*), Rosales (*C. sativa*, *Z. jujuba*), Malvales (*D. zibenthinus*, *H. umbratica*, *T. cacao*, *C. capsularis*, *C. oliotorius*), Fagales (*J. regia*), Malpighiales (*M. esculenta*, *H. brasiliensis*, *R. communis*) e Sapindales (*M. indica*). Somente há uma espécie da ordem Oxalidales no BLASTP, talvez pela escassez destas sequências no banco de dados. No NCBI, ao pesquisar por “Thaumatina” na ordem Oxalidales, há 21 resultados que correspondem ao mesmo organismo, *Cephalotus follicularis*, não havendo outro exemplar. Se a pesquisa for por

“Thaumatococcus”, mas na família Elaeocarpaceae, não há um item que corresponda aos filtros da busca. Desta forma, as AcTLPs são as primeiras TLPs a serem descritas dentre os membros da família Elaeocarpaceae.

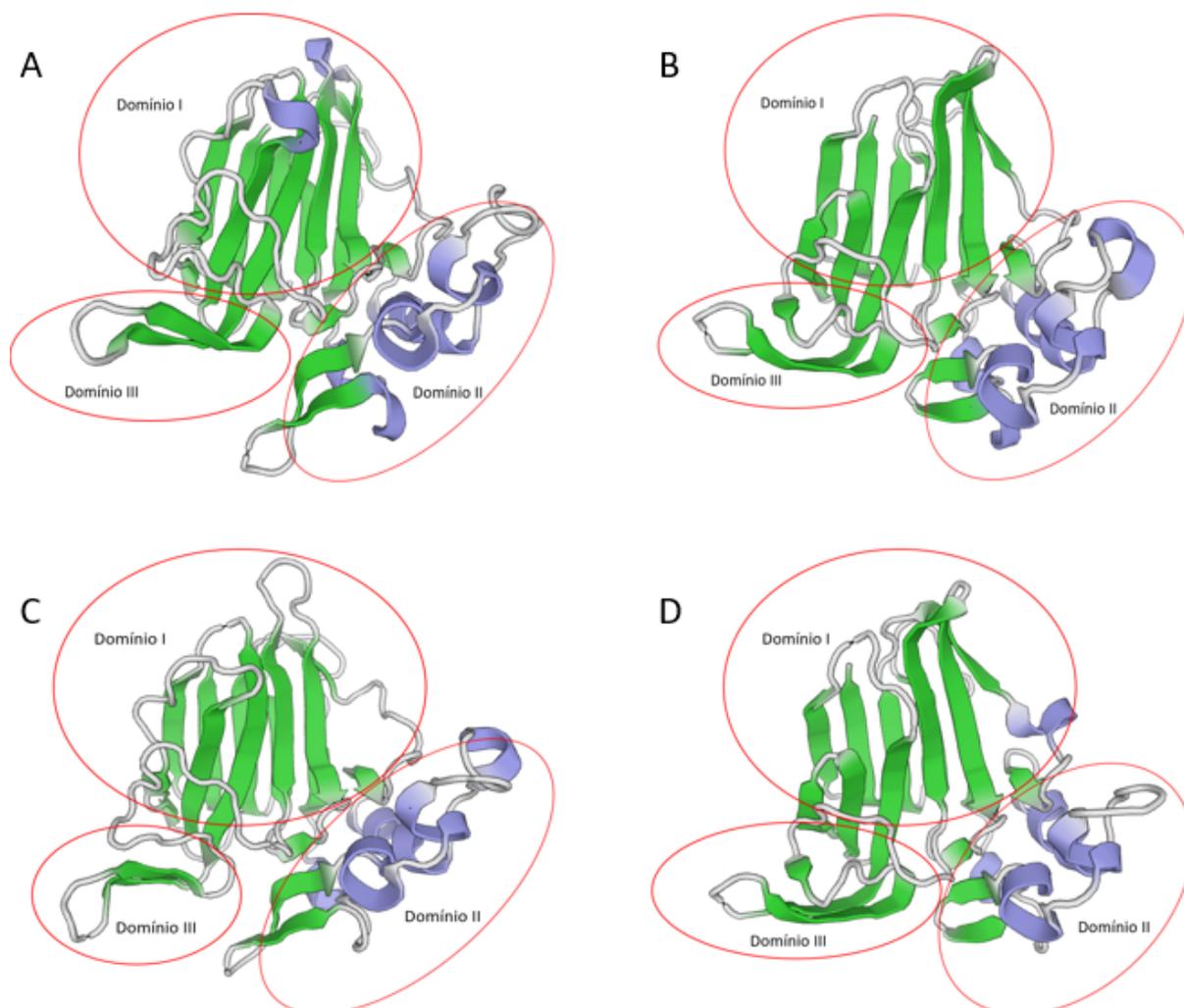
**Tabela 4. Identidade e similaridade da estrutura primária entre as AcTLPs e TLPs de outras espécies de plantas.** Os códigos de acesso do NCBI estão ao lado do nome da espécie.

Taumatina	Espécie de planta	Identidade (%)	Similaridade (%)
AcTLP1	<i>Cephalotus follicularis</i> (GAV85733.1)	83	90
	<i>Durio zibenthinus</i> (XP_022766133.1)	75	87
	<i>Herrania umbratica</i> (XP_021289565.1)	80	88
	<i>Theobroma cacao</i> (EOY07996.1)	79	89
AcTLP2	<i>Cephalotus follicularis</i> (GAV68928.1)	84	92
	<i>Cannabis sativa</i> (XP_030502231.1)	87	91
	<i>Juglans regia</i> (XP_018811920.1)	88	91
	<i>Ziziphus jujuba</i> (XP_015896184.1)	85	91
AcTLP3	<i>Cephalotus follicularis</i> (GAV76559.1)	91	96
	<i>Corchorus capsularis</i> (OMO66371.1)	91	96
	<i>Corchorus olitorius</i> (OMO97240.1)	91	96
	<i>Durio zibenthinus</i> (XP_022748176.1)	90	95
AcTLP4	<i>Manihot esculenta</i> (XP_021601423.2)	85	90
	<i>Mangifera indica</i> (XP_044484821.1)	82	90
	<i>Hevea brasiliensis</i> (XP_021657932.1)	83	89
	<i>Ricinus communis</i> (XP_002509750.2)	82	89

Até o momento, a estrutura terciária de várias TLPs já foi resolvida (LIU *et al.*, 2021). A estrutura terciária das quatro AcTLPs foi modelada de maneira homóloga com a plataforma SWISS-MODEL (Figura 12). As análises mostraram que quatro modelos podem ser usados para a modelagem das AcTLPs, um modelo para cada proteína. Para a AcTLP1, foi usada como modelo a TLP de *Prunus avium* (PDB: 2ahn.1). Para a AcTLP2, a TLP de *Musa acuminata* (PDB: 1z3q.1) foi utilizada. Para a AcTLP3, foi usada a TLP de *Malus domestica* (PDB: 3zs3.1). Por fim, para a AcTLP4, foi utilizada como modelo a TLP de *Vitis vinifera* (PDB: 4jru.1).

Como apresentado na Figura 12, a estrutura da AcTLP1 mais se assemelha com a da AcTLP3, enquanto que a AcTLP2 tem uma estrutura semelhante à AcTLP4, o que está de acordo com os resultados obtidos nos alinhamentos das proteínas (Figura 8) e na construção da árvore filogenética das mesmas (Figura 9). Cabe salientar que todas as AcTLPs possuem a

estrutura tridimensional esperada para as Taumatinas, composta por três domínios. O domínio I é o núcleo funcional N-terminal formado por duas folhas-beta antiparalelas. O domínio dois é formado por algumas pequenas alfa-hélices e o domínio três é formado por duas fitas-beta ligadas por um *loop* (LIU *et al.*, 2021).



**Figura 12.** Predição da estrutura terciária das AcTLPs de *Aristotelia chilensis*. As circunferências vermelhas indicam os três domínios de cada proteína. A: AcTLP1. B: AcTLP2. C: AcTLP3. D: AcTLP4.

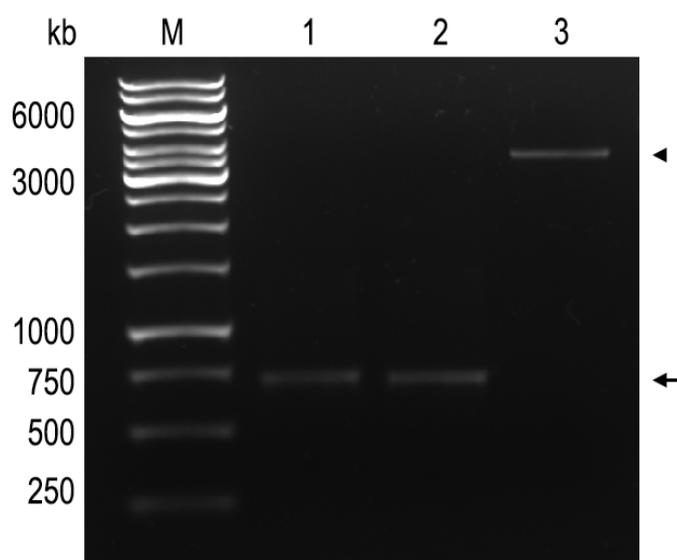
A Tabela 5 resume as características das AcTLP1 e AcTLP2 expressas em *Pichia pastoris*, com a remoção do peptídeo sinal, pois o plasmídeo de expressão já possui um peptídeo sinal de secreção próprio, e com a adição de *his-tag*, para purificação das proteínas.

**Tabela 5.** Características das AcTLPs recombinantes expressas em *Pichia pastoris*, possuindo aminoácidos oriundos do plasmídeo de expressão, incluindo a *his-tag*.

Taumatina	Aminoácidos	Massa Molecular Teórica	pI
AcTLP1	232	24.033 (Da)	4,78
AcTLP2	229	24.677 (Da)	6,05

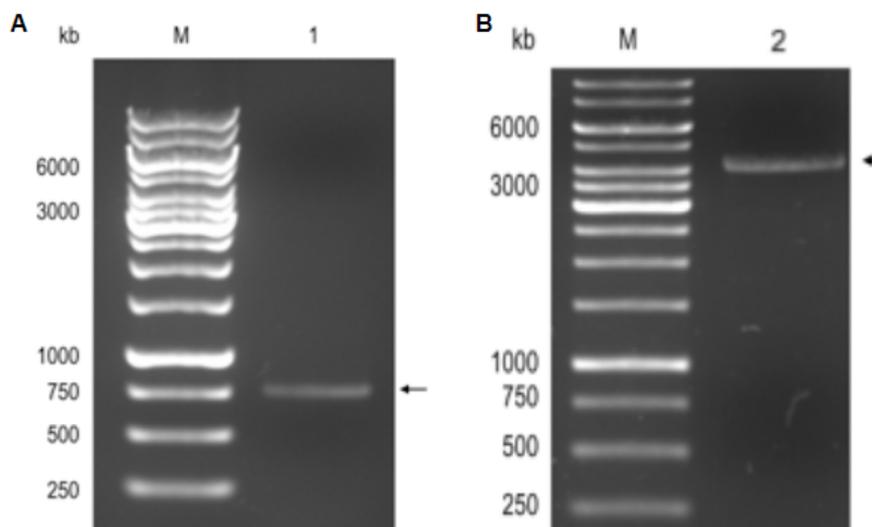
#### 4.3. Construção dos cassetes de expressão

Os oligonucleotídeos sintetizados foram desenhados para amplificar a sequência das AcTLPs retirando-se o peptídeo sinal. Como o vetor pPICZA não possui um códon iniciador (ATG) próprio, para a clonagem neste plasmídeo um códon de iniciação foi inserido no começo das sequências para posterior expressão intracelular. Além disso, o códon terminador das próprias TLPs foi retirado, visando fundir a proteína com uma cauda de histidina para purificação por cromatografia de afinidade. As ORFs das TLPs foram amplificadas por PCR e resolvida em gel de agarose 1%, apresentando o tamanho esperado de aproximadamente 660 pb para as duas sequências (Figura 13). A ORF da GFP também foi amplificada por PCR e resolvida em gel de agarose 1%, apresentando o tamanho esperado de aproximadamente 720 pb (Figura 14A). Os vetores pPICZA (Figura 13) e pPICZαA (Figura 14B) foram clivados com as enzimas de restrição *EcoRI* e *SalI*. Os *amplicons* anteriores também foram digeridos com estas enzimas e logo depois ligados separadamente aos vetores.



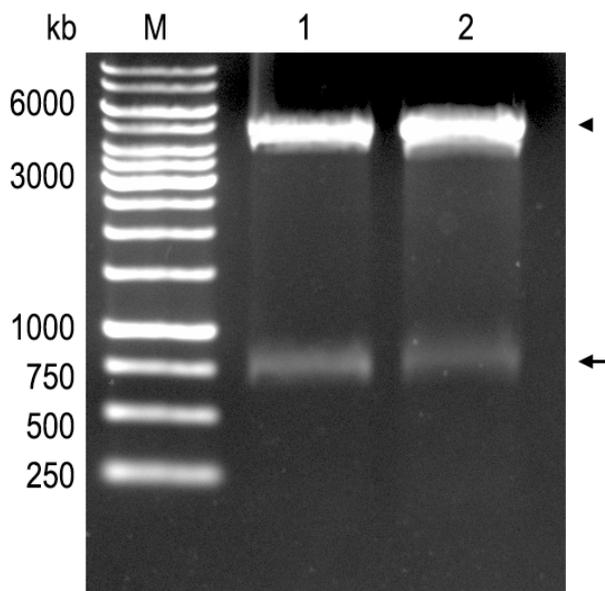
**Figura 13. Amplificação das ORFs das AcTLPs e clivagem do pPICZA.** Gel de agarose 1% contendo em M: marcador de peso molecular (GeneRuler 1kb ThermoScientific). 1: produto de PCR da AcTLP1. 2: produto de PCR da AcTLP2. 3: pPICZA digerido com *EcoRI* e *SalI*. A seta indica as bandas referentes à amplificação das

TLPs (aproximadamente 660 pb) e a cabeça de seta indica a banda referente ao pPICZA clivado (aproximadamente 3.200 pb).

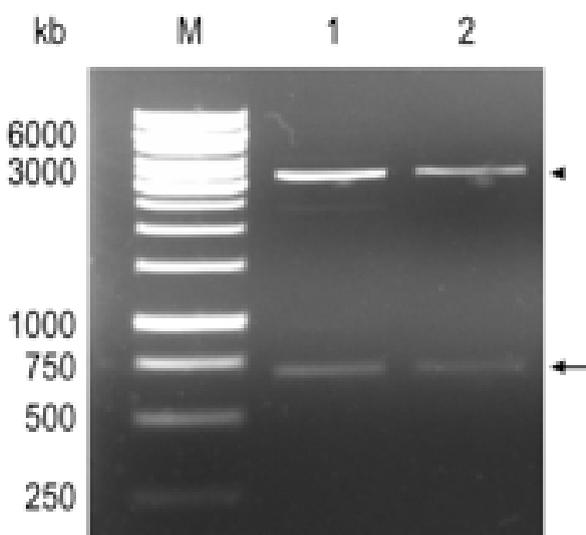


**Figura 14. Amplificação da ORF da GFP e clivagem do pPICZ $\alpha$ A para clonagem.** Gel de agarose 1% contendo em M: marcador de peso molecular (GeneRuler 1kb ThermoScientific). 1: produto de PCR da GFP. 2: pPICZ $\alpha$ A digerido com *EcoRI* e *SalI*. A: Amplificação da ORF da GFP. B: Clivagem do pPICZ $\alpha$ A para clonagem. A seta indica a banda referente à amplificação da GFP (aproximadamente 720 pb). A cabeça de seta indica a banda referente ao pPICZ $\alpha$ A clivado (aproximadamente 3.500 pb).

Os genes das TLPs foram clonados em ambos os vetores pPICZA e pPICZ $\alpha$ A. A ligação foi transformada em *E. coli* DH5a e a clonagem foi confirmada por PCR de colônia e pela clivagem dos plasmídeos com as mesmas enzimas que foram usadas na clonagem (Figuras 15 e 16). A escolha destes plasmídeos se deu pela ausência ou presença do peptídeo sinal secreção ( $\alpha$ -factor), o que resulta numa expressão intracelular ou secretada, respectivamente, das AcTLPs.



**Figura 15. Confirmação da clonagem por clivagem dos plasmídeos pPICZ $\alpha$ A-AcTLP1 e pPICZ $\alpha$ A-AcTLP2.** Gel de agarose 1% contendo em M: marcador de peso molecular (GeneRuler 1kb ThermoScientific). 1: pPICZ $\alpha$ A-AcTLP1. 2: pPICZ $\alpha$ A-AcTLP2. A seta indica as bandas referentes ao inserto clivado (aproximadamente 660 pb). A cabeça de seta indica o vetor pPICZ $\alpha$ A clivado (aproximadamente 3.500 pb).



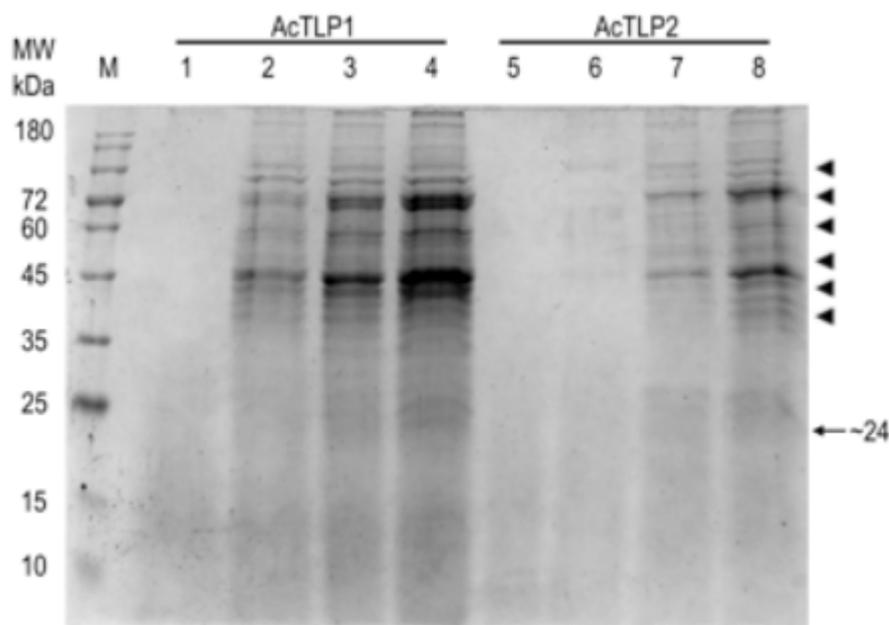
**Figura 16. Confirmação da clonagem por clivagem dos plasmídeos pPICZA-AcTLP1 e pPICZA-AcTLP2.** Gel de agarose 1% contendo em M: marcador de peso molecular (GeneRuler 1kb ThermoScientific). 1: pPICZA-AcTLP1. 2: pPICZA-AcTLP2. A seta indica as bandas referentes ao inserto clivado (aproximadamente 660 pb). A cabeça de seta indica o vetor pPICZA clivado (aproximadamente 3.200 pb).

#### 4.4. Expressão e purificação das TLPs recombinantes

O plasmídeo utilizado para a expressão das AcTLPs foi o pPICZ $\alpha$ A, o qual expressa a proteína em fusão com o fator alfa de secreção N-terminal e uma *his-tag* C-terminal. O fator

alfa é um peptídeo sinal que endereça a proteína recombinante para a rota de secreção (EasySelect Pichia Expression Kit, 2001), facilitando sua purificação, pois não há a necessidade de fazer a lise das células da levedura. A *his-tag* permite a purificação da proteína em um único passo, por cromatografia de afinidade com metal imobilizado.

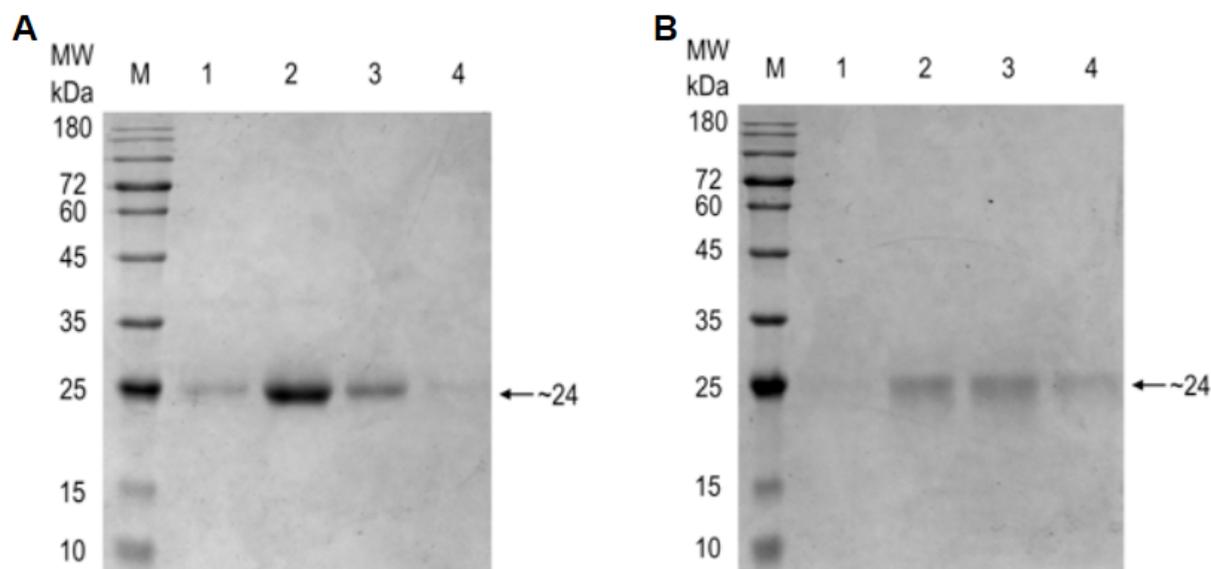
A análise por SDS-PAGE das alíquotas coletadas permitiu a avaliação do perfil de expressão das proteínas a partir do início até 72 horas de indução (Figura 17). Nesta figura, é possível perceber a aparição de bandas bem definidas e de diversos tamanhos, incomuns à expressão de proteínas secretadas por *P. pastoris*. Este padrão poderia ser causado pela ação de proteases que clivam proteínas da própria *P. pastoris*, o que já foi observado anteriormente com a expressão de proteases em *P. pastoris* no Laboratório de Biologia Molecular (SHIBAO, 2021). Entretanto, este padrão de bandas poderia também ser de proteínas intracelulares da *Pichia*, que se encontram no sobrenadante da indução porque a membrana plasmática da *P. pastoris* foi permeabilizada pela ação das TLPs. Se for este o caso, este é o primeiro indício de atividade das AcTLPs.



**Figura 17. Análise da indução das AcTLPs em *Pichia pastoris*.** SDS-PAGE 12% contendo em M: BlueClassic Prestained Protein Marker (Jena Bioscience). 1 e 5: Alíquotas não-induzida (0 h). 2 e 6: Alíquotas de 24 h de indução. 3 e 7: Alíquotas de 48 h de indução. 4 e 8: Alíquotas de 72 h de indução. As cabeças de seta indicam as várias bandas de proteínas incomuns à expressão secretada em *Pichia pastoris*. A seta aponta as possíveis bandas referentes às AcTLPs.

As TLPs de Maqui foram expressas com sucesso, tendo em vista que suas bandas aparecem claramente no gel após purificação em coluna imobilizada com níquel (Figura 18).

Após purificação e quantificação, a AcTLP1 teve um rendimento de aproximadamente 50 mg/L de cultura, enquanto que a AcTLP2 teve cerca de 8 mg/L.



**Figura 18. Análise das purificações das AcTLPs.** SDS-PAGE 12% mostrando em M: BlueClassic Prestained Protein Marker (Jena Bioscience). 1 e 2: Gradientes de 75 mM de Imidazol. 3 e 4: Gradientes de 100 mM de Imidazol. As setas indicam as bandas referentes às AcTLP1 ou AcTLP2. A: AcTLP1. B: AcTLP2.

#### 4.5. Caracterização das AcTLPs

##### 4.5.1. Atividade de $\beta$ -1,3-Glucanase

As AcTLPs foram incubadas com Laminarina, a qual é um oligômero de  $\beta$ -1,3-glucanos solúvel em água comumente usado para ensaios de  $\beta$ -1,3-Glucanases (TRUDEL *et al.*, 1998). Os resultados estão apresentados nos gráficos a seguir (Figura 19). A diferença estatística entre os resultados da amostra-teste e dos controles indicam que há atividade de Laminarinase em ambas as proteínas, sendo que na AcTLP1 esta atividade se mostra mais acentuada.

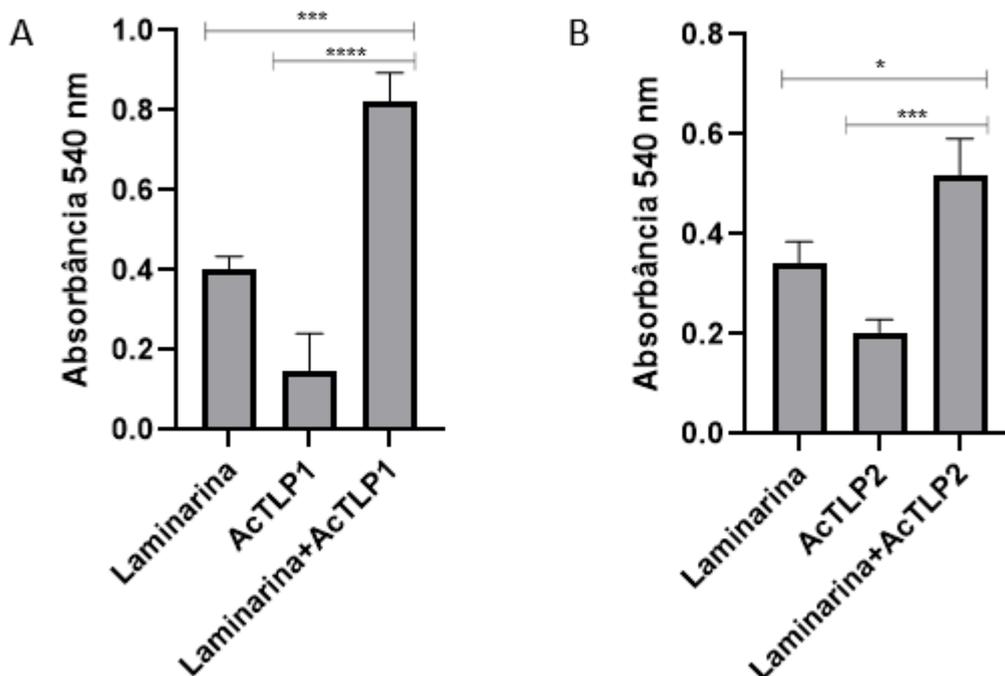


Figura 19. Gráficos de colunas representando a absorbância a 540 nm dos ensaios de atividade de  $\beta$ -1,3-glucanase e seus respectivos controles. A: Ensaio com AcTLP1. B: Ensaio com AcTLP2. Um asterisco (\*) indica que  $p \leq 0,05$ , três (\*\*\*) indicam que  $p \leq 0,001$ , quatro (\*\*\*\*) indicam que  $p \leq 0,0001$ .

Apesar de ser notório a atividade de glucanase das AcTLP1 e AcTLP2, elas possuem atividades enzimáticas baixas, menores que 1 U/mg (0,29 e 0,13 para as AcTLP1 e AcTLP2, respectivamente). As TLPs de plantas possuem atividade de glucanase variadas, indo de aproximadamente 0,004 a até 1.477 U/mg (GRENIER *et al.*, 1999). Este resultado coloca as AcTLPs dentro do espectro esperado para as TLPs, embora fiquem nos níveis inferiores de atividade enzimática quando comparadas com outras TLPs vegetais.

Já se sabe sobre a capacidade das TLPs de se ligarem em  $\beta$ -1,3-Glucanos, principalmente as oriundas de plantas monocotiledôneas, embora existam TLPs de dicotiledôneas que também possuem esta capacidade. Ainda assim, esta característica não é compartilhada por todos os indivíduos desta superfamília. Também é importante a estrutura do polissacarídeo para a ligação. Por exemplo, há TLPs que se ligam no Pachyman (polímero linear de  $\beta$ -1,3-glucanos insolúvel em água), entretanto não são capazes de se ligar ao Pachyman após ser digerido com Zymolase, sugerindo que os  $\beta$ -1,3-glucanos removidos com este tratamento eram os responsáveis pela ligação ao substrato. A ligação das TLPs com os

$\beta$ -1,3-glucanos, componentes comumente encontrados na parede celular de fungos, pode estar associada com a atividade antifúngica destas proteínas (TRUDEL *et al.*, 1998).

A interação das TLPs com os  $\beta$ -glucanos não se resume somente às suas ligações. Como demonstrado, também pode haver a clivagem dos polissacarídeos por parte das proteínas. Há casos de TLPs que não conseguem hidrolisar a Laminarina, porém são capazes de hidrolisar polissacarídeos mais complexos como o Pachyman ou o Carboximetil (CM) Pachyman, que são insolúveis e solúveis em água, respectivamente (GRENIER *et al.*, 1999). Existem também TLPs que são capazes de hidrolisar a Laminarina (LIU *et al.*, 2021). Neste trabalho, foram feitos somente ensaios de Laminarinase com as AcTLPs.

Apesar dos mecanismos pelos quais as TLPs realizam a quebra da ligação glicosídica ainda não serem totalmente elucidados, supõe-se que a fenda inter-domínio da proteína contribua para esta atividade. A TLP NP24-I, uma isoforma da NP24 de tomate, possui atividade de  $\beta$ -1,3-glucanase frente a polímeros lineares destes glucanos (GRENIER *et al.*, 1999). Gosh e Chakrabarti (2008) reportaram a estrutura cristalizada da NP24-I e fizeram estudos de *docking* para entender a base estrutural de sua interação com  $\beta$ -1,3-glucanos e tentar esclarecer as propriedades de sua ligação e atividade de glucanase. Dentre os vários aminoácidos envolvidos com a interação da proteína com o polissacarídeo, dois resíduos se destacaram. No complexo de ancoragem, o  $\beta$ -1,3-glucano passa pela fenda da proteína com Glu84 de um lado e Asp102 de outro. Esta orientação de pareada de Glu-Asp, que é encontrada em glucanases, é necessária para a quebra da ligação glicosídica (GOSH; CHAKRABARTI, 2008).

Ambas AcTLP1 e AcTLP2 possuem estes resíduos de ácido glutâmico e ácido aspártico. Interessantemente, eles fazem parte do motivo REDDD conservado, o que explica estarem posicionados na fenda acídica entre os domínios I e II das TLPs (Figura 20). Esta propriedade das AcTLPs talvez contribua para as suas atividades de glucanase demonstradas.

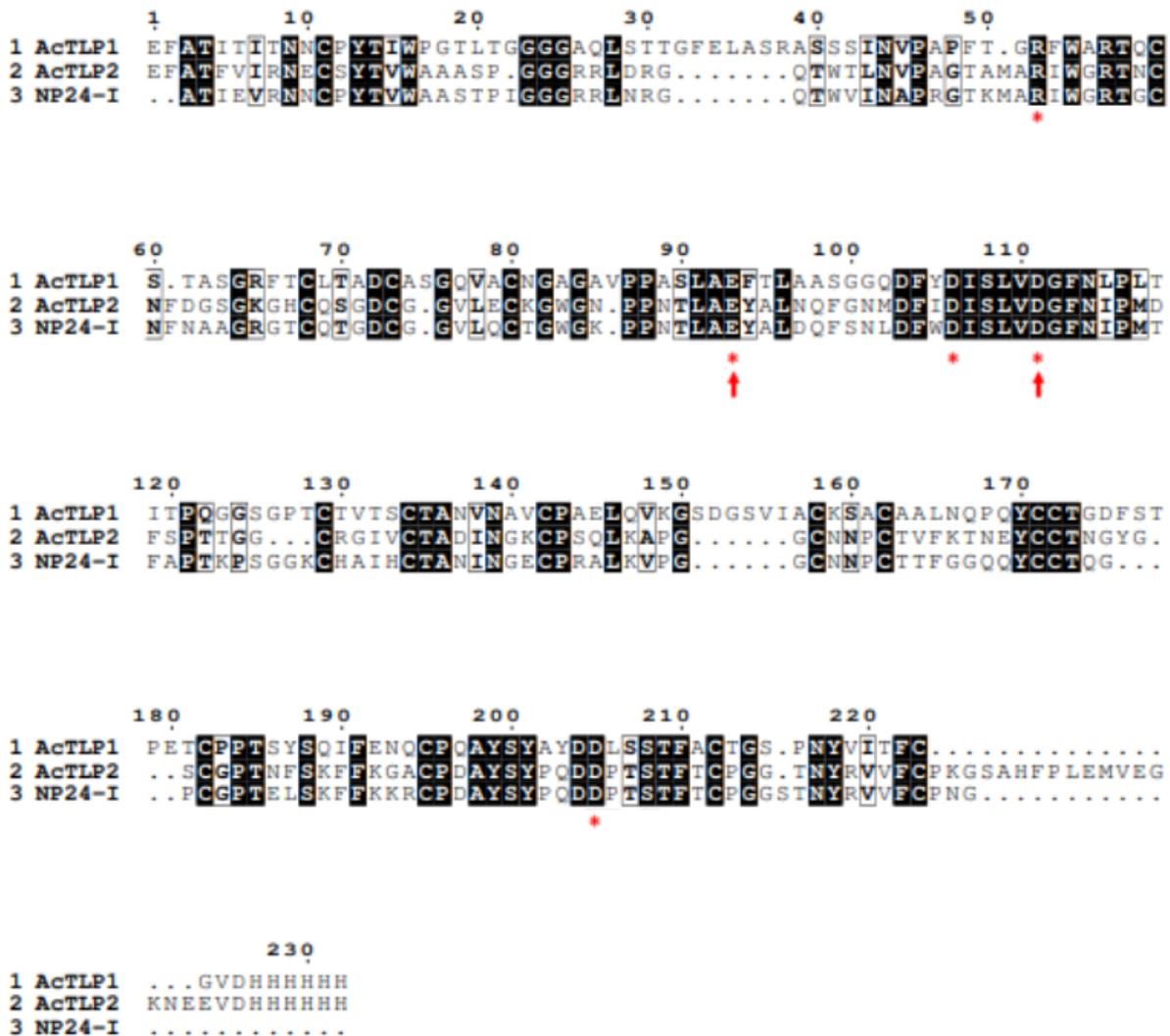


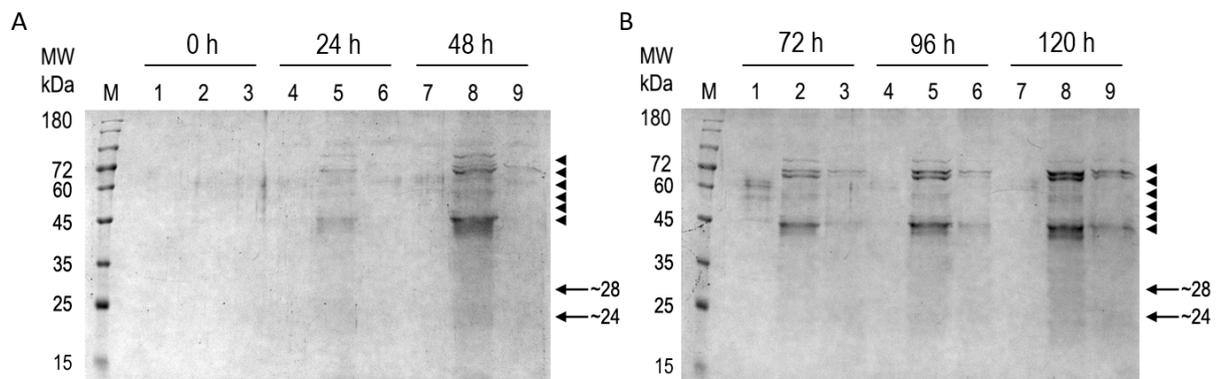
Figura 20. Alinhamento entre a NP24-I de tomate e as AcTLPs. Os asteriscos em vermelho representam os resíduos do motivo REDDD conservado. As setas em vermelho indicam os resíduos conservados de Ácido Glutâmico e Ácido Aspártico, os quais possivelmente atuam na hidrólise de  $\beta$ -glucanos. Os resíduos considerados altamente similares estão destacados em negrito e dentro de uma caixa vazia de contorno preto. Nos resíduos com identidade estrita, a coluna é colorida com fundo preto. NP24-I PDB: 2I0W.

#### 4.5.2. Permeabilização de membrana plasmática de levedura

Para este ensaio de atividade, os clones de *Pichia pastoris* transformadas com pPICZaA-AcTLP1 e pPICZaA-AcTLP2, utilizadas para a expressão das taumatinas, e também um clone de *Pichia pastoris* transformada com pPICZA-GFP, previamente preparada, a qual expressa a GFP intracelularmente.

Foi feito um co-cultivo da colônia produtora de GFP com as colônias produtoras de AcTLP1 ou AcTLP2. Também foi feito co-cultivo da produtora de GFP com uma *P. pastoris* X-33 não transformada como controle. A indução das proteínas recombinantes foi feita por

adição de metanol para concentração final 0,75% todos os dias por 144 h. Aliquotas foram recolhidas a cada 24 h e os sobrenadantes foram analisados por SDS-PAGE 12% (Figura 21).



**Figura 21. Análise da permeabilização de membrana no co-cultivo com GFP.** SDS-PAGE 12% mostrando em 1, 4 e 7: Amostras controle X-33 + GFP. 2, 5 e 8: Amostras AcTLP1 + GFP. 3, 6 e 9: Amostras AcTLP2 + GFP. As cabeças de seta indicam as possíveis bandas de proteínas oriundas das permeabilização da membrana de *Pichia pastoris*. As setas se referem às prováveis bandas de GFP (~28) e AcTLPs (~24). A: Amostras de 0 a 48 h. B: Amostras de 72 a 120 h.

Apesar de não ser bem evidente a banda referente à GFP ou às AcTLPs nos géis apresentados na Figura 21, é clara a diferença entre o padrão de bandas das amostras controle e o das amostras teste. Estas bandas proteicas que aparecem nas canaletas das amostras testes, quando em comparação com o controle, devem ser proteínas da *P. pastoris* que provavelmente sofreram permeabilização de sua membrana pelas TLPs e seu conteúdo celular vazou para o meio extracelular.

Testes fluorimétricos com as mesmas alíquotas corroboram com os resultados do SDS-PAGE 12%. Os sobrenadantes foram diluídos em água (1:30) e a fluorescência foi analisada em fluorímetro, durante 20 segundos, com excitação a 475 nm e emissão a 503 nm. As fluorescências estão apresentadas no gráfico da Figura 22. Com o passar do tempo do co-cultivo, a emissão a 503 nm das amostras teste aumentou significativamente em relação ao controle (X-33+GFP), principalmente no co-cultivo com a AcTLP1.

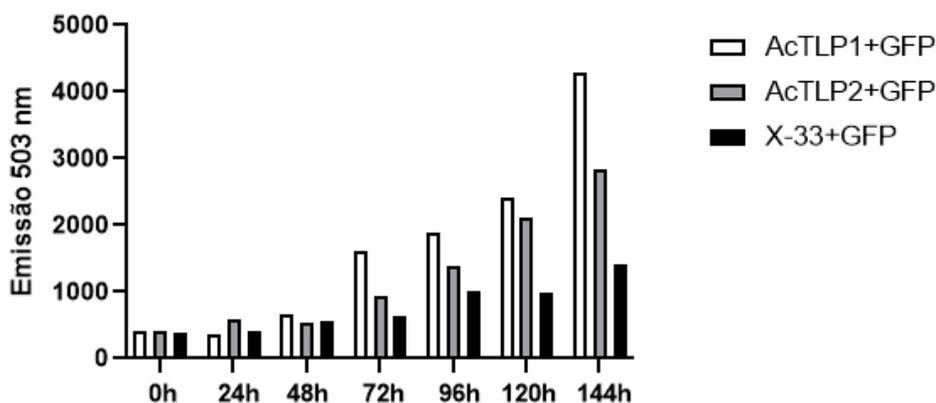
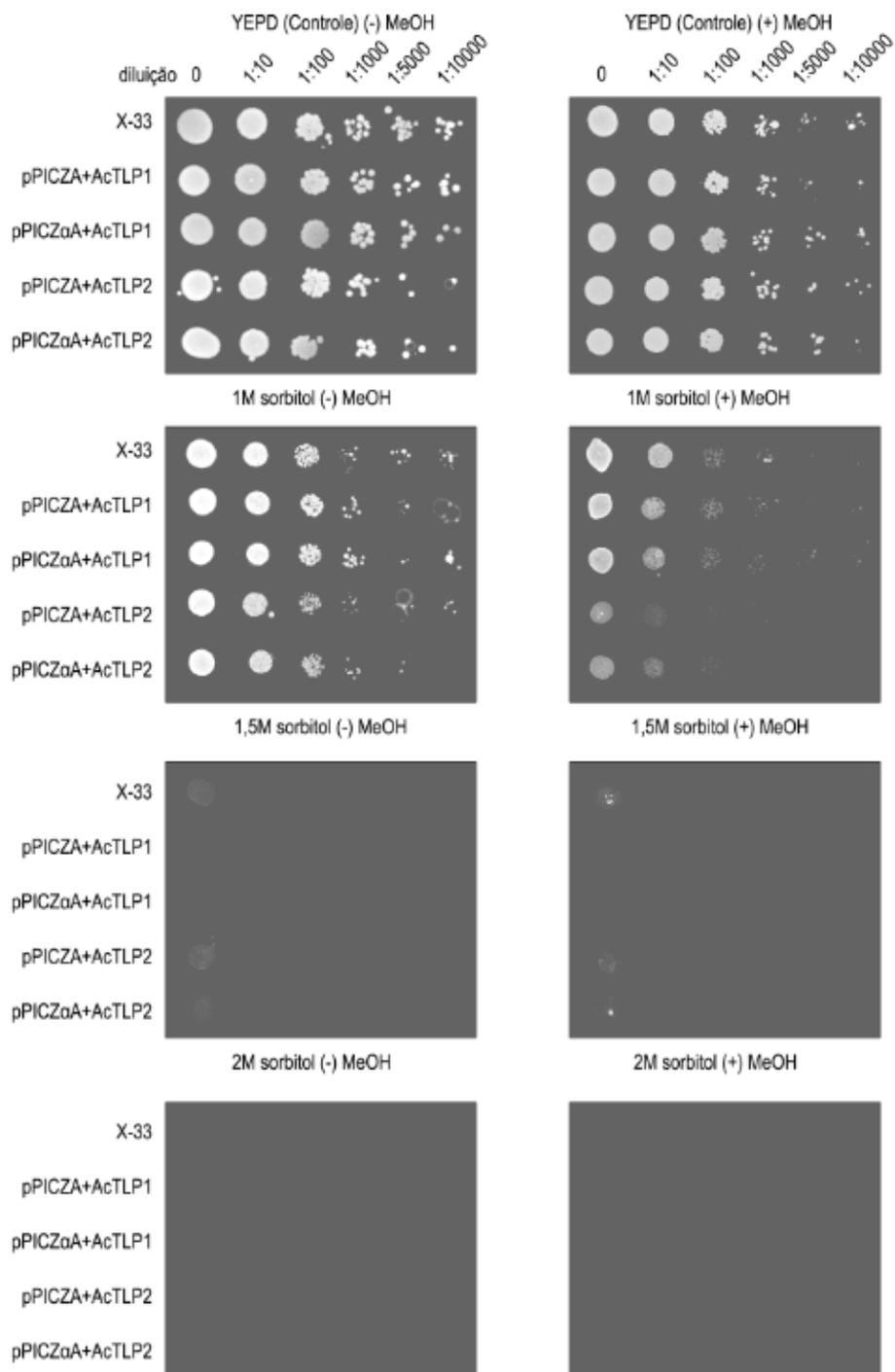


Figura 22. Gráfico representando a emissão a 503 nm das alíquotas do co-cultivo ao longo de 144 horas.

Uma característica recorrente das TLPs é possuir atividade antifúngica. Apesar do mecanismo desta ação ainda não ser bem elucidado, frequentemente se atribui às TLPs a propriedade de permeabilizar a membrana plasmática, exibindo a atividade antifúngica ao matar o fungo por vazamento do conteúdo intracelular. Por exemplo, uma TLP de banana, a BanTLP, exerce sua atividade antifúngica por induzir perturbações na membrana plasmática e desorganização da parede celular, levando o fungo *Penicillium expansum* à morte celular (JIAO *et al.*, 2018). Foi demonstrada a capacidade das AcTLPs de permeabilizarem a membrana plasmática de *P. pastoris*, mas os mecanismos de ação por detrás desta atividade, e também se esta atividade se mantém frente a outras espécies de fungo, ainda carecem de mais estudos. De fato, foram feitos ensaios antifúngicos *in vitro* em colaboração com o Prof. Dr. Luís Eduardo Aranha Camargo, do Departamento de Fitologia e Patologia da ESALQ/USP. Foram testados os fungos *Fusarium moniliforme*, *Phytophthora* spp, *Sclerotinia sclerotiorum* e *Colletotrichum falcatum*. No entanto, nenhuma delas foi capaz de inibir o crescimento destes fungos (dados não mostrados).

#### 4.5.3. Tolerância ao estresse osmótico

Para verificar se a AcTLP1 ou AcTLP2 são capazes de aumentar a tolerância do organismo ao estresse osmótico, células de *Pichia pastoris* expressando as TLPs foram comparadas com células controle sob condições mimetizadas de estresse osmótico. O grau da condição foi mimetizado por aumentar a concentração de sorbitol no meio de cultura da placa. Sob cada condição osmótica, as leveduras foram inoculadas nas placas em seis diluições seriadas (Figura 23).



**Figura 23. Ensaio de tolerância ao estresse osmótico de *Pichia pastoris* em placa.** A figura mostra células não transformadas (X-33), células transformadas expressando AcTLPs internamente (pPICZA) ou células transformadas expressando e secretando AcTLPs (pPICZαA). O painel esquerdo não foi induzido com metanol [(-) MeOH], enquanto que o direito foi induzido com metanol [(+) MeOH].

Nem as células transformadas, nem as não transformadas, cresceram bem na presença de sorbitol quando comparadas com a condição controle (uma placa de YEPD sem adição de sorbitol). É notável a diminuição no crescimento celular conforme a concentração de sorbitol

aumenta. Na presença de 1,5 M de Sorbitol, somente o crescimento de células das frações que não foram diluídas é evidente. Não houve crescimento nem mesmo das frações não diluídas na condição de 2 M de Sorbitol.

Salienta-se a indiferença no crescimento celular das placas tratadas com [(+) MeOH] ou sem [(-) MeOH] a adição de Metanol 100%. Este resultado pode ser atribuído à presença de glicose no meio YEPD, a qual atua como um repressor do AOX1 (HARTNER; GLIEDER, 2006), promotor que está controlando a expressão das AcTLPs nestas *P. pastoris*.

## 5. CONSIDERAÇÕES FINAIS

Quatro sequências de *Thaumatin-like Proteins* foram identificadas no transcriptoma de *Aristotelia chilensis* realizado por sequenciamento de nova geração a partir do RNA de plantas cultivadas em biorreatores de imersão temporária. Após análises *in silico* duas sequências foram amplificadas a partir de cDNA e clonadas em vetores de expressão.

A expressão das TLPs de Maqui na levedura *Pichia pastoris* X-33 usando o vetor pPICZαA se mostrou eficiente, possibilitando a purificação de proteínas solúveis, com rendimento satisfatório para a realização de todos os ensaios propostos.

As duas TLPs de Maqui conseguem de hidrolisar β-1,3-Glucanos, apesar de esta atividade ser baixa (menor que 1 unidade por miligrama de proteína). Além disso, possivelmente elas também podem permeabilizar a membrana plasmática de *Pichia pastoris*. As AcTLPs não foram capazes de aumentar a tolerância ao estresse osmótico de *Pichia pastoris* expressando de forma recombinante as proteínas.

## 6. PERSPECTIVAS FUTURAS

Considerando-se a vasta gama de possíveis atividades das TLPs, ainda se faz necessário continuar realizando ensaios de atividades com as AcTLP1 e AcTLP2, pois é possível que elas tenham alguma atividade não testada neste trabalho, como por exemplo *antifreezing* ou antiinflamatória. A realização de estudos estruturais destas proteínas também está prevista.

Também pretende-se clonar e expressar as AcTLP3 e AcTLP4 em *Pichia pastoris* e realizar tanto ensaios de atividade quanto estudos estruturais.

Por fim, planeja-se realizar mais ensaios de atividade antifúngica *in vitro* com as AcTLPs, focando em fungos fitopatogênicos relacionados ao Maqui, como por exemplo *Macrophomina phaseolina*, *Crepitodus aristoteliae*, *Mycenella aristoteliae*, *Septoria maqui* (SALINAS *et al.*, 2012), dentre outros.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

ANVISA. AGÊNCIA NACIONAL DE VIGILÂNCIA SANITÁRIA. Resolução da Diretoria Colegiada. Resolução RDC nº 18, de 24 de março de 2008. Disponível em: <<https://www.gov.br/agricultura/pt-br/assuntos/inspecao/produtos-vegetal/legislacao-1/biblioteca-de-normas-vinhos-e-bebidas/resolucao-rdc-no-18-de-24-de-marco-de-2008.pdf/view>>. Acessado em: 05 mar. 2022.

ARAUJO SOUSA, Bruna et al. Identification of the Active Principle Conferring Anti-Inflammatory and Antinociceptive Properties in Bamboo Plant. **Molecules**, v. 26, n. 10, p. 3054, 2021.

TEUFEL, Felix et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. **Nature biotechnology**, p. 1-3, 2022.

CREGG, J. M. **Methods in Molecular Biology: Pichia Protocols**. 2<sup>a</sup>. ed. Totowa, New Jersey: Humana Press, v. 389, 2007.

DE JESÚS-PIRES, Caroline et al. Plant thaumatin-like proteins: function, evolution and biotechnological applications. **Current Protein and Peptide Science**, v. 21, n. 1, p. 36-51, 2020.

EasySelect™ Pichia Expression Kit. A Manual of Methods for Expression of Recombinant Proteins Using pPICZ and pPICZa in Pichia pastoris. Version G, Invitrogen, Carlsbad, 2001.

EDGAR, Robert C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. **BMC bioinformatics**, v. 5, n. 1, p. 1-19, 2004.

GHOSH, Raka; CHAKRABARTI, Chandana. Crystal structure analysis of NP24-I: a thaumatin-like protein. **Planta**, v. 228, n. 5, p. 883-890, 2008.

GRABHERR, Manfred G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature biotechnology**, v. 29, n. 7, p. 644-652, 2011.

GRENIER, Jean et al. Some thaumatin-like proteins hydrolyse polymeric  $\beta$ -1, 3-glucans. **The Plant Journal**, v. 19, n. 4, p. 473-480, 1999.

HARTNER, Franz S.; GLIEDER, Anton. Regulation of methanol utilisation pathway genes in yeasts. **Microbial cell factories**, v. 5, n. 1, p. 1-21, 2006.

JANGPROMMA, Nisachon et al. Molecular cloning and characterization in eukaryotic expression systems of a sugarcane cysteine protease inhibitor gene involved in drought tolerance. **Turkish Journal of Botany**, v. 38, n. 4, p. 724-736, 2014.

JIAO, Wenxiao et al. Antifungal activity of an abundant thaumatin-like protein from banana against *Penicillium expansum*, and its possible mechanisms of action. **Molecules**, v. 23, n. 6, p. 1442, 2018.

KUMAR, Sudhir et al. MEGA X: molecular evolutionary genetics analysis across computing platforms. **Molecular biology and evolution**, v. 35, n. 6, p. 1547, 2018.

LIU, Jun-Jun; STURROCK, Rona; EKRAMODDOULLAH, Abul KM. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. **Plant cell reports**, v. 29, n. 5, p. 419-436, 2010.

LIU, Yufeng et al. Molecular Identification and Antifungal Properties of Four Thaumatin-like Proteins in Spruce (*Picea likiangensis*). **Forests**, v. 12, n. 9, p. 1268, 2021.

LUCENA, Severino A. et al. Miniaturization of hydrolase assays in thermocyclers. **Analytical biochemistry**, v. 434, n. 1, p. 39-43, 2013.

MASOODI, H.; VILLAÑO, D.; ZAFRILLA, P. A comprehensive review on fruit *Aristotelia chilensis* (Maqui) for modern health: towards a better understanding. **Food & function**, v. 10, n. 6, p. 3057-3067, 2019.

MISLE, E. et al. Maqui (*Aristotelia chilensis* (Mol.) Stuntz) the amazing chilean tree: a review. **Journal of Agricultural Science and Technology B**, v. 1, n. 4b, p. 473-482, 2011.

MIRANDA-ROTTMANN, Soledad et al. Juice and phenolic fractions of the berry *Aristotelia chilensis* inhibit LDL oxidation in vitro and protect human endothelial cells against oxidative stress. **Journal of Agricultural and Food Chemistry**, v. 50, n. 26, p. 7542-7547, 2002.

MURASHIGE, Toshio; SKOOG, Folke. A revised medium for rapid growth and bio assays with tobacco tissue cultures. **Physiologia plantarum**, v. 15, n. 3, p. 473-497, 1962.

NAWROT, Robert 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.

NIU, X. et al. The antifungal activity of a thaumatin-like protein from oyster *Crassostrea gigas*. **Invertebrate Survival Journal**, v. 15, n. 1, p. 210-222, 2018.

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

SALINAS, Jaime et al. Monografía de MAQUI *Aristotelia chilensis* (Mol.) Stuntz. **Programa de Investigación de productos forestales no madereros. Institut o Forestal**, 2012.

SAMBROOK, J.; RUSSEL, D. **Molecular Cloning: A Laboratory Manual**. 3a ed. Nova Iorque: Cold Spring Harbor Laboratory Press, 2000.

SANTIAGO, A. C. et al. A new sugarcane cystatin strongly binds to dental enamel and reduces erosion. **Journal of Dental Research**, v. 96, n. 9, p. 1051-1057, 2017.

SHARMA, Alok et al. Molecular characterization revealed the role of thaumatin-like proteins of bread wheat in stress response. **Frontiers in plant science**, v. 12, 2021.

SHIBAO, P. Y. T. **Cistatinas de cana-de-açúcar: produção recombinante em vacúolos de cana-de-açúcar transgênica (e outros sistemas de expressão em plantas) e estudos de interação com uma pseudo protease do coleóptero *Sphenophorus levis***. 2021. Tese de Doutorado. Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos.

SIMÃO, Felipe A. et al. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. **Bioinformatics**, v. 31, n. 19, p. 3210-3212, 2015.

SINGH, Naveen Kumar et al. Characterization of a pathogen induced thaumatin-like protein gene AdTLP from *Arachis diogeni*, a wild peanut. **PLoS one**, v. 8, n. 12, p. e83963, 2013.

SOUZA, Eduardo Pereira de. **Cistatinas de maqui: produção recombinante e inibição de cisteíno peptidases**. 2022. Dissertação de Mestrado. Departamento de Genética e Evolução, Universidade Federal de São Carlos, São Carlos.

TRENTINI, Giulia E. et al. Elicitation of phenylpropanoids in maqui (*Aristotelia chilensis* [Mol.] Stuntz) plants micropropagated in photomixotrophic temporary immersion bioreactors (TIBs). **Plant Cell, Tissue and Organ Culture (PCTOC)**, v. 146, n. 3, p. 607-619, 2021.

TRUDEL, Jean et al. Several thaumatin-like proteins bind to  $\beta$ -1, 3-glucans. **Plant physiology**, v. 118, n. 4, p. 1431-1438, 1998.

ULLAH, Abid et al. Osmotin: a plant defense tool against biotic and abiotic stresses. **Plant Physiology and Biochemistry**, v. 123, p. 149-159, 2018.

VAN LOON, Leendert C.; REP, Martijn; PIETERSE, Corné MJ. Significance of inducible defense-related proteins in infected plants. **Annu. Rev. Phytopathol.**, v. 44, p. 135-162, 2006.

VELAZHAHAN, Rethinasamy; DATTA, Swapan K.; MUTHUKRISHNAN, Subbaratnam. The PR-5 family: thaumatin-like proteins. **Pathogenesis-related proteins in plants**, p. 107-129, 1999.

WANG, Hexiang; NG, Tzi Bun. Isolation of an antifungal thaumatin-like protein from kiwi fruits. **Phytochemistry**, v. 61, n. 1, p. 1-6, 2002.

WATERHOUSE, Andrew et al. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic acids research**, v. 46, n. W1, p. W296-W303, 2018.

YASMIN, Nusrat et al. Molecular characterization, structural modeling, and evaluation of antimicrobial activity of Basrai thaumatin-like protein against fungal infection. **BioMed Research International**, v. 2017, 2017.

## APÊNDICE I

### MEIOS DE CULTURA, SOLUÇÕES E TAMPÕES UTILIZADOS

#### ***Meio Luria-Bertani low salt***

O meio Luria-Bertani (LB) *low salt* é composto por 1% de tripton, 0,5% de extrato de levedura e 0,5% de cloreto de sódio em água destilada. Em caso de meio sólido, há a adição de 1,5% de ágar. O meio de cultura foi esterilizado por autoclavagem a 121 °C, 1 atm por 20 minutos e mantido à temperatura ambiente.

#### ***Meio Extrato de Levedura, Peptona e Dextrose (YEPD)***

O meio YEPD é composto por 1% de extrato de levedura, 2% de peptona e 2% de dextrose em água destilada. Em caso de meio sólido, há a adição de 2% de ágar. O meio de cultura foi esterilizado por autoclavagem a 121 °C, 1 atm por 20 minutos e mantido à temperatura ambiente.

#### ***Meio Extrato de Levedura, Peptona, Dextrose e Sorbitol (YPDS)***

O meio YPDS é composto por 1% de extrato de levedura, 2% de peptona, 2% de dextrose, 2% de ágar e 1 M de sorbitol em água destilada. O meio de cultura foi esterilizado por autoclavagem a 121 °C, 1 atm por 20 minutos e mantido à temperatura ambiente.

#### ***Meio Glicerol Complexo Tamponado (BMGY)***

O meio BMGY é composto por 1% de extrato de levedura, 2% de peptona, Tampão Fosfato de Potássio 100 mM pH 6,0, 1,34% de YNB, 0,00004% de biotina e 1% de glicerol. O meio BMGY é montado na hora do uso.

O extrato de levedura, a peptona e o tampão fosfato de potássio pH 6,0 são esterilizados por autoclavagem a 121 °C, 1 atm por 20 minutos e mantidos à temperatura ambiente até o uso. O glicerol é esterilizado por autoclavagem a 121 °C, 1 atm por 20 minutos e mantido à 4° C até o uso.

O YNB e a biotina são esterilizados por filtração em filtros de 0,22 µm. Aquele é mantido à 4 °C e esta à -20 °C até o uso.

#### ***Meio Metanol Complexo Tamponado (BMMY)***

O meio BMMY é composto por 1% de extrato de levedura, 2% de peptona, Tampão Fosfato de Potássio 100 mM pH 6,0, 1,34% de YNB, 0,00004% de biotina e 0,5% de metanol. O meio BMMY é montado na hora do uso.

O extrato de levedura, a peptona e o tampão fosfato de potássio pH 6,0 são esterilizados por autoclavagem a 121 °C, 1 atm por 20 minutos e mantidos à temperatura ambiente até o uso.

O YNB, o metanol e a biotina são esterilizados por filtração em filtros de 0,22 µm. Os dois primeiros são mantidos à 4 °C e o último a -20 °C até o uso.

#### ***Tampão de equilíbrio e tampão de eluição***

O tampão de equilíbrio é constituído por 10 mM Tris-HCl, 50 mM fosfato de sódio e 100 mM cloreto de sódio, pH 8,0. O tampão de eluição utilizado é constituído dos mesmos

componentes do tampão de equilíbrio, com a adição de imidazol em diferentes concentrações (10, 25, 50, 75, 100 e 250 mM).

#### ***Solução de ácido-3,5-dinitrossalicílico (DNS)***

A solução de DNS é composta por 32,9 mM de ácido-3,5-dinitrossalicílico, 350 mM de hidróxido de sódio, 76,6 mM de tartarato de sódio e potássio, 0,54% de fenol e 30 mM de metabissulfito de sódio em água destilada. A solução foi mantida a 4 °C sob o abrigo de luz.

#### ***Tampão fosfato salino (PBS)***

O tampão PBS é constituído por 137 mM de cloreto de sódio, 2,7 mM de cloreto de cálcio, 10 mM fosfato de sódio dibásico e 1,8 mM de fosfato de potássio monobásico em água destilada e pH 8,0. O PBS foi esterilizado por autoclavagem a 121 °C, 1 atm por 20 minutos e mantido à temperatura ambiente.

#### ***Tampão de amostra Laemmli 3X***

O tampão de amostra Laemmli 3X é composto por 188 mM de Tris-HCl pH 6,8, 30% de glicerol, 6% de sódio dodecil sulfato, 0,5% de azul de bromofenol e 15% de  $\beta$ -mercaptoetanol. A solução foi mantida a temperatura ambiente sob o abrigo de luz.

**APÊNDICE II**  
**SEQUÊNCIAS DE NUCLEOTÍDEOS E DE AMINOÁCIDOS**  
**DAS *THAUMATIN-LIKE PROTEINS* DO MAQUI**

**Sequência de nucleotídeos da AcTLP1**

AAAAGCAATTAATCTAAGTATGAATTCAGCAATGAAAACCTCAAGTACTCTTTGGCCTTACCTTGGCCT  
 TCCTCTTCTCAGGTGTTCTCTCGGCTACTATCACCATCACAAACAACCTGCCGTACACAATCTGGCCA  
 GGAACCCTAACTGGTGGTGGCGGTGCTCAATTATCAACTACTGGGTTTGAGTTAGCCTCCAGAGCAT  
 CATCATCCATTAATGTTCTGCTCCATTTACAGGCCGGTTCTGGGCCCGAACACAGTGCTCCACGGCC  
 TCCGGTAGGTTCACTTGTCTTACCGCAGACTGTGCTTCTGGCCAGGTCGCATGCAATGGTGCTGGTG  
 CAGTTCACCCGCTTCCCTCGCAGAGTTCACTCTAGCAGCAAGTGGTGGACAAGATTTCTACGATAT  
 TAGCCTTGTTGATGGCTTCAACTTGCCCTTACAATTACCCCGCAAGGCGGATCTGGTCCTACTTGCA  
 CTGTCACAAGCTGTACCGCCAATGTGAACGCAGTTTGCCTGCAGAGCTTCAAGTTAAAGGGTCTG  
 ATGGGAGTGTTATCGCTTGCAAAGCGCTTGTGCGGCTTTAAATCAGCCACAGTATTGTTGCACTGG  
 TGATTTTAGTACACCAGAAACGTGCCCGCCACGAGCTATTCTCAGATCTTCGAGAATCAGTGCCCT  
 CAAGCTTATAGTTATGCTTATGATGACCTTTCTAGCACTTTTGCCTGCACTGGCTCTCCTAACTACGTT  
 ATAACCTTCTGTGGATGAGATAGAAACACACCTCCTCTATACATATGATAATTCAGTTGTCATGTAAT  
 AAATAAGAGAAGTCATCATAGCCAAGGATATATCCGTCATGAAAATTGATATGAATAAATATGTTTT  
 AATTTTATGTTTG

Destacado azul: UTR 5'.

Destacado amarelo: sequência codificadora do peptídeo sinal.

Destacado vermelho: UTR 3'.

**Sequência de aminoácidos da AcTLP1**

MNSAMKTQVLFGLTLAFLFSGVLSATITITNNCPYTIWPGTLTGGGGAQLSTTGFE LASRASSSINVPAPF  
 TGRFWARTQCSTASGRFTCLTADCASGQVACNGAGAVPPASLAEFTLAASGGQDFYDISLVDGFNLPLTI  
 TPQGGSGPTCTVTSTANVNAVCPAELQVKGSDGSVIACKSACAALNQPQYCCTGDFSTPETCPPTSYS  
 QIFENQCPQAYSAYDDLSTFACTGSPNYVITFCG

Destacado amarelo: peptídeo sinal putativo.

**Sequência de aminoácidos da AcTLP1 recombinante expressa em *Pichia pastoris***

EFATITITNNCPYTIWPGTLTGGGGAQLSTTGFE LASRASSSINVPAPFTGRFWARTQCSTASGRFTCLTAD  
 CASGQVACNGAGAVPPASLAEFTLAASGGQDFYDISLVDGFNLPLTITPQGGSGPTCTVTSTANVNAV  
 PAELQVKGSDGSVIACKSACAALNQPQYCCTGDFSTPETCPPTSYSQIFENQCPQAYSAYDDLSTFAC  
 TGSPNYVITFCGVDHHHHHHH

Sublinhado: *his tag*.

Destacado cinza: aminoácidos oriundos do pPICZaA.

### Sequência de nucleotídeos da AcTLP2

CTTCATTAAGATTTTTGAATAACTTTATGAAGATAATGAAGTAGGCTAGTAGCCGCCTCCCACCTTAT  
 TAAATACAATGTACACTCCCTACCTATATAAACTCCCCTTGTGTGCCTCTTGAACCAAGCACAAAGACA  
 CATAAAAAAAGGTAGCAACAATGAGCTACTACACCAATCTCCCCTTCTCTTGCCTTCTCTTGAGCCTC  
 CTCTTCATTTCAACCCATTCAGCCACCTTTGTAATCCGCAATGAATGCTCTTACACTGTCTGGGCCGC  
 AGCCTCGCCTGGTGGAGGCCGCCGCCTAGACCGTGGCCAACTTGGACCCTAACGTGCCCGCTGG  
 CACGGCTATGGCACGTATTTGGGGCCGAACAAATTGCAATTTTGATGGAAGTGGTAAGGGTCATTGC  
 CAAAGTGGTACTGTGGTGGAGTCTAGAGTGTAAGGCTGGGGTAACCCTCCAAACACACTAGCT  
 GAATACGCACTAAACCAATTTGGTAACATGGATTTTATTGATATATCCCTAGTAGATGGATTTAATATCC  
 CCATGGATTTTAGTCCAACACTACTGGTGGTTGCCGAGGAATTGTGTGCACTGCAGATATTAACGGGAA  
 ATGCCCTAGCCAATTGAAGGCTCCAGGTGGGTGTAACAATCCATGTACAGTTTTTAAGACTAATGAAT  
 ATTGTTGCACTAATGGGTATGGTAGCTGTGGACCGACCAATTTTCAAAGTTTTTCAAGGGTGCCTGC  
 CCTGATGCTTATAGCTACCCCCAGGATGATCCTACAAGCACATTTACATGCCCTGGTGGGACTAACTA  
 CAGGGTTGTCTTTTGCCTAAAAGGGTCTGCTCATTCCCTCTTGAGATGGTTGAAGGGAAGAATGAG  
 GAGTAGAGGGGTCTTGAAGTGGTTGTGACTTATGAGGCTTCAAAAACCTATAGTGAAAGTAAGGTC  
 TGCCTAAGATGCATCTATAGTTGACTTGTATGTGTTCTTTTTGTTTTTGTGTGGTGTATGATGGGGT  
 GTAATGAAAAATCCATCGAAATAAAGATCTTTGATCAAAGATATCTTTAATCCATCGAATTTGATGCAT  
 AAGATCAATGTTTTTTTTAATAAGATCAATGTTTGATCAGTTTga

Destacado azul: UTR 5'.

Destacado amarelo: sequência codificadora do peptídeo sinal.

Destacado vermelho: UTR 3'.

### Sequência de aminoácidos da AcTLP2

MSYYTNLPFSCLLLSLLFISTHSATFVIRNECSYTVWAAASPGGRRRLDRGQTWTLNVPAGTAMARIWG  
 RTNCNFDGSGKGHCQSGDCGGVLECKGWGNPPNTLAEYALNQFGNMDFIDISLVDGFNIPMDFSPTTG  
 GCRGIVCTADINGKCPSQLKAPGGCNPCTVFKTNEYCCTNGYGSCGPTNFSKFFKGACPDAYSYPQDD  
 PTSTFTCPGGTNYRVVFCPKGSAHFPLEMVEGKNEE

Destacado amarelo: peptídeo sinal putativo.

### Sequência de aminoácidos da AcTLP2 recombinante expressa em *Pichia pastoris*

EFATFVIRNECSYTVWAAASPGGRRRLDRGQTWTLNVPAGTAMARIWGRRTNCNFDGSGKGHCQSGDC  
 GGVLECKGWGNPPNTLAEYALNQFGNMDFIDISLVDGFNIPMDFSPTTGGCRGIVCTADINGKCPSQLK

APGGCNPCTVFKTNEYCCTNGYGSCGPTNFSKFFKGACPDAYSYPQDDPTSTFTCPGGTNYRVVFCPK  
GSAHFPLEMVEGKNEEVDHHHHHH

Sublinhado: *his tag*.

Destacado cinza: aminoácidos oriundos do pPICZaA.

### Sequência de nucleotídeos da AcTLP3

CACTGCAAAGCAACAACAAAGATTTTCATAAACCAAATTCAAACCTTGAAAAATGCTAACTCCTTC  
AGTCTTTTCACTTCTCCTCTATCTACTAGCATCATTCTCCACTACAAGTGGATCACAAATTATTCTAGT  
CAACAACTGCAAGGAAAGCATATGGCCCGGAATACTAGGCGGTGCAGGCCAGCCAACCCCTAAAGA  
TGGTGGTTTCTTCTTAGCAGTGGCGAGGAAGTAGTCCTTGATGTGCCCGAAAAGTGGTCAGGGAG  
AATATGGGGCAGGCAAGGATGTTCTTTTGACAATAATGGGAAAGGCTCTTGTGACTGTTGACTGC  
TCCGGGTTGCTTCGTTGCCAAGGCACTGCTGGCTTGCCCCAGCAACTGTGGTTGAAATGACTCTTG  
GGACATCATCTTCGCTCTACATTTCTATGATGTTAGTTTGGTTGATGGTTTCAATTTGCCGGTTTCAA  
TGAAACCTGTTGGGGGTGGAATTGGGTGTGGTGTGCATCGTGTGAAGTTGATTTGAACATTTGTTG  
TCCTTCAGCGTTGGAAGTGAAGAAAGGGAACAAAGTGGTGGGATGCAAGAGTGCCTGCTTGGCTAT  
GCAATCTGCTAAGTATTGTTGCACAGGAGAGTATGCAAACCCAAACACTTGCAAGCCTACACTTTTT  
GCTCATCTATTTAAGGCTATATGCCCAAAGGCCTATAGTTATGCTTTTGATGACTCCTCCAGCCTAAC  
AAATGCAGGGCTTACGATATGTTATCACCTTTTGCCCCCAAATGAGGAGTGATGCAAGAATAAG  
AGGGCAGTCATGGCTATGCATTCTGCAAGGAAGTCAATGTCATTAGTGACAGATTATAATAGTTAAAT  
TACGGAATGGTTTCCGTTTTTTTATTTTCTTTTTAATAATTTTATTTCTGTTTCTGTTTCTGTCTCTGTAA  
GTAAATAATAGTAACAAGTGGAGTATGGAGTAAGAAATGGATAGAAAAAGAAAATTTTCTCTTGGG  
ATAGGGCTACTGGTTAGATGTTTTGGCATGTAAGTTTATGAATAGAATT

Destacado azul: UTR 5'.

Destacado amarelo: sequência codificadora do peptídeo sinal.

Destacado vermelho: UTR 3'.

### Sequência de aminoácidos da AcTLP3

MLTPSVFSLLLYLLASFSTTSGSQIILVNNCKESIWPILGGAGQTPKDGGLSSGEEVLDVPEKWSG  
RIWGRQGCSEFDNNGKGSCTDGDSCGLLRQCQTAGLPPATVVEMTLGTSSSPLHFYDVSIVDGFNLPVS  
MKPVGIGGIGCGVASCEVDLNICPSALEVKKGNKVVGCKSACLAMQSAKYCCTGEYANPNTCKPTLFA  
HLFKAICPKAYSFAFDDSSSLNKRASRYVITFCPPK

Destacado amarelo: peptídeo sinal putativo.

### Sequência de nucleotídeos da AcTLP4

CAAAATGAGCTCCTTCAAACCCCTTTCTTTCTTTTCCTTCCTTCTTGTTACCATTTACTTCACCTTAGC  
 CCATGCTGCCACCTTTGATATAACAAACAGCTGCCCCTACACTGTCTGGGCCGCGGGTGTGCCCGGT  
 GGTGGCAAGCAACTTAACCAAGGTGAAACATGGACCATCAGTGCTAACCCCTGGCACTACACAAGCA  
 CGTATTTGGGCTCGAACCAATTGCCAATTTGATGCATCAGGCAGAGGCAAGTGTGAGACTGGTGACT  
 GTAATGGTCTCCTCCAATGTGATGCCATGGTGTACCTCCTAACACTCTAGCTGAGTACGCACTAAAA  
 CAATATATGGACCAAGACTTCATTGATATCTCTGTAATCGATGGATTTAATGTTCCAATGGAGTTTAGC  
 GCAGCCTCTGGTGGGTGCACCCGTGTGATCAAATGTACAGGAGATATTATTGGTCAATGTCCGAATG  
 AGTTGAAGGTTACTGGAGGATGTAACGGGGCATGTCCAGTTTTCAAGACTGACCAGTATTGTTGCAA  
 TTCTGGTAACTGTGGACCTACAGATTTTTCTAAGTTTTTTAAGGATAGGTGCCCAGATGCTTATAGTTA  
 CCCTAAAGATGATCCACAAGTTTGTTTACTTGTCTTAGTGGAACCTAATAAGGTTATATTCTGCC  
 CTAAAGGCTAAAAGTAATCCCAAGAGTCATAGAAAAAAGTAGTCTATATATGGTCCAGCAAAAAA  
 AAAAGTACTCTATATATGACATGGAATAATTGAATAAAAGTTCAAAGTATTTTTGAGCACGTCTTGTTA  
 CCTGCTCAAGTTTTAATATTTTATGTGTAATTGTTTTTTTTTTTTATTAGAACAAAGTTTTAAAATACGA  
 GGTCAAGAAAGCGCTTGAGACACCACCAGGACCACTAAAAAGAATTAGACGTCCCCAGTTATCACC  
 CTGACATCCAGCATCTTTGAATAGAATATGCGTTTACACTAATATTTCTCTTTATCATCAGGATTCAG  
 AAATTCCAATTCTCATGTGAAACTCGATCATGCGTACAATAAAAAATAATAAATGAAGAGGTGGTGTTA  
 TTTAAAATTAT

Destacado azul: UTR 5'.

Destacado amarelo: região codificadora do peptídeo sinal.

Destacado vermelho: UTR 3'

### Sequência de aminoácidos da AcTLP4

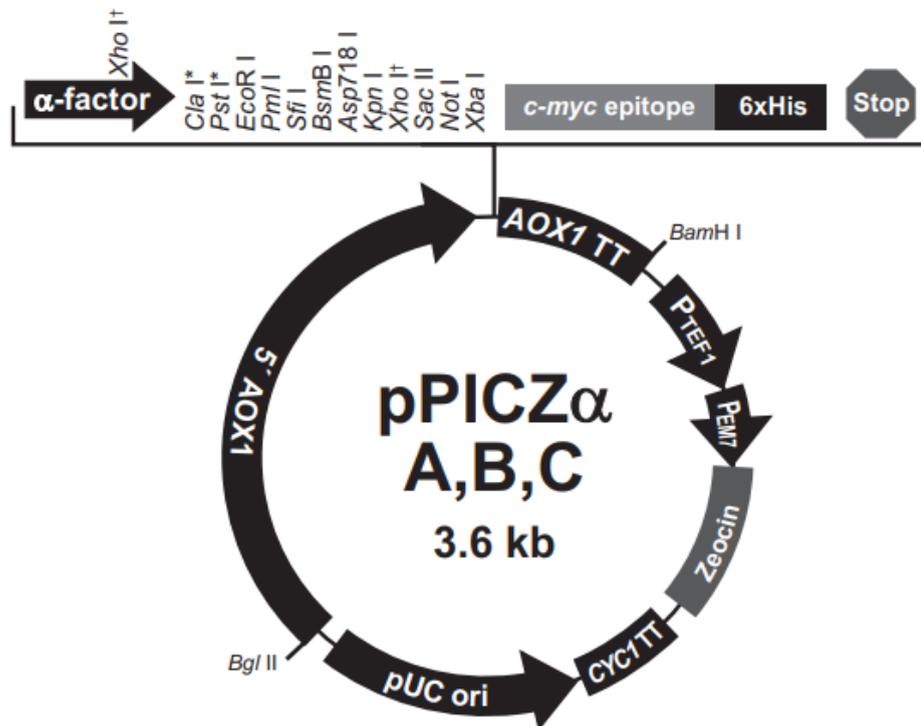
MSSFKLSFFSFLLVTIYFTLAHAATFDITNSCPYTVWAAGVPGGGKQLNQGETWTISANPGTTQARIWA  
 RTNCQFDASGRGKCETGDCNGLLQCDAYGVPPNTLAEYALKQYMDQDFIDISVIDGFNVPMEFSAASG  
 GCTRVIKCTGDIIGQCPNELKVTGGCNGACPVFKTDQYCCNSGNCGPTDFSKFFKDRCPDAYSYPKDDP  
 TSLFTCPSGTNYKVIFCP

Destacado amarelo: peptídeo sinal putativo.

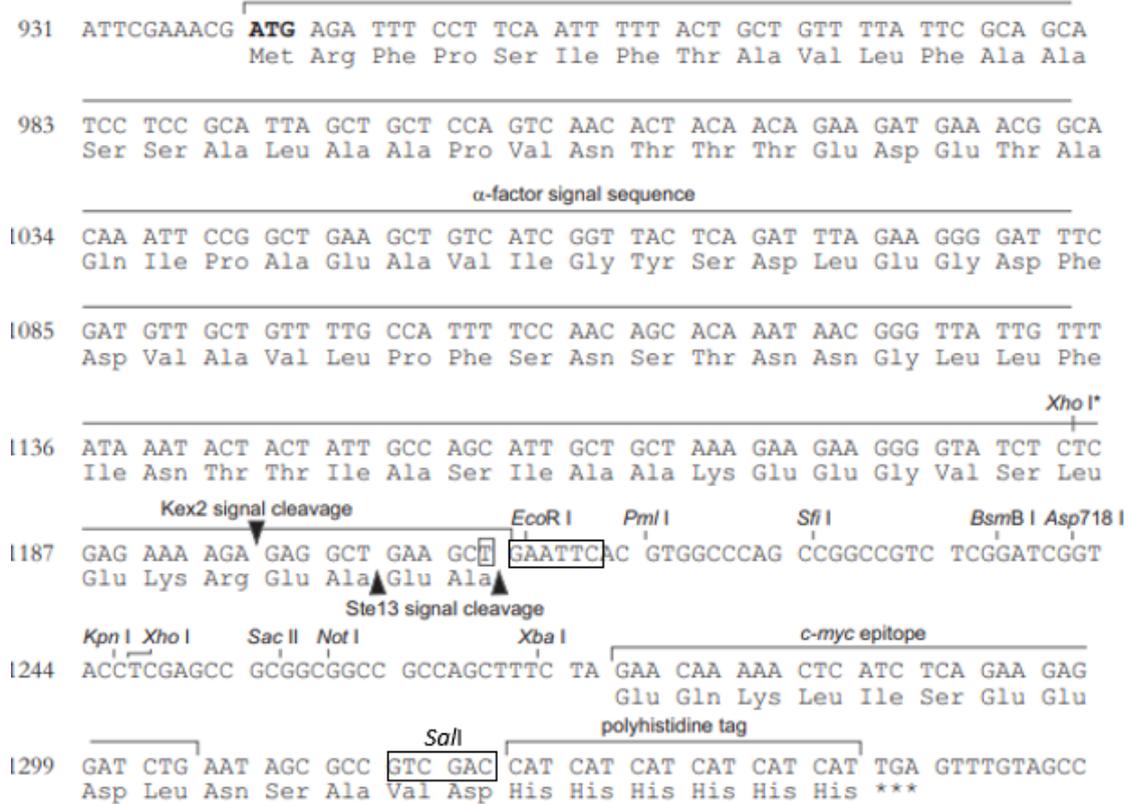


**Figura 24. Alinhamento global entre as AcTLPs e TLPs de outras espécies vegetais.** Os resíduos considerados altamente similares estão destacados em negrito e dentro de uma caixa vazia de contorno preto. Nos resíduos com identidade estrita, a coluna é colorida com fundo preto. Os códigos de acesso estão no nome de cada sequência de aminoácidos. As sequências utilizadas estão no Anexo II. Art: *Arabidopsis thaliana*. Pim: *Pinus monticola*. Php: *Physcomitrella patens*. Ors: *Oryza sativa* subsp. *japonica*.

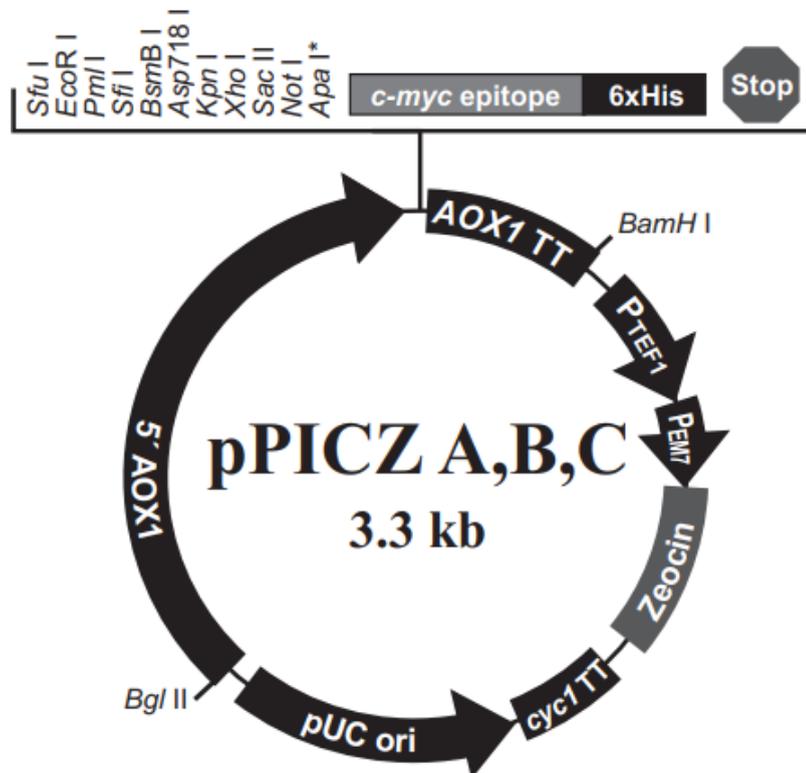
**ANEXO I**  
**MAPA E CARACTERÍSTICAS DOS VETORES pPICZA e pPICZaA**



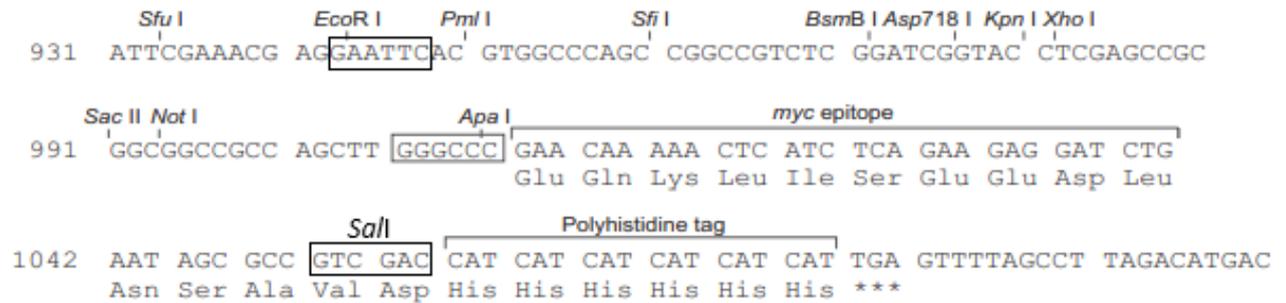
**Figura 25. Mapa geral do vetor pPICZaA.** 5' AOX1: promotor do gene da álcool oxidase 1.  $\alpha$ -factor: peptídeo sinal de secreção fator alfa de *Saccharomyces cerevisiae*. 6xHis: cauda de polihistidina. AOX1 TT: terminador de transcrição do gene da álcool oxidase 1. Zeocin: gene de resistência ao antibiótico Zeocin™ (Fonte: EasySelect Pichia Expression Kit, 2001).



**Figura 26.** Sequência do sítio múltiplo de clonagem (MCS) do vetor pPICZaA. O códon iniciador (ATG) está indicado em negrito. As caixas representam as regiões das enzimas de restrição utilizadas (*EcoRI* e *SalI*) (Fonte: EasySelect Pichia Expression Kit, 2001).



**Figura 27. Mapa geral do vetor pPICZA.** 5' AOX1: promotor do gene da álcool oxidase 1. 6xHis: cauda de polihistidina. AOX1 TT: terminador de transcrição do gene da álcool oxidase 1. Zeocin: gene de resistência ao antibiótico Zeocin™ (Fonte: EasySelect Pichia Expression Kit, 2001).



**Figura 28. Sequência do sítio múltiplo de clonagem (MCS) do vetor pPICZA.** As caixas representam as regiões das enzimas de restrição utilizadas (*EcoRI* e *Sall*). O códon iniciador foi inserido após o sítio da *EcoRI*, usando os oligonucleotídeos indicados nas tabelas 1 e 2 (Fonte: EasySelect Pichia Expression Kit, 2001).

**ANEXO II**  
**SEQUÊNCIAS DE TAUMATINA E *THAUMATIN-LIKE PROTEINS* UTILIZADAS**  
**PARA ALINHAMENTO GLOBAL**

>AcTLP1

MNSAMKTQVLFGLTLAFLFSGVLSATITITNCPYTIWPGTLTGGGGAQLSTTGFEASRASSINVPAPFTGRFWA  
 RTQCSTASGRFTCLTADCASGQVACNGAGAVPPASLAEFTLAASGGQDFYDISLVDGFNLPLTITPQGGSGPTCTVT  
 SCTANVNAVCPAELQVKGSDGSVIACKSACAALNQPQYCCTGDFSTPETCPPTSYSQIFENQCPQAYSAYDDLST  
 FACTGSPNYVITFCG

>AcTLP2

MSYYTNLPFSCLLLSLLFISTHSATFVIRNECSYTVWAAASPGGGRRLDRGQWTWLNVPAGTAMARIWGRNCFDG  
 SGKGHCQSGDCGGVLECKGWGNPPNTLAEYALNQFGNMDFIDISLVDGFNIPMDFSPPTGGCRGIVCTADINGKCP  
 QLKAPGGCNPCTVFKTNEYCCTNGYGSCGPTNFSKFFKACPDAYSYPQDDPTSTFTCPGGTNYRVVFCPKGSAHF  
 PLEMVEGKNEE

>AcTLP3

MLTPSVFSLLLYLLASFSTTSGSQIILVNNCKESIWPGILGGAGQPTPKDGGFLLSSGEEVLDVPEKWSGRIWGRQ  
 GCSFDNNGKGSCTGDGCSGLLRQCAGTGLPPATVVEMTLGTSSSPLHFYDVSLVDGFNLVSMKPVGGGIGCGVASC  
 EVDLNICCPSALEVKKGNKVVGCKSACLAMQSAKYCCTGEYANPNTCKPTLFAHLFKAICPKAYSAYAFDDSSSLNKC  
 RASRYVITFCPPK

>AcTLP4

MSSFKTLFFFLLVITYFTLAHAATFDITNCPYTVWAAAGVPGGKQLNQGETWTISANPGTTQARIWARTNCQFD  
 ASGRGKCTGDCNGLLQCDAYGVPNTLAEYALKQYMDQDFIDISVIDGFNVPMEFSAASGGCTRVIKCTGDIIGQC  
 PNELKVTGGCNGACPVFKTDQYCCNSGNCGPTDFSKFFKDRCPDAYSYPKDDPTSLFTCPSGTNYKVIFCP

>BAF44567.1 preprothaumatin I [Thaumatococcus daniellii]

MAATTCFFFLFPFLLLLSRAATFEIVNRCSYTVWAAASKGDAALDAGGRQLNSGESWTINVEPGTNGG  
 KIWARTDCYFDDSGSGICKTGDCGGLLRCKRFRPPTLAEFSLNQYKDYIDISNIKGFNVPMDFSPTT  
 RGCRGVRCAADIVGQCPAKLKAPGGCNDACTVFQTSYCCCTTGKCGPTEYSRFFKRLCPDAFSYVLDKP  
 TTVTCPGSSNYRVTFCTALELEDE

>GAV58933.1 Thaumatin domain-containing protein [Cephalotus follicularis]

MHSAMKTQVIFGLTLAFLLAGVFSATVTITNCPYTIWPGTLTGGGGNQLSTTGFEASGASNSFNVPPP  
 FTGRFWARTQCSTFGAVTCATADCASGQVACNGAGAIPPASLAEFTIAPSGGQDFYDISLVDGYNIPLSI  
 APKGGSGPTCTTTSCSADVNAVCPAELQVKGSDGRVIACKSACDAFSQPQYCCTGAYNTPATCPPTSYSN  
 IFKQCPQAYSAYDDSTSTFSCTGSPNYVITFCP

>XP\_021601423.2 thaumatin-like protein [Manihot esculenta]

MSSFNISFFLVAALYYFTFAHAATFDITNKCPYTVWGAASPGGGRELKTGETWTITANPGTTQARIWART  
 NCQFDASGKGCETGDCNGLLVCQGYGAAPNTLAEYALDQFERQDFIDISVIDGFNVPMEFSSASGSCSR  
 VIKCTADIIGQCPNELKVPGGCNGPCPVFKTEEHCCNSGNGCPTNFSKYFKDRCPDAYSYPKDDPTSFLT  
 CPTGTNYKVIIFCP

>NP\_173261.1 Pathogenesis-related thaumatin superfamily protein [Arabidopsis thaliana]

MASINLFLFAFLLLLSSHASASTVIFYNKCKHPVWPGIQPSAGQNLLAGGGFKLPANKAHSLLQLPPLWGR  
 FWGRHGCTFDRSGRHCATGDCGGSLSCNGAGGEPATLAEITLGPELDIFYDVSLLVDGYNLAMSIMPVK  
 SGQCSYAGCVSDLNQMCVGLQVRSRNGKRVVACKSACSAFNSPQYCCTGLFGNPQCKPTAYSKIIFKVA  
 CPKAYSAYDDPTSIAATCSKANYIVTFCPHHRH

>NP\_001049473.1 Os03g0233200 [Oryza sativa Japonica Group]

MEFGRTPALQIIVGVLWSQLQFGAEAVGTTVFTLRNNTYTVWPATLSGNTAVAVGGGGFELSPGANVSF  
 PAPAGWSGRLWARTDCAPSGTASLACVTGDCGGAVSCSLGGAPPVTLAEFTLGGTDGKDFYDVSLLVDGYN  
 VGIGVAATGARVNRSTCGYAGCVGDVNALCPAELQVAGKENDQQSGAAATTTVACRSACEAFGTAEYCCT  
 GAHGGPDSGCPTRYSLFKAACPAAYSAYDDPTSTFTCGTGAQYVITFCPAQQQ

>ADB97925.1 thaumatin-like protein L1 [Pinus monticola]

MGLILRSRALALLLILFSGGAIMVICAKTSPLTITIMNSCPTTIWPGQLQASEGHDVLEQGGFALESLSRSFSVANP  
 WNGVWVWARTGCSFSGEKGSCLTGDCDGKLCNGTGGRNVPVTLAQLSLHHGGNDVSSYTLVSLVNGFNPLPLTFTPHGGR  
 GRCGIPRCMANMLESCPKEQLQVKNVDNVIACKTACQAFRTDAYCCTSQHYNGSRTCPTTYSQIFKRACPNFAFAYPD  
 DNPALVHNCVTPNEIKLIFCH

>XP\_001765997.1 predicted protein [Physcomitrella patens]

MMLVMKIKSYILDHALCAEGVTFTFINHCKFTVWVGVPNGGLPLLVGGGFELAAGKQDAVTASASWGR  
 FWGRTGCKFDSAGKNCETGDCGGVLKCGGAGGNPPASLAEITLNGADGNDFYDISLVDGYNLPLSMAPS  
 GGTGKCGAPGCISNLNDNCPAALQFLAEGVLVGCNSACNAFNTPEYCCTGAFFGGPTTCPPTQYSMAFKSA  
 CPTAYSAYDDATSTFTCKGANYAITFCPTGIFIPDLACTMKQ

>NP\_502360.1 THaumatIn family [Caenorhabditis elegans]

MALIKLTLAVLLLALGAEARKITIYNKCPFTIWPGLGPNPSGGGFKLNAGQSRDINVDDAWTAGRVWA  
 RTGCDGNFNCETGFCRNSEQCNGAGGVPPASLAEFTLKAAGGQDFYDVSLLVDGYNLPLVLDPHGGSGCKR  
 AGGCVKDINAECPAALSVKHNGKTVACKSGCLGYNTDQECCRGAYGTPDKCHRSATAQMFKDACPTAYS  
 YAYDDGSSTFTCQPSASYTVQFC

>XP\_660683.1 hypothetical protein AN3079.2 [Aspergillus nidulans FGSC A4]

MMFTKALVAATLTLTAALPQPTVVRREGGDAGVTIVNMDSDVYAWSVTDGVSKMHTLSSGGGSYTENF  
 QANPNGGGVSIKLSHQDQTDVLQFEYTKSGETIFWDMSCIDMDRAASTFTKNGFDVSPSQTSGDCPAVN  
 CHAGDTSACAEAYLQPKDDHATHGCPIDTSFTLTLGAPSEFLFAVLIQEYSYIFPISVTICPQSSPFLATY

LPLFKSA

>XP\_001765997.1 predicted protein [Physcomitrella patens]

MMLVMKIKSYILDHALCAEGVTFTFINHCFTVWVGVQPNGGLPLLVGGGFELAAGKQDAVTASASWGGR  
 FWGRTGCKFDSAGKGNCTGDCGGVLKCGGAGGNPPASLAEITLNGADGNDFYDISLVDGYNLPLSMAPS  
 GGTGKCGAPGCISNLNDNCPAALQFLAEGVLVGCNSACNAFNTPEYCCTGAFFGGPTTCPPTQYSMAFKSA  
 CPTAYSAYDDATSTFTCKGANYAITFCPTGIFIPDLACTMKQ

>XP\_001769333.1 predicted protein [Physcomitrella patens]

MSRANLVPFLIVLQVCCMLQVAEAAATFNIINGCQFTVWVGVQPNKLPVLANGGFACASGDKVAVTAPQ  
 GWGGRFWGRTDCKFDAAGKGLCVTGDCGNVLCNGAGGNPPASLAEITLNGSGGLDFYDISLVDGYNLPI  
 SMKPIGGTGECCGATGCISDLNTKCPEALKFWSGFKVAGCKSACAAFNPEQYCCTGAYNTAATCPPTQYSK  
 AFKAACPTAYSAYDDATSTFTCKALVYDITFCPPGTA

>XP\_001784610.1 predicted protein, partial [Physcomitrella patens]

FTFVNRCKFPVWVGVQPNGGLAILADGGFELGAGKQDAITAPASWGGRFWGRTGCEFDSDAGKGTCDTGDC  
 GGALKCGSGGNPPATLAEITLHGADGNDFYDISLVDGYNLPLAMTPAGGTGKCGAPECCSSNLNDMCPEV  
 LQFLSEGSVIGCNSACNAFSEPKYCCTGAHDQPATCPPTQYSNAFKNACPLAYSYAYDDATSTFTCKAAN  
 YAIAFCPNGT

>ADB97926.1 thaumatin-like protein L2 [Pinus monticola]

MAGRNSLSGSIWITVTVTLVAVNVYLQVAEGATFTVRNQCTYTVWAAASPGGGKALGQGQTWTFNVAAGT  
 KGARIWGRGTGCSFDASGRGRCNTGDCGGLLNCQGYGVSVPATLFEYALNQYQNDYDISLVDGFNIPLSA  
 TPSNSNCKKIGCTSNINAICPSQLKVTGCKSACAAFNTPQYCCTGAYLNNCSPTNYSKFFKQCPQAYS  
 YAKDDATSTFTCPSGANHNVVFCG

>ADB97927.1 thaumatin-like protein L3 [Pinus monticola]

MARALGVLFTFMAALYICVQVGCATFEIQNECPFTVWAAAGTPGVGGKQLERQSWTVQFPAGTTGRFWG  
 RTGCSFDGSGKGSCKTGDCGGLLNCQGSVSVPATLAEFAVNQFQNLDFYDVSLVDGFNLPMSSIIPSNQGC  
 TKSICSSDINSKCPAELKVSDGCKSACVAFNTPQYCCTGSFLDNCPPSDYSRFFKKECPQAYSYAKDDTT  
 STFTCPGGTNYKVVFCGTGTSHEN

>ADB97928.1 thaumatin-like protein L4 [Pinus monticola]

MEARTLSVSISVAVTLALIFNVYLPVGTGATFAIQNQCSTYTVWAAAGIPGGGKALQGGSWSVDVPAGTSA  
 GRFWGRTSCSFDASGQKCSSTGDCGGLLNCQASGQPPATLAEYTLNGGNNRDTYDISLVDGFNIPLSITP  
 SDASCTAPTCCSNINTICPSELKVTGDCNSACVAFNTPQYCCTGDYLTNCSPTNYSKIFKDQCPQAYS  
 YAKDDATSTFTCASGANYKIVFCP

>ADB97929.1 thaumatin-like protein L5 [Pinus monticola]

MVARNLSGSISWIWITVTVTLAINVYLQVAEGATLTVKNQCRYTVWAAASPGGGKALRQGQTWTFKVAAGT  
 TGRIWGRGTGCSFDASGRGKNCNTGDCGGLLNCQGYGVSVPATLFEYALNQYQNDYDISLVDGFNVPLSAT  
 PSNSNCKKIGCTSNINAICPSQLKVTGCKSASAAFNTPQYCCTGAYINNCSPNTNYSKFFKQCPQAYS

SNDDATGTFTCPSGANYNVFCG

>NP\_177640.1 thaumatin-like protein 3 [Arabidopsis thaliana]  
 MAKISSIHILFFVFITSGIADSATVFTLQNSCAYTVWPGTSLGNSITLGDGGFPLTPGASVQLTAPTGWS  
 GRFWARTGCNFDASGHGTCVTGDCGGVLKCTGGGVPPATLAEFTVGSSNAGMDFYDVSLVDGYNVKGMIK  
 PQGGFGNCKYAGCVSDINEICPSELRIMDPNSGSVAACKSACAAAFSSPEFCCTGAHATPQTCSPTYSSM  
 FKNACPSAYSAYDDASSTFTCTGSNYLITFCPTQS

>NP\_177641.1 pathogenesis-related protein 5 [Arabidopsis thaliana]  
 MANISSIHILFLVFITSGIAVMATDFTLRNNCPTTVWAGTLAQGPKLGDGGFELTPGASRQLTAPAGWS  
 GRFWARTGCNFDASGNRGCVTGDCGGRLRCNGGGVPPVTLAEFTLVGDGGKDFYDVSLVDGYNVKGIRPS  
 GSGDCKYAGCVSDLNAACPDMLKVMQNNVVACKSACERFNTDQYCCRGANDKPETCPPTDYSRIFKNA  
 CPDAYSAYDDETSTFTCTGANYEITFCP

>NP\_177642.2 Pathogenesis-related thaumatin superfamily protein [Arabidopsis thaliana]  
 MANYSSIHILFLVFITSGIAVSAATVFTLQNSCPYTVWPGILSGNDNTLGDGGFPLTPGASVQLTAPAGW  
 SGRFWARTGCNFDASGHGNCCTGDCGGVLKCNNGGVPPVTLAEFTLVGDGGKDFYDVSLVDGYNVEMGIK  
 PQGGSGDCHYAGCVADVNAVCPNELRLMDPHTGIIAACKSACAAFNSEFCCTGAHATPQTCSPTHYSAM  
 FKSACPGAYSAYDDATSTFTCTGSNYLISFCPTRN

>NP\_173365.2 Pathogenesis-related thaumatin superfamily protein [Arabidopsis thaliana]  
 MAIFSTSHLLFISFIIATCTISVSGTFTLTNHCSTIWPGLTANGAQLGDGGFALASGSSVFTVSPG  
 WSGRFWARTYCNFDASGSGKCGTGDCGSKLKCAGAGAPPATLAEFTIGSSGKKNVQDFYDVSLVDGYN  
 VQMKITPQGGSGDCKTAGCVSDVNAICPKELQVTGPSVAAACKSACEAFNKPEYCCTGAYSTPATCPPTN  
 YSKIFKQACPSAYSAYDDASSTFTCTNANYEISFCS

>NP\_195324.4 Pathogenesis-related thaumatin superfamily protein [Arabidopsis thaliana]  
 MECNCGSAEPPATLAEFTLGISGSLDFYDVSLVDGYNLPMIVEGSGLCPTTGCVTDLNQNCPTELKAKRS  
 LACRSACEAFGRPEYCCSGAYGSPDSCKPSMYSQVFKSACPKSYSAYDDATSTFTCSGADYITITFCPKL  
 SSLKSSKEPPPEETGGESGSGTDSSTEIQQPIEEASLANSYYAAMAIGESTRNTASFLLLLIFGWISL

>NP\_194149.2 THAUMATIN-LIKE PROTEIN 1 [Arabidopsis thaliana]  
 MIITVLHSHVSFYFIILSFLFFHALHLVGS DGATITIVNRCSTVWPGILSNVSGSDIGTTGFELVPGGS  
 RSFQAPASWSGRFWARTGCNFDNSDTGQGTCLTGDCSNQVECNAGAKPPATLAEFTIGSGPADPARKQD  
 FYDVSLVDGYNVPMLEASGGSEGTCLTTGCVTDLNQKCPTELRFSGSACKSACEAFGSPEYCCSGAYA  
 SPTECKPSMYSEIFKSACPRSYSYAFDDATSTFTCTDADYITITFCPSLPR

>NP\_568046.1 Pathogenesis-related thaumatin superfamily protein [Arabidopsis thaliana]

MMNLSSRRASSTLTLSFTVLLLASTGSYGSTFTFANRCGYTVWPGILSNAGSPTLSTTGFE L PKGTSRSL  
 QAPTGWSGRFWARTGCKFDSSSGTCKTGDCGSNAVE CAGLGAAPPVTLAEFTLGTGGDDFYDVS LVDGY  
 NIPMIVEVAGGSGQCASTGCTTDLNIQCPAELRFGDGDACKSACLAFRSPEYCCSGAYATPSSCRPSVYS  
 EMFKAACPRSYSYAYDDATSTFTCAGGDYTVTFPCSSPSQKSTSYSPVTDSSSTSQGS DVPVPGSDTG YA  
 GQQQQT PGQGNVYGSQGTGSEMGTGETMLQDGSW MAGLAMGEASRPAGVSLTVLLAAFTFPFIFS

>NP\_001031809.1 Pathogenesis-related thaumatin superfamily protein  
 [Arabidopsis thaliana]

MLKGSYGSTFTFANRCGYTVWPGILSNAGSPTLSTTGFE L PKGTSRSLQAPTGWSGRFWARTGCKFDSSG  
 SGTCKTGDCGSNAVE CAGLGAAPPVTLAEFTLGTGGDDFYDVS LVDGYNIPMIVEVAGGSGQCASTGCTT  
 DLNIQCPAELRFGDGDACKSACLAFRSPEYCCSGAYATPSSCRPSVYSEM FKAACPRSYSYAYDDATSTF  
 TCAGGDYTVTFPCSSPSQKSTSYSPVTDSSSTSQGS DVPVPGSDTG YAGQQQTPGQGNVYGSQGTGSEM  
 GTGETMLQDGSW MAGLAMGEASRPAGVSLTVLLAAFTFPFIFS

>NP\_001068535.1 Os11g0703000 [Oryza sativa Japonica Group]

MALAFVAVLATFAAAATPTTTNLT LHNLC THPVWPLVTANAGLPAIADAAGAATRLDGN DGLATLAFPP  
 GAWSGRVVARTGCRNGSSRCDTGDAPPVTV AQVSVHGAGGLAEYSVSLVDGFNVAVV VTPHGFEQGRLC  
 PSLGCAVDLAADCPGDGGRGGCMAAGQAEAFKARCPDTRTTP TDVEVTPQRCIQPAELKVVFCPPDNTSS  
 TH

>NP\_001067074.1 Os12g0569500 [Oryza sativa Japonica Group]

MANKLQLAIAFVVVVVVLGAMAASAAA AVAPMLTMHNLCPYTVWPIVSPDSGSPPIADGIRLEGRGVGLR  
 SLNLPAGFWSGRVVPRTWCRDGGRC DTGNAPPATVVRLS FNGAGGLAEYSVNLGEGFNVPTV VSPHAIGG  
 GMC PALGCTADLNAGCAAGQRVYGGDTGGDVVACRGPASYFKQRCPLTRTGGDVEPVPQH CISPGEIKL  
 VFCQAAMVAGEPELIRTVDVADN

>NP\_001067073.1 Os12g0569300 [Oryza sativa Japonica Group]

MAKQLALVLLVAVVAAAATSVAAATK LTLHNLC PYPVWPLVTPNTGFPSISGNTARLDGGGRGLVSYDFP  
 ASFWAGRVAR TGCGGGGLVRCETGNAPPATVVQLV VHSPEGAQDLAAYSVSLVDGFNVPAV VSPQAIA  
 GGGQCPALGCAADLNAGCPRSQRVVGAGAVVACRGTADYFKARCP LTRTTGSDVEPVPQHCLAPGELKV  
 VFCQPSMVAAAVPELIRTVVANI

>NP\_001067069.1 Os12g0568900 [Oryza sativa Japonica Group]

MASALAFVAVLLAAAAAATSPA AVAATTLTIQNLC PHPVWPLVTP TSGQPI SDNTARLDPN SLISLAFPP  
 TPWSGRVAARTGCDAASPPAGCETGASPPSTVAQLSVHGGDVATYSVSLVDGFNVV VVSPQAVGGGQ  
 CPALGCVVDLNCDCPLGQRFSDGAACRGPPEYFKGRCPQTRTTPGDVEPVPQSCRSPGELKVI FCPPTML  
 TAAAAAASDMLIRTVVASS

>NP\_001063786.1 Os09g0536400 [Oryza sativa Japonica Group]

MASWII TRSTWIMSVFFVILSFFQGLVMVGGVTFTFTNRCGGTVWPGVLSNSGSSALGTTGFALGAGETR  
 SLAAPAGWSGRFWARTGCTFDDDGKGT CATGDCGS GEVECRGAGATPPATLVEFTLGS GGGGGKDYDVS

LVDGYNLPMVVEAAAAGCPATGCVVDLNQRCPAELKAGHGQACRSACEAFGTPEYCCSGDHGNPDTCHPS  
 VYSQMFKRACPRSYSYAYDDATSTFTCTGTDYSITFCPRPGNPNSQKSSNDPSPRPKDPQLEDDSWLASL  
 ATGEVDGAAPASTSLLLQATLAVAVMALLVLH

>NP\_001064158.1 Os10g0146200, partial [Oryza sativa Japonica Group]  
 VCADWGVEAATFTFVNRCTDTVWPGVLSNAGSARLATTGFELPPGVARAVPAPAGWSGRMWARTGCAVVQ  
 DGGAGGGRMVCATGDCGSGGAECNGAGAAPPATLAEFTLDGSGGLDFYDVSLVDGYNLPVLVEPSSSGGG  
 GGGGSLTSAATCAAAGCAADLNAMCPAELRAGGGAACRSACDAFGRPEFCCSGAFANPSTCRPTAYSQV  
 FKSACPRSYSYAFDDPTSTFTCSGGPDYTLTFCPASSPSGSQKSTTATPTPAAMMPGTGTPTTPTTATAM  
 PGATMPGTATATTMPGTTFTDAVPDTSMPMPMGDAGGGGEEGVVLSGSETWIANMATGELTAAASLSRP  
 SPAAALALFLVHALRLLVLR

>NP\_001049531.1 Os03g0243900 [Oryza sativa Japonica Group]  
 MMGIQRICIVLGMFLILVREGGAVTFTFVNRCTGTVWPGILSNAGSARMDPTGFELPPGAARAVPAPTGW  
 SGRLWARTGCTQDGTGKVCATGDCGSGTLECAGGAAPPATLAEFTLDGGGRNDFYDVSLVDGYNLPLL  
 VEPSGALGATATCAAAGCAADLNARCPAELRAVGGAAACRSACDAFGKPEFCCSGAYANPNTCRPTAYSQ  
 VFKSACPRSYSYAYDDPTSTFTCAGGRDYTITFCPVATPSLKSAGGTTTTPTTTPGVTTDAPPDMARPIG  
 SQGGGGGGGAGQGVMLGDNSWLASLAMDASSRRASRLALLAAPLALLTLHLPRL

>2I0W\_1|Chain A|Protein NP24|Solanum lycopersicum (4081)  
 ATIEVRNCPYTVWAASTPIGGRRRLNRGQTWVINAPRGTKMARIWGRTGCNFNAAGRGTCQTGDCGGVLQCTGWGK  
 PPNTLAEYALDQFSNLDFWDISLVDGFNIPMTFAPTKPSGGKCHAIHCTANINGECPRALKVPGGCNNPCTTFGGQQ  
 YCCTQGPGCPTLSKFFKKRCPDAYSYPQDDPTSTFTCPGGSTNYRVVFCPNG