

WESLEI BRUNO BOTERO

Avaliação da especificidade e da capacidade produtiva da enzima friedelina sintase a partir de estudos de mutação sítio-dirigida

Tese de Doutorado apresentada no Instituto de Química, Universidade Estadual Paulista, como parte dos requisitos para obtenção do título de Doutor em Química.

Orientador: Prof^a. Dr^a. Maysa Furlan

Araraquara
2023

B748a Botero, Weslei Bruno
Avaliação da especificidade e da capacidade produtiva da enzima friedelina sintase a partir de estudos de mutação sítio-dirigida / Weslei Bruno Botero. -- Araraquara, 2023
108 f. : il.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Instituto de Química, Araraquara
Orientadora: Maysa Furlan

1. Maytenus. 2. Espinheira-santa. 3. Biossíntese. 4. Terpenos. 5. Genômica. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Química, Araraquara. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: "Avaliação da especificidade e da capacidade produtiva da enzima friedelina sintase a partir de estudos de mutação sítio-dirigida"

AUTOR: WESLEI BRUNO BOTERO

ORIENTADORA: MAYSA FURLAN

Aprovado como parte das exigências para obtenção do Título de Doutor em QUÍMICA, pela Comissão Examinadora:

Prof.^a Dr.^a MAYSA FURLAN (Participação Virtual)
Departamento de Bioquímica e Química Orgânica / Instituto de Química - UNESP - Araraquara

Prof. Dr. IAN CASTRO GAMBOA (Participação Virtual)
Departamento de Bioquímica e Química Orgânica / Instituto de Química - UNESP - Araraquara

Prof. Dr. MARCELO JOSÉ PENA FERREIRA (Participação Virtual)
Departamento de Botânica / Instituto de Biociências - USP - São Paulo

Prof. Dr. FERNANDO COTINGUIBA DA SILVA (Participação Virtual)
Instituto de Pesquisas de Produtos Naturais Walter Mors - UFRJ - Rio de Janeiro

Prof. Dr. LUIS OCTÁVIO REGASINI (Participação Virtual)
Departamento de Química e Ciências Ambientais / Instituto de Biociências, Letras e Ciências Exatas - UNESP - São José do Rio Preto

Araraquara, 14 de abril de 2023

Maysa

Furlan:060

03111801

Assinado de forma
digital por Maysa
Furlan:06003111801
Dados: 2023.04.18
10:15:00 -03'00'

DADOS CURRICULARES

IDENTIFICAÇÃO

Nome: Weslei Bruno Botero

Nome em citações bibliográficas: BOTERO, W. B.

Endereço profissional: Rua Prof. Francisco Degni, n° 55, Bairro Quitandinha, Araraquara/SP, CEP: 14800-060.

FORMAÇÃO ACADÊMICA/TITULAÇÃO

Doutorado em Química (2018-2023)

Área: Química Orgânica.

Agência financiadora: CNPq (Processo n° 141024/2018-6).

Orientadora: Prof. Dra. Maysa Furlan.

Título da tese: Avaliação da especificidade e da capacidade produtiva da enzima friedelina sintase a partir de estudos de mutação sítio-dirigida.

Mestrado em Química (2016-2018)

Área: Química Orgânica.

Agência financiadora: CNPq (Processo n° 131233/2016-5).

Orientadora: Profa. Dra. Lourdes Campaner dos Santos.

Título da dissertação: Bioprospecção e isolamento de metabólitos de *Microsphaeropsis arundinis*, um fungo endofítico de *Paepalanthus planifolius* (Eriocaulaceae).

Iniciação científica (2013-2015)

Área: Química Orgânica.

Agência financiadora: CNPq (Processo n° 137948/2015-8).

Orientadora: Profa. Dra. Lourdes Campaner dos Santos.

Título do trabalho de conclusão de curso: Estudo químico dos escapos de *Paepalanthus geniculatus* Kunth.

Bacharelado em Química (2012-2015)

Local: Instituto de Química de Araraquara – IQ UNESP

FORMAÇÃO COMPLEMENTAR

Estágio docência (2018)

Curso: Farmácia.

Local: Faculdade de Ciências Farmacêuticas do Câmpus de Araraquara.

Disciplina: Química Orgânica I.

Docentes responsáveis: Profa. Dra. Cíntia Duarte Freitas Milagre.

Profa. Dra. Lourdes Campaner dos Santos.

Estágio docência (2016)

Curso: Bacharelado em Química.

Local: Instituto de Química de Araraquara.

Disciplina: Química Orgânica II.

Docente responsável: Profa. Dra. Isabele Rodrigues Nascimento.

Organização de minicurso (2018)

Curso: 1ª Escola de Versão IQ-Araraquara.

Local: Instituto de Química de Araraquara – UNESP.

Minicurso: Plantas medicinais: Uma abordagem prática de preparação e estudo do perfil químico de extratos vegetais.

Cursos complementares

Curso: Métodos de Bioinformática I.

Local: Universidade de Toronto.

Realização: Virtual.

Curso: Métodos de Bioinformática II.

Local: Universidade de Toronto.

Realização: Virtual.

PRODUÇÃO BIBLIOGRÁFICA

Artigos completos publicados em periódicos

BOTERO, W. B.; AMORIM, M. R.; CARLOS, I. Z.; POLES, M. C.; SANTOS, L. C.
Aromatic Polyketides and Macrolides from *Microsphaeropsis arundinis*. **Journal of the Brazilian Chemical Society**, v. 31, n. 2, p. 364-369, 2020.

Trabalhos publicados em anais de eventos científicos

AMORIM, M. R.; BOTERO, W. B.; SILVA, A. C. Z.; ARAUJO, A. R.; CARLOS, I. Z.; SANTOS, L. C. Cytotoxic activities of endophytic fungi extracts from *Paepalanthus planifolius*: A Brazilian Evergreen. In: 9th Joint Natural Products Conference. **Planta Medica**, v. 82, n. 1, 2016.

DOURADO, C. P.; AMORIM, M. R.; BOTERO, W. B.; PEITOXO, T.; SANTOS, L. C. Avaliação do perfil químico do fungo endofítico *Paraphaeosphaeria sporulosa* isolado das folhas de *Paepalanthus planifolius* (Bong) Korn (Eriocaulaceae). In: VI Congresso Científico da UNESP. **Revista de Ciências Farmacêuticas Básica e Aplicada**, v. 37, n. 1, 2016.

Apresentações de trabalhos e palestras em eventos científicos

BOTERO, W. B.; SOUZA-MOREIRA, T. M.; REMLINGER, M.; ZANELLI, C. F.; VALENTINI, S. R.; FURLAN, M. Use of site-directed mutagenesis to determine key positions in friedelin synthase (FRS), a monofunctional enzyme from *Monteverdia ilicifolia*. **44ª Reunião Anual da Sociedade Brasileira de Química**, 2021.

BOTERO, W. B.; AMORIM, M. R.; POLES, M. C.; CARLOS, I. Z.; SANTOS, L. C. Cytotoxic activity and isolation of metabolites from *Microsphaeropsis arundinis*, an endophytic fungus from *Paepalanthus planifolius*. **46th World Chemistry Congress – IUPAC**, 2017.

BOTERO, W. B.; AMORIM, M. R.; SANTOS, L. C. Avaliação da atividade antirradicalar dos extratos e das substâncias isoladas nos escapos de *Paepalanthus geniculatus* (Eriocaulaceae). **XXVII Congresso de Iniciação Científica da UNESP**, 2015.

BOTERO, W. B.; SANTOS, L. C. Estudo fitoquímico dos escapos de *Paepalanthus geniculatus* Kunth. **XXVI Congresso de Iniciação Científica da UNESP**, 2014.

“Evito contar o processo extraordinário que empreguei na composição dessas memórias, trabalhadas cá no outro mundo. Seria curioso, mas nimiamente extenso, e aliás desnecessário ao entendimento da obra. A obra em si mesmo é tudo: se te agradar, fino leitor, pago-me da tarefa; se não te agradar, pago-te com um piparote, e adeus”

-Brás Cubas.

AGRADECIMENTOS

Primeiramente gostaria de agradecer a Deus por todas as realizações e oportunidades em minha vida.

Agradeço imensamente aos meus pais Edna e Eduardo por todo o apoio, confiança e amor durante todos os anos na academia. Ao meu irmão Wander, minha cunhada Andréa e meu pai Humberto, um agradecimento especial.

Agradeço a UNESP, instituição que me formou como pesquisador e ser humano, pelos mais de dez anos de aprendizado e de muitas histórias. Agradeço ao CNPq (Processo nº 141024/2018-6) pelo financiamento concedido, essencial para o desenvolvimento da pesquisa e aos demais órgãos financiadores CAPES e FAPESP.

A minha orientadora Profa. Dra. Maysa Furlan, um agradecimento pela orientação e carinho. Também deixo um agradecimento especial ao Prof. Dr. Sandro Roberto Valentini e ao Prof. Dr. Cleslei Zanelli, pelos ensinamentos valiosos adquiridos no laboratório de Biologia Molecular. Agradeço à colaboração dos membros da banca na correção deste trabalho.

Agradeço a menina Gabriela Oliveira por aparecer em minha vida em um momento ímpar e fazer par comigo, sempre acreditando em mim.

Deixo aqui um agradecimento especial ao Laboratório de Biologia Molecular e as pessoas ali presentes, que transformaram meu maior desafio profissional na melhor experiência de minha vida, em especial a Dra. Tatiana Maria e ao Ms. Otávio Aguiar, que me deram um suporte incrível no fim do trabalho. Obrigado Fernanda Manaia (Warrior), Suélen Fernandes (Susu), Angélica Klippel, Marco Aurélio (Marcão), Maria Olívia (Marô), Mariana Santoni, Ana Paiva, Anderson Pinheiro, Marcela Forgerini, Bianca Fernandes, Enzo Corvello, Ieda Gabriela e Anna Júlia.

Agradeço ao NuBBE pela receptividade, conhecimento e pelas amizades ali feitas, em especial à Maria do Carmo, Naira Anhesine, Bruna Mazzeu, Karina, Wellington, Aline e demais.

Aos demais amigos que estiveram sempre presentes, Gabriel Alvarenga, Danielle Nogueira, Renan Zanetti, Rodrigo Cesar, Luis Felipe Bricks Bim, Murilo Araújo, Léo Inácio, Andressa Nascimento, Letícia Borges e Ana Zanatta.

Por fim agradeço aos servidores do Instituto de Química, em especial ao João Bronzel e Juliana Rodrigues, e aos amigos da biblioteca.

RESUMO

Plantas constituem uma das principais fontes de produtos naturais bioativos, sendo capazes de sintetizar substâncias complexas e estereoespecíficas utilizadas para o tratamento de diversas doenças. Dentre essas substâncias, a friedelina se destaca pelo potencial gastroprotetor, antimicrobiano, hipolipemiante e por possuir uma carbonila na posição C-3 que faz com que a mesma seja precursora de substâncias com propriedades citotóxicas como a pristimerina, o celastrol e a maitenina, marcadores taxonômicos da família Celastraceae. Ademais o alto potencial terapêutico, a quantidade produzida desses metabólitos pelas plantas é pequena, de modo que alternativas de engenharia metabólica, com a expressão dos genes de plantas em sistemas heterólogos aliada à engenharia de proteínas para otimizar e aumentar a atividade das enzimas, vêm se mostrando promissoras. Desse modo, esse estudo avaliou a expressão da enzima friedelina sintase de *Monteverdia ilicifolia* (Celastraceae) e *Populus davidiana* (Salicaceae) em *Saccharomyces cerevisiae*. Resíduos específicos de ambas as enzimas foram modificados geneticamente por meio de reações de mutação sítio-dirigida, sendo as posições selecionadas a partir do uso do alinhamento múltiplo de sequências (MAS) e a análise da interação entre o sítio catalítico da enzima modelada e o produto friedelina. Após a obtenção dos genes mutantes, a sequência codificadora das enzimas foi clonada em um vetor de expressão, sendo o DNA recombinante transformado em levedura. Posteriormente, essa linhagem foi cultivada em meio seletivo e os terpenos extraídos e avaliados por cromatografia gasosa acoplada à espectrometria de massas (GC-MS). O teor de friedelina produzido pelas leveduras foi quantificado através de uma curva analítica utilizando colesterol como padrão interno. O teor de friedelina obtido para as enzimas selvagens de *M. ilicifolia* e *P. davidiana* foram de $2,024 \pm 0,169$ mg/L e $1,024 \pm 0,072$ mg/L, respectivamente. Dentre os mutantes avaliados, G533A apresentou um aumento na produção de friedelina de 1,8-3,6x ($3,633 \pm 0,203$ mg/L), enquanto as alterações na posição 729 resultaram em perda da função. As demais mutações apresentaram resultados mistos de acordo com a natureza da enzima mutada. Os resultados aqui apresentados são importantes para melhorar o entendimento sobre a estabilidade e capacidade produtiva dessas enzimas que produzem a friedelina, o único triterpeno carbonilado da cascata de rearranjos da via biossintética.

Palavras-chave: Celastraceae, *Monteverdia ilicifolia*; *Populus davidiana*; *Saccharomyces cerevisiae*; friedelina; engenharia metabólica; mutação sítio-dirigida

ABSTRACT

Plants are one of the main sources of bioactive natural products, being capable of synthesizing complex and stereospecific substances used for the treatment of diseases. Among these substances, friedelin is a triterpene with a lot of biological activities described, including gastroprotective, antimicrobial and hypolipidemic potential. This compound is known for having a carbonyl group at C-3 position, which makes it a precursor of cytotoxic quinonemethide triterpenes such as pristimerin, celastrol and maytenin, chemical markers of the Celastraceae family. Despite this potential, the amount of friedelin produced by host plants are low. In that case, the use of metabolic engineering alternatives with the expression of plant genes in heterologous systems combined with protein engineering can be a promising way to increase the production of this compound, while site-directed mutagenesis studies can determine the influence of key residues in the activity and specificity of the enzyme friedelin synthase. This study evaluated the expression of the enzyme friedelin synthase from *Monteverdia ilicifolia* (MiFRS) and *Populus davidiana* (PdFRS) in *Saccharomyces cerevisiae*. Specific residues of both enzymes were genetically modified through site-directed mutagenesis reactions, with key-positions for mutagenesis being selected using multiple sequence alignment (MAS) and molecular modelling. After these PCR reactions, the coding sequences were cloned into an expression vector and the recombinant DNA was transformed into yeast. Subsequently, the strains were cultivated in a selective medium and the terpenes were extracted and evaluated by gas chromatography-mass spectrometry (GC-MS). The mutations that produce friedelin were quantified through a calibration curve using cholesterol as standard. The friedelin content obtained for the wild-type enzymes were 2.024 ± 0.169 mg/L for MiFRS and 1.024 ± 0.072 mg/L for PdFRS. Among the mutants, G533A achieved a production of 3.633 ± 0.203 mg/L, an increase from 1.8-3.6 fold in comparison with the wild type enzymes. Mutations at the position 729 led to a loss of the enzyme function, while other residues showed mixed results. The results presented here are important to the understanding of the stability, specificity and productive capacity of friedelin synthase, the enzyme that produces the only carbonyl pentacyclic triterpene in the cascade of rearrangements of the biosynthetic pathway.

Keywords: Celastraceae, *Monteverdia ilicifolia*; *Populus davidiana*; *Saccharomyces cerevisiae*; friedelin; metabolic engineering; site-directed mutagenesis.

LISTA DE FIGURAS

Figura 1 – Folhas de <i>Monteverdia ilicifolia</i>	15
Figura 2 – Principais metabólitos bioativos encontrados no extrato de <i>Monteverdia ilicifolia</i>	16
Figura 3 – Produção do esqualeno a partir dos blocos construtores de IPP e DMAPP	18
Figura 4 – Biossíntese de triterpenos a partir do 2,3-oxidoesqualeno em eucariotos	19
Figura 5 – Estrutura química da friedelina.....	20
Figura 6 – Exemplos de metabólitos produzidos a partir de sistemas heterólogos..	22
Figura 7 – Representação da reação BP para a produção do clone de entrada	33
Figura 8 – Representação da reação LR para a produção do clone de expressão .	34
Figura 9 – Triterpenos produzidos a partir das enzimas utilizadas no alinhamento múltiplo.....	40
Figura 10 – Modelo construído para a enzima <i>Mi</i> FRS (roxo) utilizando como molde a enzima lanosterol sintase de humano (amarelo)	47
Figura 11 – Gráfico de Ramachandran para a enzima <i>Mi</i> FRS modelada na lanosterol sintase	48
Figura 12 – Modelo construído para a enzima <i>Mi</i> FRS (azul) utilizando como molde a enzima lanosterol sintase de humano (amarelo)	49
Figura 13 – Gráfico de Ramachandran para a enzima <i>Pd</i> FRS modelada na lanosterol sintase	50
Figura 14 – Árvore filogenética produzida a partir das enzimas alinhadas	51
Figura 15 – Visão geral do alinhamento múltiplo e destaque para regiões importantes.	53
Figura 16 - Representação do MSA com destaque para a posição 416.....	55
Figura 17 – Interação entre os resíduos L416 e F473	56
Figura 18 – Aminoácidos utilizados nos estudos de mutação sítio-dirigida da posição 416	57
Figura 19 – Representação do MSA com destaque para a posição 533.....	58
Figura 20 – Interação entre os resíduos G533, L482, W534 e a posição C-3	59
Figura 21 – Aminoácidos utilizados nos estudos de mutação sítio-dirigida para a posição 533.....	60
Figura 22 – Representação do MSA com destaque para a posição 729.....	61
Figura 23 – Interação entre os resíduos K729 e Y609.....	62
Figura 24 – Aminoácidos utilizados nos estudos de mutação sítio-dirigida da posição 729	63
Figura 25 – Representação do MSA com destaque para a posição 609.....	65
Figura 26 – Interação entre os resíduos Y609, W608 e K729	66
Figura 27 – Aminoácidos utilizados nos estudos de mutação sítio-dirigida para a posição 609.....	66

Figura 28 – PCR de colônia do DNA selvagem e mutante após a transformação em levedura	68
Figura 29 – Expressão da enzima <i>MiFRS</i> e dos mutantes da posição 729. A friedelina foi identificada a partir da biblioteca NIST14	69
Figura 30 – Expressão da enzima <i>PdFRS</i> e dos mutantes da posição 729	70
Figura 31 – Expressão da enzima <i>MiFRS</i> e os mutantes <i>Mi-L416Q</i> e <i>Mi-L416H</i> por GC-MS	71
Figura 32 – Expressão da enzima <i>PdFRS</i> e os mutantes <i>Pd-L414Q</i> e <i>Pd-L414H</i> por GC-MS	72
Figura 33 – Expressão da enzima <i>MiFRS</i> e o mutante <i>Mi-G533A</i>	73
Figura 34 – Expressão da enzima <i>PdFRS</i> e o mutante <i>Pd-G530A</i>	74
Figura 35 – Expressão da enzima <i>MiFRS</i> e o mutante <i>Mi-Y609A</i>	75
Figura 36 – Expressão da enzima <i>PdFRS</i> e o mutante <i>Pd-Y606A</i>	75
Figura 37 – Curva analítica obtida para o padrão de friedelina.....	76
Figura 38 – Comparação da concentração de friedelina em mg/L para os pares de genes selvagens e mutantes	78
Figura 39 – Comparação da concentração de friedelina em mg/L para o gene selvagem e os mutantes de <i>MiFRS</i>	79
Figura 40 – Comparação da concentração de friedelina em mg/L para o gene selvagem e os mutantes de <i>PdFRS</i>	80

LISTA DE TABELAS

Tabela 1 – Primers utilizados para a inserção da sequência <i>attB</i> nos genes de interesse.....	32
Tabela 2 – Componentes utilizados na reação de PCR para inserção da sequência <i>attB</i> nos genes de interesse.....	32
Tabela 3 – Condições de PCR utilizadas para a amplificação dos genes.....	32
Tabela 4 – Mistura reacional da reação BP.....	34
Tabela 5 – Mistura reacional da reação LR.....	35
Tabela 6 – Mistura reacional da reação de sequenciamento.....	36
Tabela 7 – Condições de PCR utilizadas para o sequenciamento.....	36
Tabela 8 – Mix utilizado para a reação de transformação em levedura.....	37
Tabela 9 – Mix utilizado para a reação de transformação em levedura.....	38
Tabela 10 – Enzimas utilizadas no alinhamento múltiplo de sequências (MSA).....	41
Tabela 11 – Mix utilizado para as reações de mutação sítio-dirigida.....	43
Tabela 12 – Programa de temperatura utilizado para a reação de mutação sítio-dirigida.....	43
Tabela 13 – Programa de temperatura utilizado para a o anelamento da reação de mutação.....	43
Tabela 14 – Mutações propostas para a posição 416.....	57
Tabela 15 – Alterações realizadas para a obtenção dos mutantes da posição 416.....	57
Tabela 16 – Mutações propostas para a posição 533.....	60
Tabela 17 – Alterações realizadas para a obtenção dos mutantes da posição 533.....	60
Tabela 18 – Mutações propostas para a posição 729.....	63
Tabela 19 – Primers utilizados para a realização da PCR de mutação sítio-dirigida na posição 729.....	64
Tabela 20 – Mutações propostas para a posição 609.....	66
Tabela 21 – Primers utilizados para a realização da PCR de mutação sítio-dirigida na posição 609.....	67
Tabela 22 – Concentração de friedelina em mg/L de meio de cultura obtida para as amostras após interpolação dos dados na curva analítica. n = 4.....	77

SUMÁRIO

1 INTRODUÇÃO	14
1.1 A família Celastraceae e <i>Monteverdia ilicifolia</i>	14
1.2 A biossíntese dos triterpenos	17
1.3 Biologia molecular e engenharia metabólica aplicada à Química dos Produtos Naturais	20
1.4 Mutação sítio-dirigida	23
2 OBJETIVOS	26
3 MATERIAIS E MÉTODOS	27
3.1 Equipamentos.....	27
3.2 Soluções Preparadas	27
3.3 Enzimas e reagentes	28
3.4 Genes, Plasmídeos e linhagens de microrganismos.	29
3.5 Meios de cultura	29
4 METODOLOGIA	30
4.1 Subclonagem das sequências codificadoras dos genes selvagens de friedelina sintase	30
4.1.1 Seleção dos genes de friedelina sintase	30
4.1.2. Purificação dos genes (Miniprep)	30
4.1.3 Clonagem por Gateway®	31
4.1.4 Sequenciamento dos genes clonados	35
4.1.5 Transformação em <i>S. cerevisiae</i>	37
4.1.6 PCR de colônia	38
4.2 Realização de mutações sítio-dirigidas e quantificação dos produtos	38
4.2.1 Alinhamento dos genes de OSCs através de técnicas de bioinformática.....	39
4.2.2 Realização das reações de mutação sítio-dirigida.....	42
4.2.3 Expressão em sistema heterólogo.....	44
4.2.4 Extração da friedelina e quantificação por GC-MS	44
5 RESULTADOS E DISCUSSÕES.....	46
5.1 Estudos de modelagem molecular	46
5.1.1 Estudos de modelagem para a enzima MiFRS.....	46
5.1.2 Estudos de modelagem para a enzima PdFRS	49
5.2 Alinhamento das sequências e seleção de mutantes	50
5.2.1 Análise evolutiva das enzimas alinhadas.....	50
5.2.2 Alinhamento múltiplo de sequências.....	52
5.3 Expressão dos mutantes em sistema heterólogo e quantificação da friedelina	67
5.3.1. Mutações que não produziram friedelina	68
5.3.2. Mutações que produziram friedelina	71
5.4 Quantificação dos mutantes ativos.....	76
5.4.1 Construção da curva analítica.....	76
5.4.2 Quantificação da friedelina produzida pelas leveduras.....	77
6 CONCLUSÕES	81
7 REFERÊNCIAS.....	82
ANEXOS	88

1 INTRODUÇÃO

1.1 A família Celastraceae e *Monteverdia ilicifolia*

A família Celastraceae é composta atualmente por mais de 1300 plantas classificadas em 106 diferentes gêneros. As espécies pertencentes a essa família estão distribuídas globalmente em regiões tropicais e subtropicais, com destaque para o norte do continente africano, o leste do continente asiático e o sul do continente americano, onde o Brasil se destaca com 49 táxons no Brasil e 36 espécies endêmicas (BIRAL et al., 2017; SANTOS-OLIVEIRA; COULAUD-CUNHA; COLAÇO, 2009; VELOSO et al., 2017). Os principais gênero dentro da família são *Maytenus*, *Monteverdia*, *Euonymus*, *Cassine*, *Tripterygium* e *Celastrus* (PERESTELO, 2009).

O histórico etnofarmacológico da família, e principalmente do gênero *Maytenus*, é extenso. Como exemplo, infusões das folhas de *Maytenus ilicifolia* e *Maytenus aquifolium* são amplamente utilizadas para o tratamento de distúrbios gastrointestinais (ANTUNES et al., 2020). Além disso, *Euonymus alatus* têm importância ornamental e infusões com seus galhos são ingeridas para o controle de diabetes e hiperglicemia (CHANG et al., 2012), sementes de *Celastrus paniculatus* são consumidas principalmente na Índia para o tratamento de disfunção cognitiva e epilepsia (SHEN et al., 2019) e espécies como *Tripterygium wilfordii* são utilizadas pela medicina tradicional chinesa para o tratamento de artrite reumatoide e inflamações (BRINKER et al., 2007)

Estudos recentes publicados por Biral e colaboradores (2017) reclassificaram espécies dentro do gênero *Maytenus* por similaridades morfológicas e análises genéticas, resultando na criação do gênero *Monteverdia* com 123 espécies. Dentre as espécies desse gênero, *Monteverdia ilicifolia* (Mart. ex Reissek) Biral, anteriormente conhecida como *Maytenus ilicifolia*, destaca-se pela ampla gama de atividades biológicas e rica diversidade metabólica (Figura 1).



Figura 1 – Folhas de *Monteverdia ilicifolia*. Fonte: Portal CIBFar.

Conhecida popularmente como “espinheira-santa”, “cancerosa” e “espinho-de-Deus”, *Monteverdia ilicifolia* é uma espécie amplamente utilizada pela população para o tratamento de dispepsias e úlceras gástricas, com os primeiros registros da utilização datando de 1920 (ANTUNES et al., 2020). A espécie, nativa do sul do continente, possui suas folhas em formato de espinho e galhos vendidos no país para serem consumidos em infusões, tinturas e tererês. O extrato seco é atualmente comercializado em cápsulas produzidas por laboratórios como Bionatus® e Herbarium®, sendo esses fitoterápicos recomendados pela ANVISA. Além disso, a espécie está presente na Relação Nacional de Plantas Medicinais de Interesse ao Sistema Único de Saúde – REINISUS (CHEN et al., 2023).

O potencial biológico dessa planta vai além das capacidades gastroprotetoras. Estudos na literatura mostram a presença de substâncias com potencial antioxidante, antimicrobiano, citotóxico, antiinflamatório, antineoplásico, hipotensivo e antidiabético. Os metabólitos que conferem essas atividades para o extrato bruto fazem parte das classes dos flavonoides, taninos condensados, alcaloides e triterpenos (Figura 2) (ANTONISAMY; DURAIPANDIYAN; IGNACIMUTHU, 2011; ANTUNES et al., 2020; CHEN et al., 2023; VELOSO et al.,

2017).

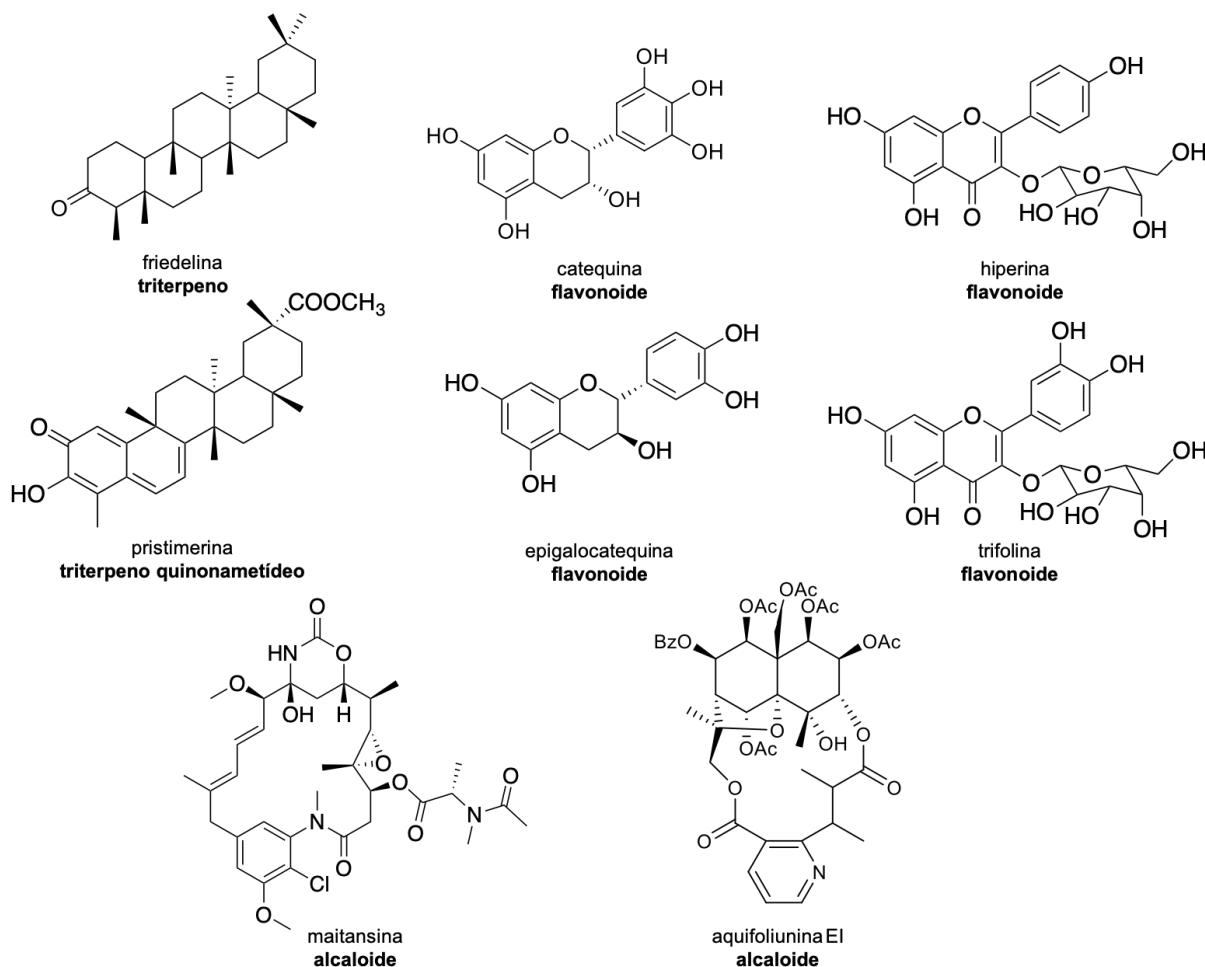


Figura 2 – Principais metabólitos bioativos encontrados no extrato de *Monteverdia ilicifolia* Fonte: Imagem fornecida pelo autor.

É reportado que os flavonoides auxiliam na atividade gástrica a partir da captura de radicais livres, protegendo o estômago e o intestino da peroxidação lipídica, e por inativar a enzima ATPase da bomba de hidrogênio e potássio gástrico, responsável por acidificar o estômago (NIERO; ANDRADE; FILHO, 2011). Baggio e colaboradores (2007) verificaram que uma fração do extrato bruto rica em flavonoides inibiu o esvaziamento gástrico e a motilidade intestinal em níveis próximos à atropina.

Taninos condensados possuem a capacidade de formar complexos com proteínas, polissacarídeos e íons metálicos, o que resulta na produção de uma camada protetora sobre a mucosa do intestino, favorecendo a recuperação do mesmo após lesões. A epigallocatequina e epigallocatequina-3-galato demonstraram atividade *in vitro* contra células tumorais gástricas visto que estimulam a liberação do

fator de crescimento transformador beta (TFG- β), que controla a proliferação dessas células de adenocarcinoma (OKABE et al., 1999; ZHENG, 1994).

Os alcaloides não tem relação direta com a atividade gastroprotetora, porém os metabólitos isolados como aquifoliunina E1 apresentam atividade antiprotozoária contra *Trypanosoma cruzi* e *Leishmania chagasi* e atividade anti-HCV (SANTOS et al., 2012).

Outra classe de compostos que merece destaque pelas atividades biológicas apresentadas são os terpenos, em especial os triterpenos pentacíclicos friedelânicos e quinonametídeos presentes nas folhas e nas raízes de *M. ilicifolia*, respectivamente. Essas substâncias auxiliam na proteção da mucosa gástrica e na inibição da secreção ácida a partir do estímulo à produção de prostaglandina (ALVES et al., 2018).

A friedelina, um derivado 3-oxofriedelano apresenta atividade antiproliferativa e potencial antimicrobiano, anti-pirético e hipolipemiante já descritos na literatura (SUBASH-BABU; LI; ALSHATWI, 2017). Acredita-se que este triterpeno é sintetizado nas folhas e posteriormente transportado para as raízes onde é convertido em triterpenos quinonametídeos (QMTs) como celastrol, maitenina e a pristimerina, compostos com expressiva antiprotozoária e antitumoral (SANTOS et al., 2013; SEO et al., 2018). Estudos recentes com a pristimerina demonstraram que a mesma inibiu o crescimento de células tumorais humanas de pulmão (A-549), mama (MCF-7) e fígado (Hep-G2 e Hep-3B) (CHEN et al., 2023).

1.2 A biossíntese dos triterpenos

Os terpenos representam uma das mais importantes e amplamente encontradas classes de compostos existentes na natureza, com mais de 35 mil substâncias isoladas e identificadas de fontes naturais como plantas, organismos marinhos e demais microrganismos (HAPPI et al., 2022).

A diversidade química dos terpenos vem de sua biossíntese que ocorre a partir da condensação de unidades de isopreno (C_5H_8) que funcionam como blocos construtores para a produção de monoterpenos (C_{10}), sesquiterpenos (C_{15}), diterpenos (C_{20}), sesterterpenos (C_{25}), triterpenos (C_{30}), sesquariterpenos (C_{35}) e carotenoides (C_{40}).

Duas vias biossintéticas explicam a produção desses compostos, a via do mevalonato (MEV) e a via do metileritritol-fosfato (MEP). A diferença entre elas é o precursor, de modo que as unidades construtoras da via do mevalonato são oriundas da condensação de unidades de acetil-CoA por reação de Claisen, resultando no ácido mevalônico. Na via do metileritritol-fosfato, o ácido pirúvico se condensa com tiamina pirofosfato para a produção do 2-C-metil-D-eritritol-4-fosfato (THIMMAPPA et al., 2014).

A via do mevalonato ocorre fora do plastídeo e resulta nos blocos construtores pirofosfato de isopentenila (IPP) e pirofosfato de dimetilalila (DMAPP), sendo esses utilizados pelo metabolismo de plantas, animais, fungos e algas principalmente na produção de esteroides, sesquiterpenos e triterpenos. A via do metileritritol-fosfato engloba ciclização que ocorrem dentro dos plastídeos resultando na produção de monoterpenos, diterpenos, carotenoides e isoprenoides fotossintéticos.

Para a produção dos triterpenos, os terpenos de maior destaque em *M. ilicifolia*, três unidades dos blocos construtores IPP e DMAPP se condensam para a produção de farnesilpirofosfato (FPP). A junção de duas unidades de FPP cauda-cauda leva à produção do esqualeno, como demonstrado abaixo (Figura 3).

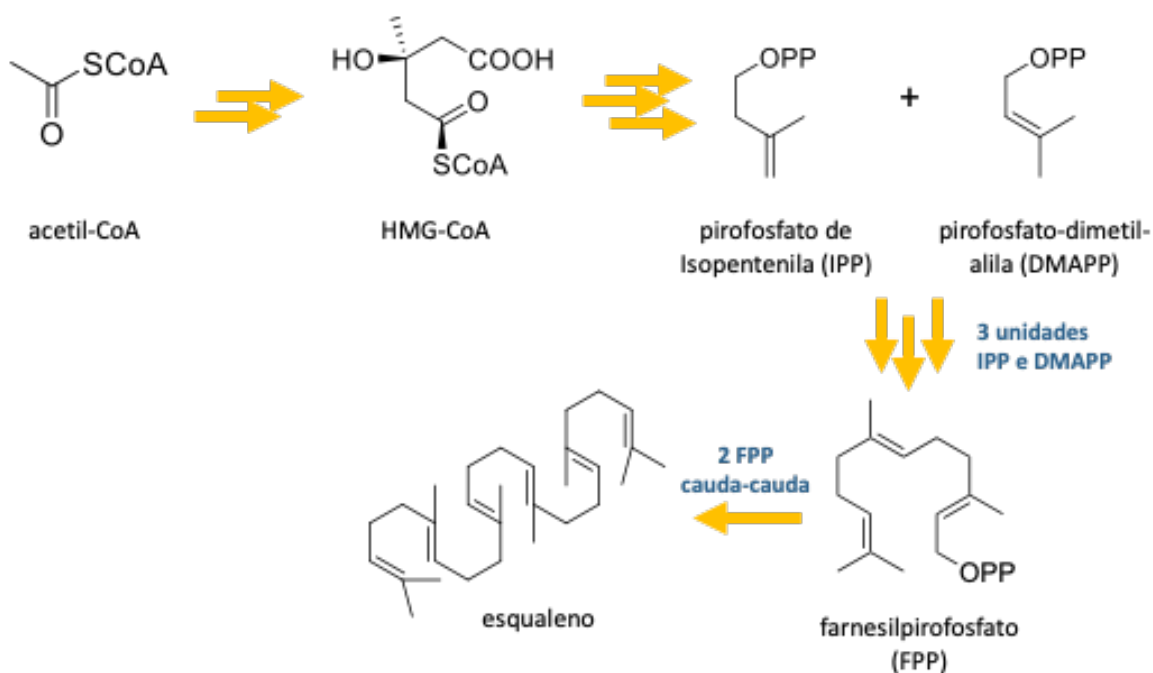


Figura 3 – Produção do esqualeno a partir dos blocos construtores de IPP e DMAPP. Fonte: Imagem adaptada de THIMMAPPA et al., 2014.

Nos eucariotos, o esqualeno é oxidado para 2,3-oxidoesqualeno e pode assumir a conformação cadeira-barco-cadeira, levando à produção de esteroides em fungos e animais, e a conformação cadeira-cadeira-cadeira, a qual, geralmente em plantas, leva à produção de triterpenos. A quebra do anel epóxi resulta na formação do cátion damarenila em plantas. Esse carbocátion assume diversas posições ao longo da estrutura do esqualeno, rearranjando-se através de migração de átomos de hidrogênio ou grupos metila.

A partir disso, uma série de rearranjos ocorre dando origem a diferentes esqueletos triterpênicos, sendo essas policiclizações catalisadas por enzimas conhecidas como oxidoesqualeno ciclases (OSCs).

A estabilização dos cátions leva à produção de triterpenos como lupeol, β -amirina e glutinol. O último rearranjo da cadeia leva à formação do cátion intermediário em C-4 friedelanila, que é posteriormente reduzido para friedelina (Figura 4) (THIMMAPPA et al., 2014). Dentre as enzimas, destaca-se a friedelina sintase, que recebe o substrato 2,3-oxidoesqualeno, promove o rearranjo e a estabilização do cátion friedelanila e realiza a redução desse cátion para o triterpeno friedelina.

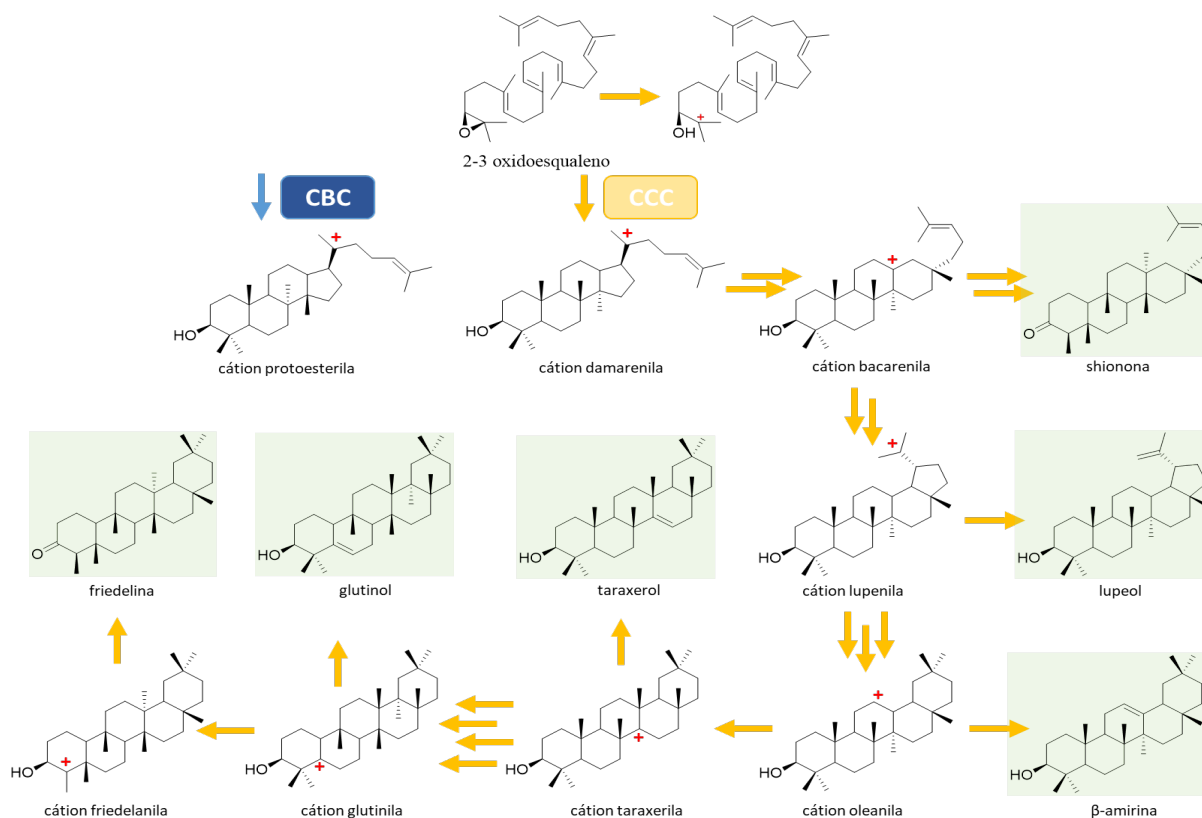


Figura 4 – Biossíntese de triterpenos a partir do 2,3-oxidoesqualeno em eucariotos. Fonte: Imagem adaptada de THIMMAPPA et al., 2014.

Além da ampla gama de atividades descritas, a friedelina é o único triterpeno pentacíclico da cascata de rearranjos da via biossintética que possui uma carbonila em C-3 (Figura 5). Essa especificidade estrutural é um dos motivos pelos quais essa substância e a enzima correlata, chamada de friedelina sintase, são os alvos de estudos do nosso grupo de pesquisa nos últimos anos (CORSINO et al., 2000; MAZZEU et al., 2021).

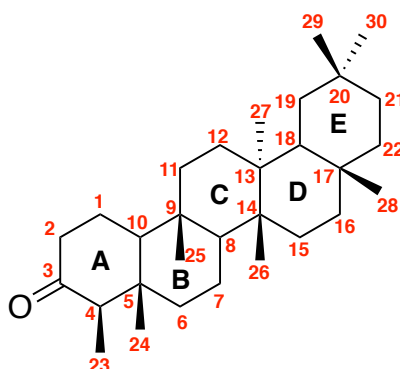


Figura 5 – Estrutura química da friedelina. Imagem fornecida pelo autor.

A fim de entender a especificidade e a capacidade produtiva, atualmente realiza-se a associação da química de produtos naturais com a engenharia metabólica e a engenharia de proteínas para compreender a biossíntese e aprimorar a produção desse triterpeno.

1.3 Biologia molecular e engenharia metabólica aplicada à Química dos Produtos Naturais

Plantas, fungos e bactérias são as principais fontes de produtos naturais bioativos. Esses organismos possuem um especializado metabolismo com a capacidade de sintetizar moléculas complexas e estereoespecíficas, que são posteriormente utilizadas pelo homem para o tratamento de variados tipos de doenças (PICKENS; TANG; CHOOI, 2011).

Ademais o alto potencial biológico dessas substâncias, a quantidade produzida desses metabólitos secundários, principalmente por plantas, é pequena. Desse modo, alternativas como a síntese química e a engenharia metabólica vêm sendo exploradas ao longo dos anos para a obtenção desses compostos de alto valor agregado em quantidades suficientes para suprir a demanda crescente por insumos farmacêuticos ativos (COURDAVAULT et al., 2021).

A obtenção de substâncias através de síntese e de semissíntese vêm permitindo, ao longo dos anos, a produção de variados tipos de metabólitos em maior escala. Como exemplo, podemos citar a síntese completa da artemisinina em cinco etapas (OLIVEIRA; CRUZ; DIAS, 2022; ZHU; COOK, 2012) e a obtenção do paclitaxel a partir do intermediário 10-desacetilbacatina-III em quatro etapas (DENIS et al., 1988), respectivamente.

O maior entrave para a obtenção de maior rendimento nessas sínteses é a estereosseletividade da maioria das reações específicas. Por isso, nas últimas décadas, dedicou-se uma atenção especial e crescente ao estudo das vias biossintéticas de vegetais e microrganismos, buscando conhecer as enzimas e os mecanismos que levam à produção dos compostos de interesse.

Recentemente, o avanço tecnológico e científico possibilitou o desenvolvimento de técnicas de sequenciamento de nova geração, estudos de genômica e de transcriptoma, os quais tem permitido o amplo mapeamento das vias biossintéticas, fornecendo alternativas para a manipulação do fluxo metabólico (YANG et al., 2022).

Desse modo, a engenharia metabólica, definida como a reprodução/modificação das vias biossintéticas de organismos superiores para a otimização da produção de um composto de interesse em um sistema heterólogo, vem se estabelecendo como uma alternativa sustentável, otimizada e promissora para a obtenção de metabólitos derivados de plantas (MOSES et al., 2017).

Nesse contexto, a utilização de um sistema heterólogo como uma bactéria ou levedura, permite a transferência dos genes de organismos derivados para um sistema de crescimento rápido e de fácil manipulação, gerando fábricas celulares para a produção desses compostos (BRÖKER et al., 2020). Ao mesmo tempo, técnicas de biologia molecular e biologia sintética permitem a manipulação e otimização dessas vias biossintéticas buscando uma alteração no fluxo metabólico e, conseqüentemente, uma maior produção (BIRCHFIELD; MCINTOSH, 2020).

Dentre os sistemas heterólogos mais utilizados, destacam-se bactérias (*Escherichia coli*), leveduras (*Saccharomyces cerevisiae*) e plantas (*Nicotiana benthamiana*). Atualmente, diferentes substâncias de alto valor agregado são produzidas a partir de engenharia metabólica, como β -amirina, ácido artemisínico e

taxadieno (Figura 6) (AJIKUMAR et al., 2010; CAO et al., 2015; MATSUMURA et al., 2018; PADDON; KEASLING, 2014; WESTFALL et al., 2012; ZHU et al., 2018).

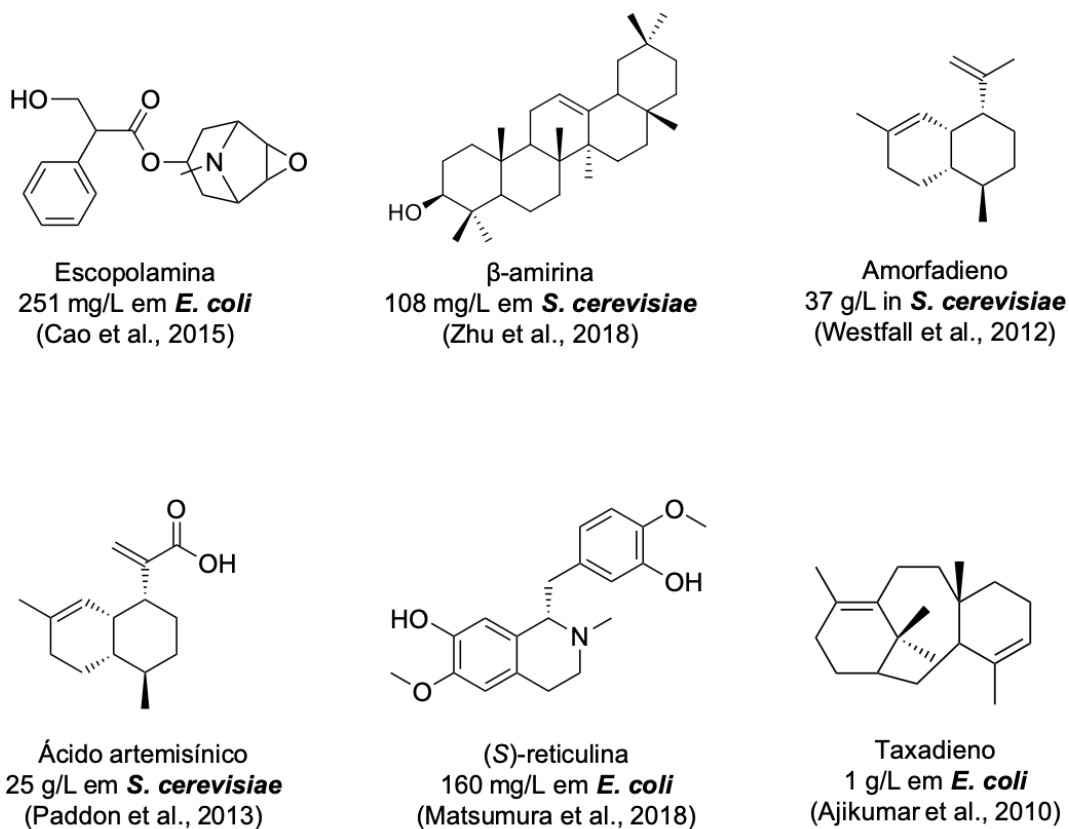


Figura 6 – Exemplos de metabólitos produzidos a partir de sistemas heterólogos. Fonte: Imagem adaptada de PYNE et al., 2019.

Para a produção de triterpenos, *Saccharomyces cerevisiae* se destaca por ser um organismo eucarioto, com modificações pós-traducionais, compartimentalizado e robusto, sendo capaz de reproduzir a via do mevalonato e expressar genes endógenos como as enzimas da família CYP450 (PICKENS; TANG; CHOOI, 2011).

Após a seleção do hospedeiro, é necessário clonar os genes que codificam as enzimas para a produção do metabólito desejado nesse sistema heterólogo. Além disso, outras modificações podem ser realizadas para fazer o ajuste fino do fluxo metabólico. Como exemplo, genes que codificam precursores da via podem ser superexpressos ou integrados ao cromossomo do sistema heterólogo, enquanto genes de vias competitivas podem ser nocauteados (LI et al., 2022).

Além disso, é possível também avaliar as relações estrutura-atividade das enzimas a partir da realização de mutações sítio-dirigidas.

1.4 Mutação sítio-dirigida

A mutação sítio-dirigida é uma técnica baseada na reação em cadeia da polimerase (PCR) utilizada para alterar nucleotídeos de uma sequência genética que é posteriormente recombinada (EDELHEIT; HANUKOGLU; HANUKOGLU, 2009). Com essas alterações, é possível avaliar a importância dos resíduos de aminoácidos para a atividade e especificidade de uma enzima.

Essas mutações fornecem informações valiosas na análise de famílias de genes com alto índice de similaridade entre si, o que ocorre com as enzimas oxidoesqualeno ciclases (OSCs). Na literatura existem mais de 98 mutações pontuais realizadas com esses genes, sendo que a partir desses estudos foi possível identificar regiões conservadas, resíduos importantes para a estabilidade e posições-chave para a atividade das enzimas (CHEN et al., 2021).

Dentre as regiões conservadas, o *motif* altamente conservado composto por aspartato-cisteína-treonina-alanina-ácido glutâmico (DCTAE) mostrou-se parte do sítio ativo da enzima, com o aspartato sendo responsável por protonar o 2,3-oxidoesqualeno e iniciar a cascata de rearranjos. Adicionalmente, pode-se citar as regiões conservadas metionina-X-cisteína-X-cisteína-arginina (MXCXCR), glutamina-X-X-X-X-X-triptofano (QXXXXXW) e C-terminal, cujo os resíduos são responsáveis pela estabilização do carbocátion e importantes para a ligação do substrato e especificidade da reação (CHEN et al., 2021).

Como exemplo de informações obtidas a partir da realização de mutações em oxidoesqualeno ciclases, é digno de nota o estudo realizado por Kushiro e colaboradores (1999), que realizou mutações na posição 256 da enzima lupeol sintase de *Olea europaea*. Como resultado da troca do resíduo de lisina por um resíduo de triptofano (L256W), a enzima converteu-se em uma β -amirina sintase, estabilizando o cátion oleanil e possibilitando uma etapa de rearranjo adicional que levou à produção majoritária de β -amirina.

A alteração do resíduo de serina pelo aminoácido fenilalanina na posição 728 da enzima β -amirina sintase de *Avena strigosa* (S728F) resultou na produção do triterpeno tetracíclico dammarano, não sendo a enzima mutada capaz de estabilizar cátions que levam à produção de triterpenos pentacíclicos (SALMON et al., 2016).

Além de mudar a função, alterações de determinados resíduos podem resultar em uma maior produção para a enzima. Zhou e colaboradores (2019) verificaram que a mutação da treonina para ácido glutâmico (T502E) na friedelina sintase de *Tripterygium wilfordii* foi capaz de aumentar a produção de friedelina sem alterar a especificidade da mesma. Um resultado semelhante foi observado por Mazzeu e colaboradores (2021), que identificaram que a alteração de um resíduo de metionina por um resíduo de serina na friedelina sintase de *Monteverdia ilicifolia* (M549S) resultou em um aumento de produção do composto de interesse.

Souza-Moreira e colaboradores (2016) realizaram os primeiros estudos de mutação sítio-dirigida com a enzima friedelina sintase de *Monteverdia ilicifolia* (MiFRS), observando que a posição 482, adjacente ao *motif* DCTAE, era essencial no reconhecimento do substrato pela enzima a partir da presença de um resíduo de leucina que não era conservado nas demais OSCs. Em 2018, Alves e colaboradores identificaram e quantificaram, a partir de expressão heteróloga, a produção de diferentes isoformas de friedelina sintase encontradas *M. ilicifolia*, identificando qual possuía a maior atividade enzimática.

Devido à atividade biológica da friedelina e dos triterpenos quinonametídeos, a enzima MiFRS vem sendo alvo de estudo de nosso grupo de pesquisa nos últimos 20 anos, inclusive com a finalidade de aumento de sua produção por expressão heteróloga. A partir do estudo publicado por Corsino (2000), foi possível verificar que a friedelina é produzida nas folhas e posteriormente transportada para as raízes, onde é oxidada para os triterpenos quinonametídeos (TQs).

Bicalho e colaboradores (2019) realizaram a identificação e a expressão heteróloga da enzima CYP712K4 em *Saccharomyces cerevisiae* e *Nicotiana benthamiana*, revelando que essa enzima é responsável por realizar a oxidação da posição C-29 da friedelina, levando à produção do ácido maitenoico.

Desse modo, o presente trabalho tem como objetivo continuar explorando resíduos-chave da enzima friedelina sintase a partir da realização de mutações sítio-dirigidas uma vez que a obtenção de uma enzima com maior capacidade produtiva pode colaborar na determinação da via biossintética dos TQMs. Além do aumento da produção, estudos da relação estrutura-atividade da FRS podem proporcionar melhor entendimento da especificidade estrutural desse triterpeno cetônico. No presente estudo, as mutações sítio-dirigidas foram realizadas nas enzimas

monofuncionais de *Monteverdia ilicifolia* (MiFRS) e de *Populus davidiana* (PdFRS) buscando identificar alterações na especificidade e capacidade produtiva da enzima, permitindo a obtenção de uma maior quantidade de friedelina.

2 OBJETIVOS

Objetivos gerais:

Entender a especificidade biossintética da friedelina sintase a partir de estudos de mutação sítio-dirigida e avaliar como essas mutações afetam a capacidade de produção da enzima.

Objetivos específicos:

- a) Clonar sequências codificadoras (cds) de friedelina sintase de *Maytenus ilicifolia* e de *Populus davidiana* no plasmídeo de expressão pESC-URA;
- b) Alinhar sequências de oxidoesqualeno ciclases (OSCs) e avaliar a estrutura tridimensional de *MiFRS* para selecionar potenciais sítios de mutação;
- c) Realizar reações de mutação sítio-dirigida com os genes de friedelina sintase;
- d) Confirmar as mutações a partir da realização de sequenciamento;
- e) Transformar os mutantes em *Saccharomyces cerevisiae* e induzir a expressão;
- f) Extrair e quantificar a friedelina produzida por GC-MS.

3 MATERIAIS E MÉTODOS

3.1 Equipamentos

- Sequenciador ABI 3130 Genetic Analyzer (Applied Biosystems®);
- Cromatógrafo Gasoso GC2010 Plus com injetor automático AOC6000 acoplado à Espectrômetro de Massas QP-2020 (Shimadzu®);
- Coluna capilar RTX-5MS – C = 30m, L = 0,25mm, E = 0,25µm (Restek®);
- Liofilizador Alpha 1-2 LD Plus (Christ®).
- Incubadora B.O.D 347-CD (Fanem®);
- Incubadora Shaker Innova 4430 (New Brunswick Scientific®);
- Incubadora Shaker Innova 4000 (New Brunswick Scientific®);
- Estufa 002-CB (Fanem®);
- Espectrofotômetro NanoDrop 2000 (Thermo Fisher®);
- Termociclador T-Advanced (Biometra®);
- Fluxo laminar (Velco®);
- Rotaevaporador 144 (Buchi®);
- pHmetro PHi 510 (Beckman Coulter®);
- Balança analítica AB204-S (Mettler Toledo®);
- Pipetadores automáticos (Gilson®);
- Sistema de purificação de água Milli-Q (Millipore®);
- Sonicador USC-2800 (Unique®);
- Centrífuga 5417-R (Eppendorf®);
- Centrífuga 5818-R (Eppendorf®);
- Centrífuga de bancada 5415-D (Eppendorf®)
- Autoclave vertical (Primatec®);
- Espectrofotômetro DU 800 (Beckman Coulter®);
- Banho de aquecimento 1100 (Fanem®);
- Chapa de aquecimento (Lab-Line®);
- Câmara UV (Syngene®);

3.2 Soluções Preparadas

- Agarose 1% (v/v) (Sigma-Aldrich®);

- Brometo de etídeo 0.5 µg/mL
- Ladder de DNA 1Kb⁺;
- Solução estoque de Ampicilina 50 mg/mL;
- Solução estoque de Gentamicina 100 mg/mL;
- Solução estoque de Cloroanfenicol 100 mg/mL;
- Solução estoque de dNTP 100 mM;
- Galactose 10% (V/V);
- Glicose 10% (V/V);
- Tampão TE pH 8.0;
- Acetato de Lítio 1 M;
- Acetato de Lítio 100 mM;
- PEG 50% (m/V);
- SS-DNA 2 mg/mL;
- NaOH 20 mM;
- Solução de NaCl 0,73% (m/V);
- Clorofórmio grau HPLC;
- Solução de clorofórmio:metanol 2:1 (V/V);

3.3 Enzimas e reagentes

- Padrão analítico de Colesterol (Sigma Aldrich®);
- Padrão analítico de Friedelina (Sigma Aldrich®);
- Platinum™ Pfx DNA polymerase (Invitrogen®);
- *GoTaq® G2 DNA Polymerase* (Promega®);
- Phusion High-Fidelity DNA Polymerase (Thermo Fisher®);
- Gateway® BP Clonase II Enzyme Mix (Invitrogen®);
- Gateway® LR Clonase II Enzyme Mix (Invitrogen®);
- PureLink™ Quick Plasmid Miniprep Kit (Invitrogen®);
- GeneJet Genomic DNA Purification Kit (Thermo Fisher®);
- BigDye™ Terminator v3.1 Cycle Sequencing (Thermo Fisher®);
- BigDye XTerminator™ Purification Kit;
- Enzima de restrição *DpnI* (New England Biolabs®);

3.4 Genes, Plasmídeos e linhagens de microrganismos.

- Genes sintetizados de friedelina sintase (*MiFRS* e *PdFRS*, Genone Biotechnologies®);
- Plasmídeo pDONR™ 207 (Invitrogen®);
- Plasmídeo pESC-URA (Agilent Technologies®) [P_{GAL1} -*tHMG1*, P_{GAL10} , *TRP1*, 2 μ];
- Bactéria competente *Escherichia coli* DH5- α (Invitrogen®);
- Levedura *Saccharomyces cerevisiae* CEN.PK.113-5D (*MATa MAL2-8c SUC2 ura3-52 P_{ERG7}::P_{KEX2}-ERG7*).

3.5 Meios de cultura

- Meio Luria-Bertani líquido - 2,52 g do caldo Luria-Bertani (Kasvi®) em 100 mL de água ultrapura;
- Meio Luria-Bertani sólido - 2,52 g do caldo Luria-Bertani (Kasvi®) e 2 g de Ágar (Kasvi®) em 100 mL de água ultrapura;
- Meio YPD líquido - 1 g de Extrato de levedura (Kasvi®), 2 g de Peptona (Kasvi®) e 10 mL de glicose 20% (V/V) em 100 mL de água ultrapura;
- Meio sintético completo para auxotrofia de uracila (SC-URA) líquido – 0,67 g do reagente Levedura Nitrogênio Base sem aminoácidos (Difco®), 0,2 g do *drop-out* de aminoácidos utilizados (receita no anexo A-2), 10 mL de açúcar 20% (V/V) em 100 mL de água ultrapura;
- Meio sintético completo para auxotrofia de uracila (SC-URA) sólido – 0,67 g do reagente Levedura Nitrogênio Base sem aminoácidos (Difco®), 0,2 g do *drop-out* de aminoácidos utilizados (receita no anexo A-2), 2 g de Ágar (Kasvi®), 10 mL da solução de glicose ou galactose 20% (V/V) em 100 mL de água ultrapura;

4 METODOLOGIA

Os ensaios de biologia molecular aqui descritos foram realizados em colaboração com os Prof. Dr. Cleslei Fernando Zanelli e Prof. Dr. Sandro Roberto Valentini no Laboratório de Biologia Celular e Molecular de Microrganismos da Faculdade de Ciências Farmacêuticas de Araraquara (UNESP). A Dra. Tatiana Maria de Souza Moreira supervisionou as etapas do trabalho desenvolvido no laboratório.

4.1 Subclonagem das sequências codificadoras dos genes selvagens de friedelina sintase

4.1.1 Seleção dos genes de friedelina sintase

Para a realização desse estudo, foram selecionados os genes que codificam as enzimas friedelina sintase de *Monteverdia ilicifolia* (Número de acesso GenBank: AVR49025.1) e de *Populus davidiana* (Número de acesso GenBank: ART66198.1).

A escolha por esses genes, dentre as mais de quatorze sequências atualmente depositadas no *National Library of Medicine* (NCBI), se deu pelo fato dessas enzimas serem monofuncionais, ou seja, capazes de produzir somente a friedelina. Desse modo, a produção de outros terpenos após a expressão dos genes modificados poderia ser verificada de modo qualitativo.

Para esse estudo, ambos os genes foram sintetizados comercialmente com otimização dos códons para expressão em *S. cerevisiae* e clonados em um plasmídeo com marca de seleção de ampicilina pela empresa Genone Biotechnologies®.

4.1.2. Purificação dos genes (Miniprep)

As bactérias contendo o DNA recombinante com as sequências codificadoras de friedelina sintase de *M. ilicifolia* e *P. davidiana* foram inoculadas em 5 mL de meio Luria-Bertani com ampicilina (LB+Amp) por 16 horas (37 °C, 200 rpm), sendo posteriormente purificados através do protocolo descrito no kit PureLink™ Quick Plasmid Miniprep Kit (Invitrogen®). A quantidade de DNA molde utilizada para as reações posteriores foi determinada após a medição da concentração da solução purificada no equipamento NanoDrop 2000 (Thermo Fisher®).

4.1.3 Clonagem por Gateway®

O sistema Gateway® (Invitrogen) é uma técnica de biologia molecular que permite a transferência eficiente de fragmentos de DNA entre plasmídeos usando sequências específicas de recombinação, chamadas de “*site-specific attachment site* (*att*)”. A metodologia se baseia na utilização da recombinação sítio-específico do bacteriófago λ , onde uma região específica da sequência é ligada nos dois flancos de um DNA e esse é recombinado com um vetor a partir de duas reações simples, usando um mix de enzimas fornecido pelo fabricante e sem a necessidade da utilização de enzimas de restrição ou ligação (REECE-HOYES; WALHOUT, 2018).

4.1.3.1 Desenho dos primers *attB* e reação de amplificação

O primeiro passo para a realização dessa clonagem foi desenhar os primers que se ligam aos genes desejados e inserem nos mesmos as sequências recombinantes (*attB*).

Como essa etapa se dá por uma reação de PCR comum, os primers devem obedecer a certos requisitos, como possuir de 40-60% das bases guanina (G) e citosina (C) e possuir temperatura de fusão (*Melting temperature* - *T_m*) entre 50-55 °C.

O primer *forward*, que é complementar à sequência 5'-3' do DNA, deve possuir de 18-25 pares de base (bp) do gene a ser clonado, até 25 pares de base da sequência recombinante *attB* e iniciar com quatro resíduos de guanidina.

O primer *reverse*, que é o reverso complementar da sequência 3'-5' do DNA, deve possuir de 18-25 pares de base do gene a ser clonado, até 25 pares de base da sequência recombinante *attB* e iniciar com quatro resíduos de guanidina.

Desse modo, temos para os genes de friedelina sintase de *Monteverdia ilicifolia* (MiFRS) e para a friedelina sintase de *Populus davidiana* (PdFRS) os seguintes primers apresentados na Tabela 1, sendo a sequência recombinante *attB* apresentada em azul e as bases utilizadas para ajustar a janela de leitura (*open reading frame*) em vermelho. Os oligonucleotídeos utilizados nesse trabalho foram obtidos comercialmente e a sequência dos primers foram checadas a partir da utilização do software PrimerX (Bioinformatics.org).

Tabela 1 – Primers utilizados para a inserção da sequência *attB* nos genes de interesse.

Gene	Primers	Sequências
<i>MiFRS</i>	forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGTGGAAAATCAAATAGCCG
	reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTATCAAGCACTGTTTATCTTCATAGCC
<i>PdFRS</i>	forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGTGGAGATTGAAGATTGC
	reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTATTACAACCTTCTTAGATGGTAATGGGAC

Para a inserção da sequência de *attB* nos genes de interesse, os primers desenhados foram empregados em reações em cadeia da polimerase (PCR) utilizando a enzima Platinum™ Pfx DNA polymerase (Invitrogen®) e seus respectivos reagentes, como apresentado na Tabela 2.

Os genes de friedelina sintase *MiFRS* e *PdFRS*, utilizados nessa reação como DNA molde foram obtidos como descrito no item 4.1.2.

Tabela 2 – Componentes utilizados na reação de PCR para inserção da sequência *attB* nos genes de interesse.

Componente	Volume utilizado
Tampão da enzima	6,25 µL
mix dNTP 100 mM	1,5 µL
MgSO ₄ 50 mM	1 µL
Primer forward 12,5 µM	0,75 µL
Primer reverse 12,5 µM	0,75 µL
DNA molde	200 ng
Pfx DNA polymerase	0,5 µL
Água	q.s.p. 50 µL

Preparou-se a mistura reacional em um microtubo de 200 µL e a mesma foi submetida a amplificação utilizando de um termociclador Analytik Jena (T-Advanced Biometra®), sendo o programa de temperatura utilizado apresentado na Tabela 3.

Tabela 3 – Condições de PCR utilizadas para a amplificação dos genes.

Etapas	Temperatura	Tempo	Ciclo
Desnaturação inicial	95 °C	10 min	-
Desnaturação	95 °C	30 seg	35x
Anelamento	68 °C	1 min/1000 bases	
Extensão	68 °C	15 min	-

Após a realização da PCR, a mistura reacional foi analisada em gel de agarose 0,8% (m/v) com auxílio de *ladder* de DNA e revelada em câmara UV

Syngene®. O inserto visualizado foi recortado do gel e purificado seguindo o protocolo do kit GeneJet Genomic DNA Purification Kit (Thermo Fisher®).

4.1.3.2 Reação BP

A reação BP é a etapa seguinte da clonagem via Gateway® onde o substrato da PCR anterior (gene-*attB*) é clonado dentro de um vetor doador, ou seja, um plasmídeo contendo o sítio *attP* que será recombinado. A Figura 7 esquematiza essa clonagem com a entrada no gene desejado no plasmídeo, mostrando que o mesmo substitui o gene que confere resistência ao cloranfenicol.

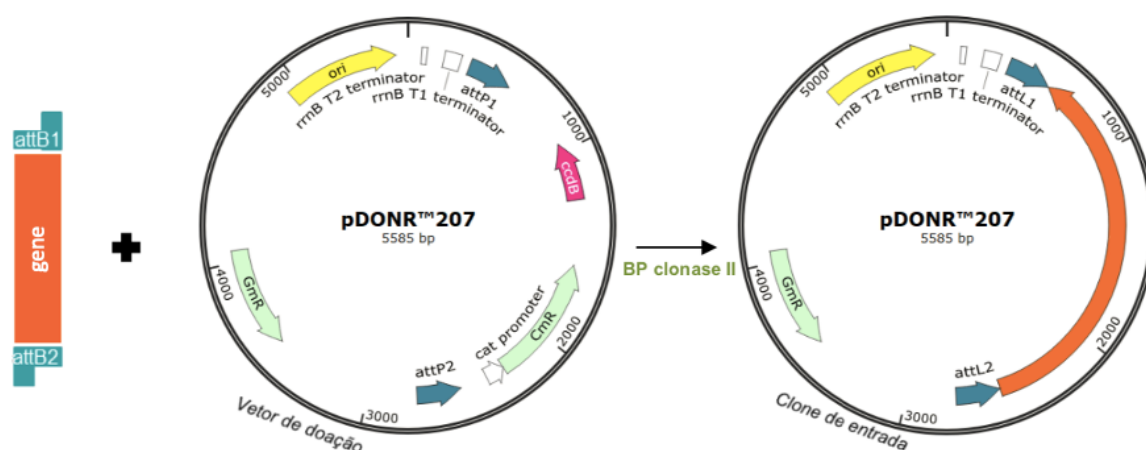


Figura 7 – Representação da reação BP para a produção do clone de entrada. Fonte: Imagem fornecida pelo autor.

Para essa etapa, o plasmídeo *pDONR™ 207* (vetor doador) foi descongelado e inoculado em meio líquido Luria-Bertani contendo os antibióticos gentamicina e cloroanfenicol (LB+Gem+Clo) em incubadora com agitação constante (37 °C, 16h, 200 rpm). Após, realizou-se a purificação desse inóculo de acordo com o descrito no item 4.1.2.

A reação BP foi realizada a partir da mistura reacional apresentada na Tabela 4. O mix de enzimas utilizado foi obtido a partir do kit Gateway® BP Clonase II Enzyme Mix (Invitrogen®).

Tabela 4 – Mistura reacional da reação BP.

Componente	Quantidade utilizada
miniprep gene- <i>attB</i>	50 fmol
miniprep pDONR™ 207	50 fmol
Mix de enzimas BP Clonase	2 µL
Tampão TE pH 8.0	q.s.p 8 µL

O microtubo contendo a reação foi incubado em banho de aquecimento *overnight* (25 °C, 18h) e após, o DNA amplificado foi adicionado em bactéria competente seguindo o protocolo apresentado no anexo A-1, sendo plaqueada em meio seletivo LB+Gem. A placa foi incubada em estufa (37 °C, 16 h) e após um dia verificou-se o aparecimento de colônias.

O produto formado pela reação consiste no gene da friedelina sintase clonado no plasmídeo pDONR™ 207.

4.1.3.3 Reação LR

A reação LR é a segunda e última da clonagem Gateway®. Nessa etapa, o produto da reação BP (gene-*attL*) é purificado e clonado para um vetor de expressão, ou seja, um plasmídeo com alto número de cópias que pode ser expresso em meio heterólogo. A Figura 8 esquematiza essa clonagem com a entrada no gene desejado no plasmídeo, mostrando que o mesmo substitui o gene que confere resistência ao cloranfenicol.

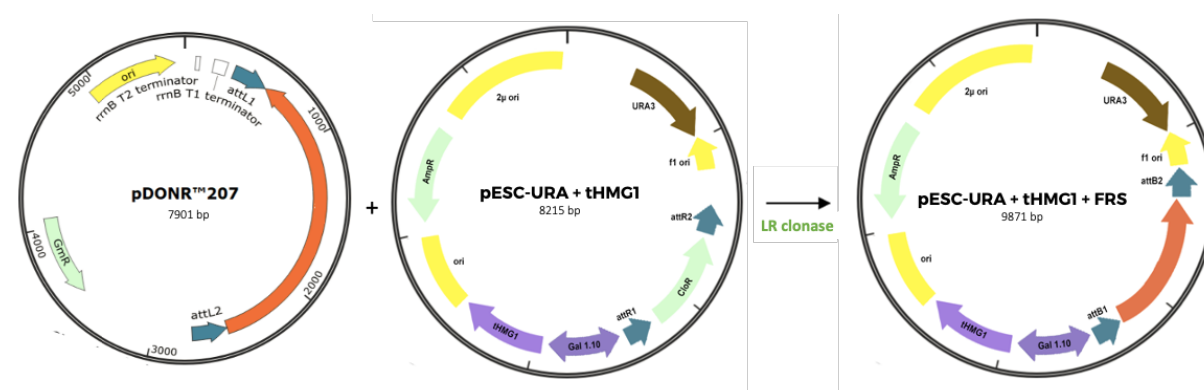


Figura 8 – Representação da reação LR para a produção do clone de expressão. Fonte: Imagem fornecida pelo autor.

Para essa etapa, o plasmídeo pESC-URA (vetor de destino) foi descongelado e inoculado em meio líquido Luria-Bertani contendo os antibióticos ampicilina e

cloranfenicol (LB+Amp+Clo) em incubadora shaker (37 °C, 16 h, 200 rpm). Após, realizou-se a purificação desse inóculo com o kit de miniprep utilizado anteriormente.

A reação LR foi realizada a partir da mistura reacional apresentada na Tabela 5. O mix de enzimas utilizado foi obtido a partir do kit Gateway® LR Clonase II Enzyme Mix (Invitrogen®).

Tabela 5 – Mistura reacional da reação LR.

Componente	Quantidade utilizada
miniprep gene- <i>attL</i>	150 ng
miniprep pESC-URA	150 ng
Mix de enzimas BP Clonase	2 µL
Tampão TE pH 8.0	q.s.p 8 µL

O microtubo contendo a reação foi incubado em banho de aquecimento *overnight* (25 °C, 18 h) e posteriormente o DNA resultante da PCR foi inserido em bactéria a partir de um procedimento conhecido como transformação do plasmídeo em bactéria (seguindo o protocolo apresentado no anexo A-1) e plaqueado em meio LB+Amp.

O aparecimento de colônias corresponde aos clones de friedelina sintase *MiFRS* e *PdFRS* inseridos no plasmídeo pESC-URA. Esse plasmídeo, foi selecionado por apresentar um promotor induzível por galactose, sendo descrito por Bicalho e colaboradores (2019) como o mais reprodutível para a expressão de genes de friedelina sintase em levedura.

Além disso, o gene truncado de 3-hidroxi-3-metil-glutaril-CoA redutase (*tHMG1*) também está clonado neste vetor de expressão. Ao ser expresso, esse gene é transcrito para a enzima HMG-CoA redutase, fornecendo o substrato HMG-CoA para a via do mevalonato (OHYAMA et al., 2007).

Para a confirmação da eficiência da clonagem, o procedimento de sequenciamento foi realizado.

4.1.4 Sequenciamento dos genes clonados

Após a realização da clonagem Gateway®, cada gene de friedelina sintase (*MiFRS* e *PdFRS*) foi inserido no plasmídeo pESC-URA juntamente com o gene *tHMG-1*. Apesar da taxa de sucesso dessa clonagem ser grande, deve-se garantir que nenhuma mutação involuntária ocorreu durante o procedimento.

Para isso, realizou-se o sequenciamento dos genes a partir da utilização do kit BigDye™ Terminator v3.1 Cycle Sequencing (Thermo Fisher®). Para essa reação, foi realizada uma miniprep (item 4.1.2) a partir do inóculo da bactéria transformada com o plasmídeo pESC-URA [P_{GAL1} -*tHMG1*, P_{GAL10} -FRS, *TRP1*, 2 μ] (pESC-URA+*tHMG1*+FRS). Além disso, utilizou-se do tampão presente no kit, da enzima BigDye™ e de primers *forward* desenhados para se alinharem os genes de friedelina sintase (anexo B-1). A mistura reacional utilizada foi preparada em um microtubo e consistiu nos componentes apresentados na Tabela 6.

Tabela 6 – Mistura reacional da reação de sequenciamento.

Componente	Quantidade utilizada
miniprep pESC-URA+ <i>tHMG1</i> + FRS	500 ng
Primer 3,2 μ M	2 μ L
Tampão 5x	3,5 μ L
BigDye™	1 μ L
Água	q.s.p. 20 μ L

Após, a mesma foi submetida a amplificação utilizando termociclador, sendo o programa de temperatura utilizado apresentado na Tabela 7.

Tabela 7 – Condições de PCR utilizadas para o sequenciamento.

Etapa	Temperatura	Tempo	Ciclo
Desnaturação inicial	96 °C	1 min	-
Desnaturação	96 °C	10 seg	
Anelamento	50 °C	5 seg	25x
Extensão	60 °C	4 min	
Extensão final	60 °C	5 min	-

Após a amplificação, as amostras foram precipitadas com o kit BigDye XTerminator™ Purification Kit (Thermo Fisher®) seguindo o protocolo do fabricante. Para cada amostra sequenciada, foram utilizados 10 μ L de solução X-Terminator, 45 μ L de solução SAM e 10 μ L da PCR. As amostras foram submetidas a homogeneização em vórtex por 30 minutos, centrifugadas (2 min, 4000 rpm) e encaminhadas para análise no sequenciador Genetic Analyzer 3130 (Applied Biosystems®) do Laboratório Multiusuários da Faculdade de Ciências Farmacêuticas de Araraquara (UNESP).

4.1.5 Transformação em *S. cerevisiae*

A transformação de alta eficiência é uma metodologia utilizada para inserir o plasmídeo de expressão em uma levedura conhecido por “método acetato de lítio/SS-DNA carreador/polietilenoglicol” (GIETZ; SCHIESTL, 2007).

Todas as transformações em levedura realizadas neste trabalho utilizaram a linhagem de *Saccharomyces cerevisiae* CEN.PK 113-5D (*MATa MAL2-8c SUC2 ura3-52 P_{ERG7}::P_{KEX2}-ERG7*).

Inicialmente, a linhagem de levedura foi descongelada e plaqueada em meio YPD (30 °C, 16 h). Posteriormente, uma colônia foi selecionada e inoculada em 5 mL de meio líquido YPD para o preparo de um pré-inóculo (30 °C, 16 h, 200 rpm). No dia seguinte, a concentração das células foi quantificada por espectrofotômetro DU 800 (Beckman Coulter®) a 600 nm.

Calculou-se o volume de células necessário para obter uma concentração de 5×10^6 células/mL e diluiu-se em 10 mL de meio YPD, incubando por 3-4 horas (30 °C, 200 rpm).

Após, centrifugou-se o meio e o pellet de células foi lavado com 5 mL de água ultrapura. Na sequência, o pellet foi ressuspendido em 1 mL de acetato de lítio 100 mM estéril e incubado por 15 minutos (30 °C).

As células foram centrifugadas e adicionou-se os componentes do mix de transformação apresentados na Tabela 8.

Tabela 8 – Mix utilizado para a reação de transformação em levedura.

Componente	Quantidade utilizada
50% PEG	480 µL
LiAc 1M	72 µL
SS-DNA 2 mg/mL	100 µL
miniprep pESC-URA+ <i>tHMG-1</i> + FRS	5 µL
Água	63 µL

A mistura reacional foi vortexada por 1 minuto e incubou-se a mesma em agitador (30 °C, 30 min, 200 rpm). Posteriormente a amostra foi levada para banho-maria (42 °C, 20 min), adicionou-se 500 µL de meio líquido YPD e a amostra foi novamente incubada sob agitação (30 °C, 2 h).

Coletou-se as células por centrifugação e o pellet foi ressuspensionado em 200 μ L de água ultrapura e plaqueado em meio seletivo SC-URA. As colônias transformadas com o plasmídeo pESC-URA apareceram após 2-3 dias de incubação em B.O.D. (30 °C, 72 h).

4.1.6 PCR de colônia

Para confirmar a presença do plasmídeo após a transformação em levedura, algumas colônias foram analisadas por PCR de colônia.

Utilizou-se o protocolo da enzima *GoTaq® G2 DNA Polymerase* (Promega®). Inicialmente foram adicionados 10 μ L de NaOH 20 mM em cada tubo de PCR para quebrar a parede celular da bactéria. Em seguida, uma pequena quantidade de cada colônia a ser analisada foi transferida para esse tubo e o mesmo foi incubado em termociclador (95 °C, 10 min).

Após, o conteúdo do tubo foi centrifugado e o sobrenadante contendo o DNA (colônia) foi utilizado para a reação de PCR apresentada na Tabela 9. Os primers utilizando estão apresentados no anexo B-2.

Tabela 9 – Mix utilizado para a reação de transformação em levedura.

Componente	Quantidade utilizada
Tampão	4 μ L
dNTP 100 mM	0,4 μ L
Primer forward 5 μ M	1,6 μ L
Primer reverse 5 μ M	1,6 μ L
DNA (colônia)	1 μ L
Enzima GoTaq	0,1 μ L
Água	11,3 μ L

Os amplicons foram verificados por eletroforese em gel juntamente em gel de agarose 0,8% (m/v) e brometo de etídeo (0,5 μ g/mL) com auxílio de *ladder* de DNA e revelada em câmara UV Syngene®.

4.2 Realização de mutações sítio-dirigidas e quantificação dos produtos

Os genes selvagens *MiFRS* e *PdFRS* foram submetidos a reações de PCR para mutação sítio-dirigida e transformados em levedura a fim de avaliar a ação de alterações de um único aminoácido na função da enzima. Para isso, utilizou-se

técnicas de bioinformática para a seleção de posições-chave onde realizar as mutações, as quais foram transformadas em levedura, sendo então realizada a indução da expressão em sistema heterólogo, extração e quantificação da friedelina por cromatografia gasosa.

4.2.1 Alinhamento dos genes de OSCs através de técnicas de bioinformática

As enzimas oxidoesqualeno ciclases (OSCs) são o ponto central para o entendimento da biossíntese dos triterpenos e derivados. Porém, até o presente momento, não existem estudos na literatura que conseguiram isolar e cristalizar enzimas dessa família.

Desse modo, faltam informações sobre a estrutura terciária dessas proteínas, sendo a sequência primária dos aminoácidos que compõem essas enzimas a principal informação disponível em bancos de dados. A partir dessas informações, é possível utilizar técnicas de bioinformática para alinhar essas sequências primárias, verificar a similaridade entre as OSCs, realizar previsões por modelagem molecular e propor posições para a realização de mutações sítio-dirigidas.

Nesse estudo, para avaliar a especificidade da friedelina sintase, as isoformas de *Monteverdia ilicifolia* (Nº Acesso GenBank: AVR49025.1), *Kalanchoe daigremontiana* (Nº Acesso GenBank: ADK35125.1) e *Populus davidiana* (Nº Acesso GenBank: ART66198.1) foram alinhadas com OSCs que produzem triterpenos similares à friedelina, como apresentado na Figura 9.

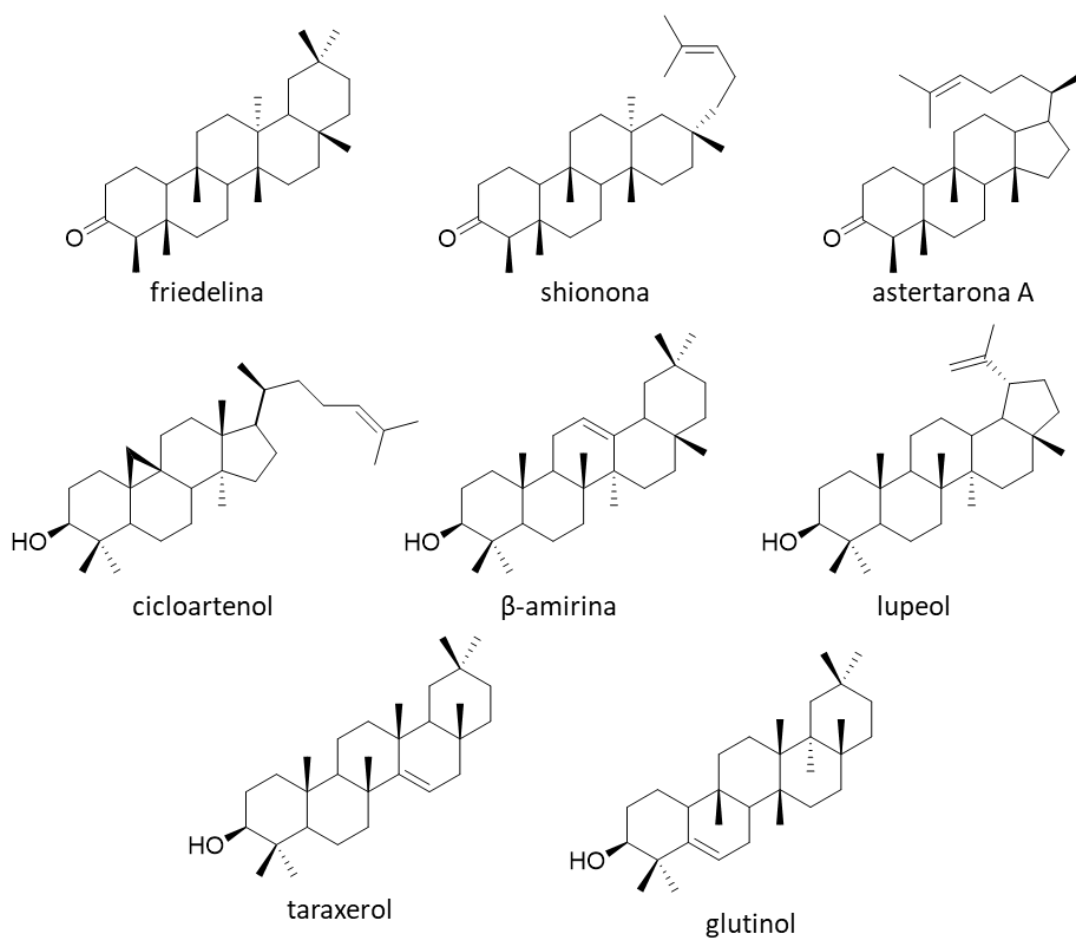


Figura 9 – Triterpenos produzidos a partir das enzimas utilizadas no alinhamento múltiplo. Fonte: Imagem fornecida pelo autor.

Além das enzimas que dão origem aos produtos da cascata de rearranjos apresentada anteriormente (Figura 4), as enzimas que produzem os triterpenos shionona e astertarona A também foram inseridas no alinhamento. Essa escolha deve-se ao fato de que essas estruturas, assim como a friedelina, possuem a posição C-3 oxidada, sendo esse grupo de enzimas denominado 3-ceto OSCs.

As enzimas utilizadas no alinhamento são apresentadas na Tabela 10.

Tabela 10 – Enzimas utilizadas no alinhamento múltiplo de sequências (MSA).

Enzima	Hospedeiro	Abreviação	Nº de acesso
friedelina sintase	<i>Monteverdia ilicifolia</i>	MiFRS	AVR49025.1
	<i>Kalanchoe daigremontiana</i>	KdFRS	ADK35125.1
	<i>Populus davidiana</i>	PdFRS	ART66198.1
shionona sintase	<i>Aster tataricus</i>	AtSHS	F8WQD0.1
astertarona A sintase	<i>Brassica rapa</i>	BrASS	XM_009142199
glutinel sintase	<i>Kalanchoe daigremontiana</i>	KdGLS	ADK35124.1
	<i>Betula platyphylla</i>	BpLUS	BAB83087.1
	<i>Lotus japonicus</i>	LjLUS	BAE53430.1
	<i>Glycyrrhiza glabra</i>	GgLUS	BAD08587.1
lupeol sintase	<i>Olea europaea</i>	OeLUS	BAA86930.1
	<i>Taraxacum officinale</i>	ToLUS	BAA86932.1
	<i>Kalanchoe daigremontiana</i>	KdLUS	ADK35126.1
	<i>Ricinus communis</i>	RcLUS	NP_001310684.1
	<i>Bruguiera gymnorhiza</i>	BgLUS	BAF80444.1
	<i>Glycyrrhiza glabra</i>	GgBAS	BAA89815.1
	<i>Pisum sativum</i>	PsBAS	BAA97558.1
	<i>Euphorbia tirucalli</i>	EtBAS	BAE43642.1
	<i>Betula platyphylla</i>	BpBAS	BAB83088.1
	<i>Panax ginseng</i>	PgBAS	BAA33461.1; BAA33722.1
β -amirina sintase	<i>Monteverdia ilicifolia</i>	MiCAS	APG38074.1
	<i>Luffa aegyptiaca</i>	LaCAS	BAA85266.1
	<i>Betula platyphylla</i>	BpCAS	BAB83085.1; BAB83086.1
	<i>Kalanchoe daigremontiana</i>	KdCAS	ADK35127.1
	<i>Brassica rapa</i>	BrCAS	XP_009136183.1; XP_009136183.1
	<i>Arabidopsis thaliana</i>	AtCAS	NP_001325220.1; NP_178722.1
	<i>Panax ginseng</i>	PgCAS	BAA33460.1
	<i>Pisum sativum</i>	PsCAS	BAA23533.1
	<i>Lotus japonicus</i>	LjCAS	BAE53431.1
	<i>Glycyrrhiza glabra</i>	GgCAS	BAA76902.1
cicloartenol sintase	<i>Glycyrrhiza glabra</i>	GgBAS	BAA89815.1

Utilizou-se da plataforma NCBI (*National Center of Biotechnology Information* - www.ncbi.nlm.nih.gov) para a obtenção das sequências primárias das enzimas, da

ferramenta Clustal- ω (EMBL-EBI - www.ebi.ac.uk/Tools/msa/clustalo/) para a realização do alinhamento e do aplicativo BoxShade ExPASy (Swiss Institute of Bioinformatics - descontinuado) para a apresentação dos dados de alinhamento.

Um modo de prever a estrutura tridimensional das OSCs de friedelina sintase é a partir da modelagem molecular. Utilizou-se da ferramenta Swiss-Model (Swiss Institute of Bioinformatics - <https://swissmodel.expasy.org>) para obter uma predição da estrutura molecular da enzima friedelina sintase por homologia com a estrutura cristalográfica da lanosterol sintase de humano (Protein Data Bank ID: 1W6K).

4.2.2 Realização das reações de mutação sítio-dirigida

Após a seleção de posições-chave para a atividade da enzima, realizou-se a PCR de mutação sítio-dirigida. Para essa reação, desenhou-se primers *forward* e *reverse* altamente complementares a sequência do gene selvagem na mesma região, sendo a única diferença a alteração pontual de 1-3 nucleotídeos, alterando o códon e consequentemente o aminoácido que é transcrito e traduzido em uma determinada posição.

Os primers para essa PCR devem possuir o(s) nucleotídeo(s) a serem mutado(s) no centro, com 10-20 nucleotídeos de cada lado da mutação, possuírem de 40-60% das bases guanina (G) e citosina (C) e temperatura de *melting* (T_m) entre 75-85 °C. A qualidade dos primers desenhados foi medida a partir da utilização da plataforma PrimerX (Bioinformatics.org), sendo os mesmos sintetizados e obtidos comercialmente.

Em sequência, a reação de mutação sítio-dirigida foi realizada seguindo a metodologia descrita por Edelheit e colaboradores (2009). Nessa metodologia, uma reação é feita para cada um dos primers e posteriormente os dois produtos da PCR são unidos e anelados em um único tubo, aumentando a eficiência e a probabilidade de sucesso na mutação.

As reações foram realizadas com a enzima Phusion High-Fidelity DNA Polymerase (Thermo Fisher®), sendo a mistura reacional apresentada na Tabela 11.

Tabela 11 – Mix utilizado para as reações de mutação sítio-dirigida.

Componente	Quantidade utilizada
Plasmídeo pESC-URA+ <i>th</i> HMG-1+ FRS	500 ng
MgCl ₂	0,25 µL
dNTP 10 mM	0,75 µL
Tampão da enzima	5 µL
Primer <i>forward</i> ou <i>reverse</i>	40 pmol
Enzima Phusion DNA Polymerase	0,3 µL
Água	q.s.p. 25 µL

Após a adição dos componentes, cada um dos tubos de PCR foi incubado em termociclador com o programa de temperaturas descrito na Tabela 12.

Tabela 12 – Programa de temperatura utilizado para a reação de mutação sítio-dirigida.

Etapas	Temperatura	Tempo	Ciclo
Desnaturação inicial	98 °C	30 seg	-
Desnaturação	98 °C	10 seg	
Anelamento	55 °C	40 seg	30x
Extensão	72 °C	5 min	
Extensão final	72 °C	10 min	-

Em seguida, o conteúdo dos dois tubos foi unido e o anelamento prosseguiu em termociclador conforme descrito na Tabela 13.

Tabela 13 – Programa de temperatura utilizado para a o anelamento da reação de mutação.

Etapas	Temperatura	Tempo
	95 °C	5 min
	90 °C	1 min
Anelamento	80 °C	1 min
	70 °C	30 seg
	60 °C	30 seg
	50 °C	30 seg
	40 °C	30 seg

Posteriormente, o conteúdo do tubo de PCR foi submetido a uma reação de digestão para a linearização do DNA parental não-mutado (25 °C, 16h) para a linearização do DNA parental não-mutado. Para isso, adicionou-se ao produto de PCR anelado 1 µL da enzima *DpnI* (Thermo Fisher®), 5 µL do tampão CutSmart™ (Thermo Fisher®) e 4 µL de H₂O ultrapura.

Após a reação de digestão, o produto da mutação foi transformado em *E. coli* (conforme descrito no anexo A-1), plaqueado em meio seletivo LB+Amp e incubado em estufa (37 °C, 16 h).

As colônias resultantes foram inoculadas em meio líquido LB+Amp (37 °C, 16 h, 200 rpm), purificadas pelo kit de miniprep e sequenciadas conforme descrito no item 4.1.4. Para confirmar as mutações pontuais e a qualidade dos sequenciamentos foram utilizados os programas SnapGene® e SeqScanner 2.0 (Applied Biosystems®).

Os genes mutantes foram posteriormente transformados em levedura pelo protocolo descrito no item 4.1.5.

4.2.3 Expressão em sistema heterólogo

Após a obtenção dos mutantes no plasmídeo de expressão pESC-URA e a transformação dos mesmos na linhagem de *S. cerevisiae* desejada, realizou-se o cultivo da levedura e a expressão da enzima.

Para isso, as colônias selvagens e mutantes foram repassadas em placa de Petri contendo o meio seletivo SC-URA e incubadas em B.O.D. (30 °C, 72 h). Após, colônias isoladas foram inoculadas em 5 mL de meio SC-URA contendo 2% de glicose (30 °C, 16 h, 200 rpm), sendo o inóculo posteriormente centrifugado (5 min, 4000 rpm) e lavado com água ultrapura (10 mL).

Após, o *pellet* de células foi ressuspendido em 10 mL de SC-URA contendo 2% de galactose em uma densidade ótica (D.O.₆₀₀) inicial de 0.6. O cultivo na presença de galactose induz o promotor do plasmídeo pESC-URA para promover a transcrição da enzima friedelina sintase clonada, dando início à expressão da enzima.

O cultivo foi realizado durante três dias em modo dinâmico (30 °C, 200 rpm) e posteriormente o inóculo foi centrifugado e o material celular seco em liofilizador.

4.2.4 Extração da friedelina e quantificação por GC-MS

A extração dos triterpenos foi realizada seguindo a metodologia descrita por Souza-Moreira (2016). O *pellet* de células foi transferido para um tubo de vidro seco e pesado, onde adicionou-se aproximadamente 7 mL do solvente extrator

CHCl_3 :MeOH 2:1 (V/V). Também realizou-se a adição de 50 μL de uma solução de colesterol à 240 $\mu\text{g/mL}$, sendo essa substância utilizada como padrão interno.

O tubo foi levado para sonicação durante 10 minutos e após essa etapa adicionou-se 1,7 mL de uma solução de NaCl 0,73% (V/V) promovendo o aparecimento de duas fases. Realizou-se a centrifugação do tubo (4000 rpm, 5 min), recolhendo a fase inferior para um frasco de vidro lavado e seco.

Para a análise desse produto por cromatografia gasosa, ressuspendeu-se o conteúdo do frasco de penicilina em 300 μL de clorofórmio grau HPLC, filtrou-se em filtro 0,22 μm e analisou-se em cromatógrafo gasoso GC2010 Plus (Shimadzu®) acoplado a espectrômetro de massas QP-2020 (Shimadzu®).

Paralelamente, para a análise quantitativa da friedelina, curvas analíticas foram construídas a partir de uma solução estoque do padrão em clorofórmio na concentração de 1 mg/mL. A partir dessa solução, diluições sucessivas (1:1, V/V) foram realizadas no intervalo de 500-0,97 $\mu\text{g/mL}$, com a adição de 50 μL do padrão interno colesterol na concentração de 240 $\mu\text{g/mL}$ em todos os pontos.

Os pontos da curva analítica foram secos em vial, ressuspendidos em 300 μL de clorofórmio, filtrados e analisados por GC-MS. Para a análise qualitativa utilizou-se do cromatograma de íons totais (TIC), enquanto a análise quantitativa da friedelina foi feita a partir da utilização do monitoramento de íon selecionado (SIM), onde o friedelina foi monitorada pelo sinal em m/z 246 e o padrão interno colesterol monitorado pelo sinal em m/z 301.

As análises de GC-MS foram realizadas no equipamento QP2020C (Shimadzu®) com coluna HP-5 (30 m x 0.25 mm x 0.25 μm ; Agilent®). Os softwares Excel (Microsoft®), GCMS Solution 4.41 (Shimadzu®), Prism 9 (GraphPad®) e Origin 9 (OriginLab®) foram utilizados para a aquisição e tratamento dos dados cromatográficos.

5 RESULTADOS E DISCUSSÕES

As sequências primárias das enzimas friedelina sintase de *Monteverdia ilicifolia* (MiFRS) e de *Populus davidiana* (PdFRS) foram analisadas em termos de suas estruturas primárias por alinhamento global múltiplo de sequências (MSA) e preditas em termos de sua estrutura terciária pela utilização de técnicas de modelagem molecular. A partir desses resultados, selecionou-se resíduos em posições chave para a realização de mutações sítio-dirigidas em ambas as sequências, observando mudanças na especificidade e na atividade das enzimas, avaliando também a capacidade produtiva desses mutantes a partir da realização da quantificação da friedelina.

5.1 Estudos de modelagem molecular

Apesar dos esforços recentes utilizando cromatografia por afinidade e por troca iônica, até o momento não existem estudos na literatura que obtiveram sucesso na purificação da enzima friedelina sintase e demais OSCs de plantas.

Um modo de prever o enovelamento da proteína é através da realização de estudos de modelagem molecular por homologia. O uso de ferramentas de bioinformática permite, através da sequência primária de aminoácidos, construir um modelo baseando-se em uma proteína já cristalizada e tridimensionalmente determinada.

5.1.1 Estudos de modelagem para a enzima MiFRS

A partir da sequência primária da enzima MiFRS com 771 aminoácidos (anexo C-1), foi possível construir um modelo tridimensional e prever as interações e a influência dos resíduos presentes.

O modelo foi construído a partir do uso da ferramenta EasyModeller, que se baseia nos algoritmos do Modeller para a obtenção e refinamento do mesmo a partir do uso de um *template* (KUNTAL; APAROY; REDDANNA, 2010; WEBB; SALI, 2016).

No presente estudo, utilizou-se da estrutura cristalizada da lanosterol sintase de humano (Código PDB: 1W6K) como molde, sendo essa a enzima que apresentou

melhor identidade sequencial com *MiFRS*. A Figura 10 apresenta a sobreposição da lanosterol sintase (amarelo) e do modelo (em roxo).



Figura 10 – Modelo construído para a enzima *MiFRS* (roxo) utilizando como molde a enzima lanosterol sintase de humano (amarelo). Imagem fornecida pelo autor.

A partir do modelo obtido, verificou-se uma homologia de 39,72% com a lanosterol sintase de humano e a presença majoritária de estruturas do tipo α -hélice, além de folhas- β e a presença alças. A qualidade do modelo foi analisada a partir do uso da ferramenta Swiss-Model (*Swiss Institute of Bioinformatics*), que permitiu a obtenção do diagrama de Ramachandran.

O diagrama de Ramachandran é um gráfico que descreve a qualidade da estrutura cristalográfica modelada a partir da identificação de resíduos em regiões energeticamente favoráveis e desfavoráveis. A construção desse diagrama baseia-se nos ângulos de torção φ (anterior ao carbono α) e ψ (posterior ao carbono α) das ligações peptídicas.

Para o modelo predito, temos o diagrama de Ramachandran apresentado na Figura 11.

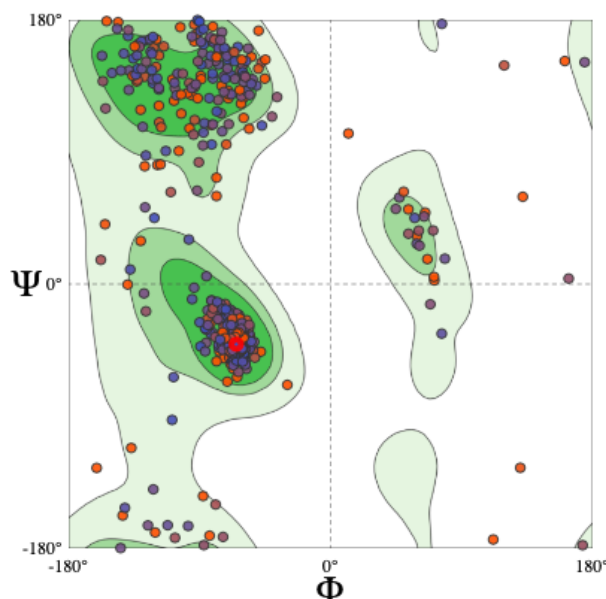


Figura 11 – Gráfico de Ramachandran para a enzima *MiFRS* modelada na lanosterol sintase. Fonte: Swiss-Model (Swiss Institute of Bioinformatics).

Cálculos teóricos da rotação dessas ligações levam a predição de diferentes regiões. Regiões onde não há a possibilidade de restrições estéricas são geometricamente favorecidas, enquanto regiões com uma pequena possibilidade de interação são vistas como regiões permitidas.

As regiões mais claras do diagrama representam conformações permitidas pela existência de uma pequena flexibilidade no ângulo diedro (verde claro) e regiões não permitidas (branco). Os pontos apresentam os aminoácidos adequados para aquela posição e a coloração dos mesmos está de acordo com a polaridade da cadeia lateral desses resíduos.

Os resultados obtidos demonstram que para o modelo temos 91,68% de regiões favorecidas geometricamente, 2,34% de regiões permitidas e 5,98% de resíduos em regiões não permitidas. Desse modo, como a ampla maioria dos resíduos estão em regiões permitidas, acredita-se que o modelo obtido é adequado para o prosseguimento dos estudos.

Esse modelo obtido para a enzima *MiFRS* foi refinado em colaboração com o Prof. Dr. Rafael Guido e o Dr. Andrew Oliveira para o trabalho de Mazzeu e colaboradores (2021). Esse modelo mais robusto foi utilizado nesse estudo para a determinação dos mutantes.

5.1.2 Estudos de modelagem para a enzima PdFRS

De modo semelhante, o programa EasyModeller também foi utilizado para a construção de um modelo a partir da sequência primária de aminoácidos da friedelina sintase de *Populus davidiana* (759 aminoácidos, Anexo C-2), utilizando novamente a lanosterol sintase de humano (Código PDB: 16KJ) como molde (Figura 12).

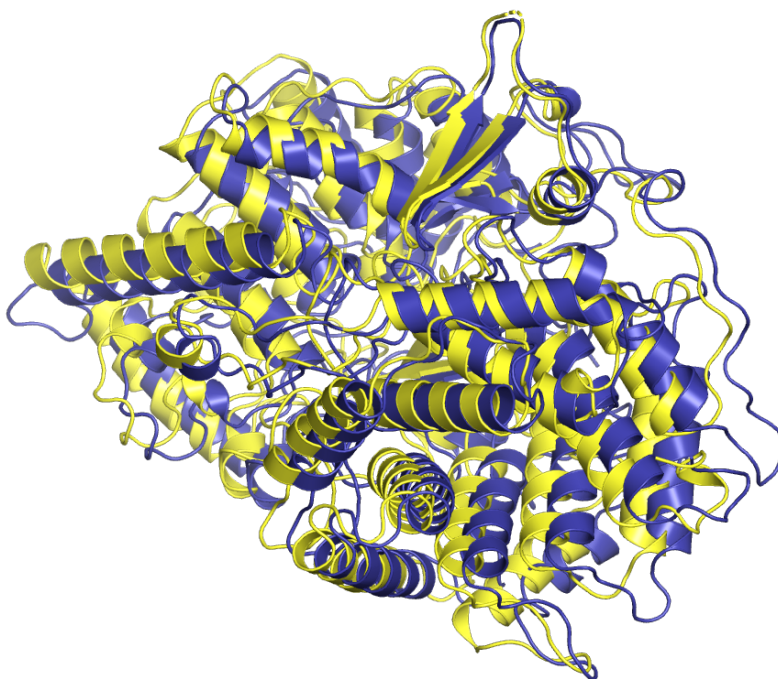


Figura 12 – Modelo construído para a enzima *MiFRS* (azul) utilizando como molde a enzima lanosterol sintase de humano (amarelo). Fonte: Imagem fornecida pelo autor.

O modelo produzido apresentou identidade de 39,58% com a lanosterol sintase de humano, possuindo majoritariamente estruturas de α -hélice e também estruturas em folhas- β e alças.

O diagrama de Ramachandran (Figura 13) demonstra que 92,60% dos aminoácidos estão em regiões favorecidas geometricamente, enquanto 2,51% aparecem em regiões permitidas e 4,89% dos resíduos estão presentes em regiões não permitidas.

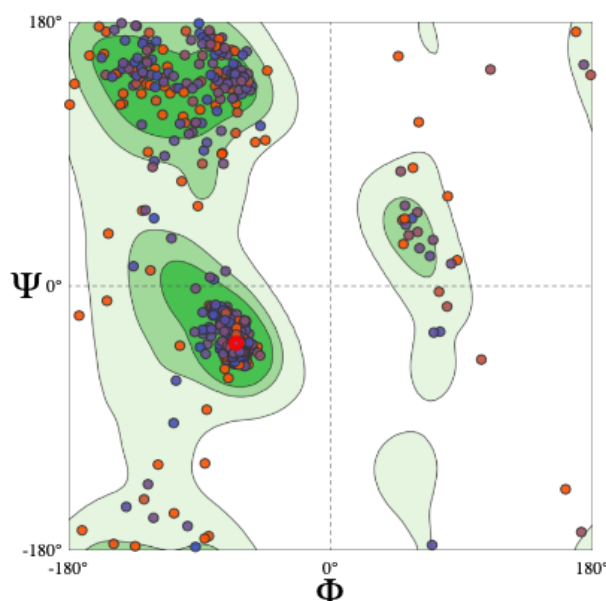


Figura 13 – Gráfico de Ramachandran para a enzima *PdFRS* modelada na lanosterol sintase. Fonte: Swiss-Model (Swiss Institute of Bioinformatics).

Os dados aqui apresentados demonstram que os modelos obtidos possuem qualidade e podem ser utilizados como guia para a seleção de posições para a realização de mutações sítio-dirigidas.

5.2 Alinhamento das sequências e seleção de mutantes

A partir do alinhamento múltiplo das sequências das enzimas oxidoesqualeno ciclases, foi possível avaliar a relação evolutiva e determinar resíduos para a realização de mutações.

5.2.1 Análise evolutiva das enzimas alinhadas

A árvore filogenética é uma representação gráfica da relação ancestral-descendente entre sequências genéticas (CALDART et al., 2018). É possível supor ancestrais comuns através dos nós e verificar diferentes grupos (*clusters*) que indicam a distância evolutiva entre as enzimas.

Para este estudo, a árvore filogenética foi construída com a sequências das enzimas apresentadas em 4.2.1 utilizando-se do método baseado em distância *Neighbor-Joining* (SAITOU; NEI, 1987) e do software MEGA 11 (TAMURA; STECHER; KUMAR, 2021). O resultado é apresentado na Figura 14.

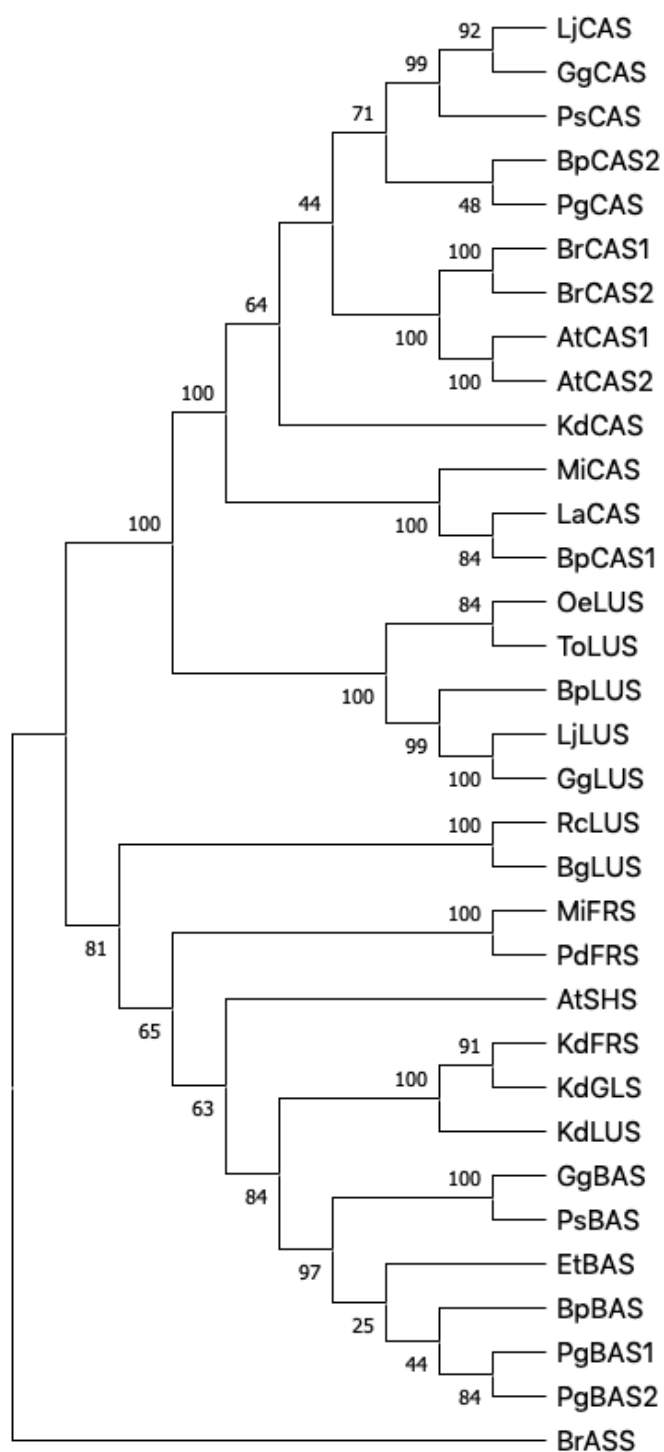


Figura 14 – Árvore filogenética produzida a partir das enzimas alinhadas. Fonte: Imagem fornecida pelo autor.

A partir da árvore filogenética foi possível perceber, baseado na identidade, que as enzimas friedelina sintase *MiFRS* e *PdFRS* são evolutivamente próximas. Além disso, em comparação com as demais enzimas que produzem compostos

carbonilados, vemos a presença de ancestrais comuns destas com *AtSHS* e *KdFRS*, sendo a exceção a enzima astertarona sintase (*BrASS*) que produz um terpeno tetracíclico.

Também foi possível observar que, apesar de possuir a mesma função que *MiFRS* e *PdFRS*, a enzima *KdFRS* é mais próxima das OSCs de *Kalanchoe daigremontiana*, o que evidencia que o grau de similaridade pode ser fortemente influenciado pela espécie de origem em vez da atividade (SOUZA-MOREIRA et al., 2016).

Como esperado, as enzimas de cicloartenol sintase que realizam ciclizações a partir do cátion protoesterila na conformação cadeira-barco-cadeira são as mais distantes das 3-ceto OSCs. Por fim, observa-se que as enzimas que produzem friedelina possuem similaridade evolutiva com as enzimas que produzem β -amirina e glutinol, o que está de acordo com a sequência de rearranjos da cascata de carbocátions observada na Figura 4.

5.2.2 Alinhamento múltiplo de sequências

O alinhamento global múltiplo de sequências (MSA) é uma ferramenta que permite identificar regiões conservadas e interações entre domínios de enzimas através de suas sequências primárias. Conforme descrito no item 4.2.1 e na Tabela 10, selecionou-se para a realização do alinhamento as 3-ceto OSCs e as demais oxidoesqualeno ciclases responsáveis pela produção de triterpenos a partir da cascata de rearranjos.

A Figura 15 traz uma visão geral das principais regiões observadas no alinhamento, utilizando a numeração dos aminoácidos da enzima *MiFRS* como referência. O arquivo completo do alinhamento está apresentado no anexo E-1.

			122				256-261
MiFRS	119	PAENAGV		253	PAKMWCYCRMVY		
KdFRS	116	PAENAGP		253	PAKMWCYCRMVY		
PdFRS	116	PAENSGV		250	PAKMWCYCRLVY		
AtSHS	116	PAENSGP		253	PGKMWSYCRLVF		
BrASS	117	PAENSGC		255	GGTVWIYLRDVF		
BpLUS	114	PGESAGP		251	PGKMLCYCRLVY		
KdLUS	116	PAENAGP		253	PGKMWCYCRMVY		
PsBAS	116	PAQIAGP		253	PAKMWCYCRLVY		
EtBAS	116	PAENAGP		253	PAKMWCYCRMVY		
MiCAS	120	PGDYGGP		256	PGRMWCHCRMVY		
KdCAS	116	PGDYGGP		252	PGRMWCHCRMVY		
			416				484-488
MiFRS	413	GSQLWDA		481	QLSDCTAEALK		
KdFRS	414	GSQHWDT		482	QLSDCTAEGLK		
PdFRS	411	GSQLWDA		479	QLSDCTAEALK		
AtSHS	413	GSQVWDA		481	QLSDCTTIGLT		
BrASS	416	GSQIWDT		484	PVSDCISESLK		
BpLUS	411	GCOMWDA		479	QVSDCTAEGLK		
KdLUS	414	GSQHWDT		482	QVSDCTAEGLK		
PsBAS	414	GSQEWDA		482	QVSDCTAEGLK		
EtBAS	414	GSQQWDT		482	QVSDCTAEGLK		
MiCAS	417	GSQLWDS		485	PISDCTAEGLK		
KdCAS	413	GSQLWDT		481	PISDCTAEGLK		
			533	609	729		
MiFRS	530	GLAGWEP		606	GSWYGNW	726	GVFKMNC
KdFRS	530	GMPGWEP		606	GSWYGNW	726	GVFKMNC
PdFRS	527	GLAGWEP		603	GSWYGNW	723	GVFKMNC
AtSHS	529	GLAAWEP		605	GSWYGEW	725	GAFKKNC
BrASS	532	GIAIWEP		608	GSWYGNW	728	GAYKMNL
BpLUS	527	GFPWEP		603	GSWYGNW	723	GVFMKNC
KdLUS	530	GLPAWEP		606	GSWYGNW	726	GVFMKNC
PsBAS	530	GLAAWEP		606	GSWYGNW	726	GVFMKNC
EtBAS	530	GLAAWEP		606	GSWYGNW	726	GVFMKNC
MiCAS	533	GFATYEL		609	GSWYGSW	729	GVFNKNC
KdCAS	529	GFATYEL		605	GSWYGSW	725	GVFNKNC

Figura 15 – Visão geral do alinhamento múltiplo e destaque para regiões importantes. Fonte: Imagem fornecida pelo autor.

Dentre os resíduos destacados, alguns já possuem sua importância e função determinados a partir do uso de mutação sítio-dirigida.

Estudos descrevem o resíduo 122 como responsável pela estabilização da conformação do 2,3-oxidoesqualeno, de modo que a presença da asparagina (N) favorece a conformação cadeira-cadeira-cadeira, enquanto a presença do

aminoácido aromático tirosina (Y) favorece a conformação cadeira-barco-cadeira (THOMA et al., 2004).

A posição 256-251 corresponde ao *motif* MWCYCR conservado na maioria das enzimas de friedelina sintase, β -amirina sintase e lupeol sintase. Os aminoácidos triptofano (W) e tirosina (Y) são úteis para a estabilização dos intermediários tetracíclicos e pentacíclicos da cascata de rearranjos, sendo essenciais para a produção dos anéis D e E por fornecerem interações do tipo cátion- π (ITO; NAKADA; HOSHINO, 2017; KUSHIRO et al., 2000).

O resíduo 484 é uma das regiões mais estudadas, correspondendo ao *motif* DCTAE altamente conservado (ITO; MASUKAWA; HOSHINO, 2013). Dados da literatura descrevem o resíduo de aspartato (D) como responsável pela protonação que dá origem a cascata de rearranjos do 2,3-oxidoesqualeno (ITO; MASUKAWA; HOSHINO, 2013; SATO; HOSHINO, 1999; SOUZA-MOREIRA et al., 2016), enquanto posições adjacentes como a leucina (L482) e fenilalanina (F473) mostram-se como essenciais para manter a atividade da enzima.

Os demais resíduos ainda não apresentam sua função determinada, de modo que estes foram selecionados para a realização de mutações sítio-dirigidas.

Posição 416: Uma posição que mostra diferença na natureza dos aminoácidos apresentados entre as 3-ceto OSCs e demais enzimas é a posição 416 (Figura 16).

416

MiFRS	405	DGMKVVSF	GSQIWDATF	GFQALVA
KdFRS	406	DGMKIQSF	GSQHWDTAF	FVQALLA
PdFRS	403	DGMKVQSF	GSQIWDATF	GFQALYT
AtSHS	405	DGMKMQTF	GSQVWDASIG	IQALLA
BrASS	408	DGLKIQTF	GSQIWDTAFLI	QMMLA
KdGLS	406	DGMKIQSF	GSQHWDTVF	SAQALLA
BpLUS	403	DGLKIQSF	GCQMWDAGFAI	QAILS
LjLUS	403	DGLKIQSF	GCQMWDAAFAI	QAILS
GgLUS	403	DGLKIQSF	GCQMWDAAFAI	QAILA
OeLUS	404	DGLKMQSF	GCQMWDAAFAI	QAILS
ToLUS	406	DGMKMQSF	GCQMWDAAFAI	QAIIFS
KdLUS	406	DGMKMQSF	GSQHWDTAF	SIQALLA
RcLUS	405	DGIKMQSF	GSQTWDTSLAL	QALIA
BgLUS	405	DGIKIQSF	GSQVWDTSEVL	QALIA
GgBAS	406	DGMTMQSF	GSQEWDAAGFAV	QALLA
PsBAS	406	DGMTMQSF	GSQEWDAAGFAV	QALLA
EtBAS	406	DGMKMQSF	GSQQWDTGFAI	QALLA
BpBAS	406	DGIKMQSF	GSQEWDTGFAI	QALLA
PgBAS1	407	DGMKMQSF	GSQEWDTGF	SIQALLD
PgBAS2	405	DGMKMQSF	GSQEWDTGFAI	QALLA
MiCAS	408	DGMKMQGYN	GSQLWDSAF	FAVQAIIS
LaCAS	411	DGMKMQGYN	GSQLWDTAF	FAVQAIMS
BpCAS1	413	DGMKMQGYN	GSQLWDTTF	FAVQAIIS
BpCAS2	403	DGMKMQGYN	GSQLWDTAF	FAVQAIIS
KdCAS	404	DGMKMQGYN	GSQLWDTAF	SVQAIIVA
BrCAS1	403	DGMKMQGYN	GSQLWDTSF	FAVQAVLA
BrCAS2	403	DGMKMQGYN	GSQLWDTSF	FAVQAVLA
AtCAS1	312	DGMKMQGYN	GSQLWDTGFAI	QAILA
AtCAS2	403	DGMKMQGYN	GSQLWDTGFAI	QAILA
PgCAS	403	DGMKMQGYN	GSQLWDTAF	FAVQAIIS
PsCAS	403	DGMKMQGYN	GSQLWDTAF	FAAQAIIS
LjCAS	403	DGMKMQGYN	GSQLWDTAF	FAAQAIIS
GgCAS	403	DGMKMQGYN	GSQLWDTAF	TAQAIIS

Figura 16 - Representação do MSA com destaque para a posição 416. Fonte: Imagem fornecida pelo autor.

Para as enzimas que produzem substâncias carboniladas verifica-se a presença de resíduos alifáticos, como a leucina (L), isoleucina (I), histidina (H) e

valina (V) na posição 416. As enzimas do tipo lupeol sintase apresentam majoritariamente a presença de metionina (M), as enzimas de β -amirina sintase conservam os resíduos com cadeia lateral carbonilada ácido glutâmico (E) e glutamina (Q) e as cicloartenol sintases possuem leucina (L).

A região ao redor da posição 416 mostra-se conservada, com a presença de resíduos de fenilalanina (F412) e triptofano (W417). Estudos na literatura descrevem que o resíduo de fenilalanina da posição 412 (F412) estabiliza a cavidade do sítio ativo a partir da presença de interações de Van der Waals, estando esse resíduo relacionado a ciclização no anel D (ITO; NAKADA; HOSHINO, 2017). Mazzeu e colaboradores (2021) realizaram mutações no triptofano da posição 417 (W417) e concluíram que as alterações levaram a perda de função da enzima, sendo o efeito estérico e a interação espacial deste aminoácido com o resíduo F473 os principais responsáveis pela estabilização dessa cavidade.

A partir do modelo construído para *MiFRS*, é possível verificar que o resíduo de leucina (L) é capaz de doar densidade eletrônica para resíduos vizinhos (Figura 17).

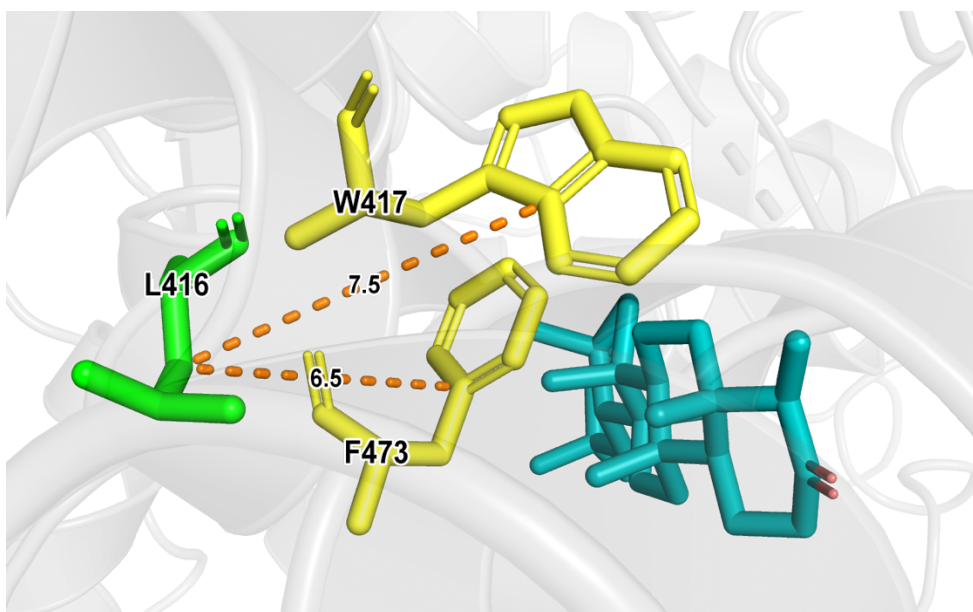


Figura 17 – Interação entre os resíduos L416 e F473. Fonte: Imagem fornecida pelo autor.

Desta forma, visto que essa região é importante para a atividade da enzima e que até o presente momento mutações não foram realizadas na posição 416, propusemos a troca do resíduo de leucina (L416) para os resíduos de glutamina

(L416Q) e histidina (L416H) Figura 18, sendo a justificativa para a escolha desses mutantes apresentada na Tabela 14.

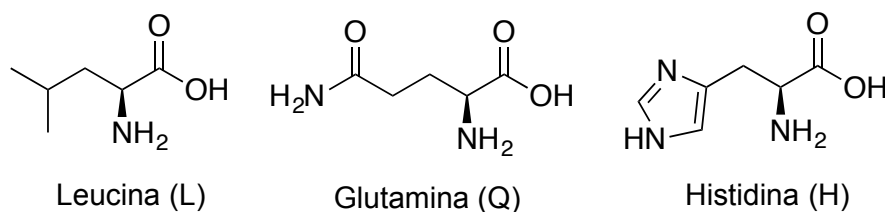


Figura 18 – Aminoácidos utilizados nos estudos de mutação sítio-dirigida da posição 416. Fonte: Imagem fornecida pelo autor.

Tabela 14 – Mutações propostas para a posição 416.

Aminoácido	Mutação	Característica físico-química	Justificativa
L416	Q	Polar não carregado	Alterar a característica físico-química do aminoácido.
	H	Polar carregado positivamente	Adicionar o resíduo presente nas demais OSCs.

As reações de PCR para a obtenção dos mutantes foram realizadas como descrito no item 4.2.2 para ambas as enzimas e os nucleotídeos alterados para as posições 416 (*MiFRS*) e para a respectiva posição 414 (*PdFRS*) são apresentados abaixo (Tabela 15).

Tabela 15 – Alterações realizadas para a obtenção dos mutantes da posição 416.

Gene	Alteração
<i>MiFRS</i> - L416Q	L → Q TTG por CAA
<i>MiFRS</i> - L416H	L → H TTG por CAT
<i>PdFRS</i> - L414Q	L → Q TTA por CAA
<i>PdFRS</i> - L414H	L → H TTA por CAT

O alinhamento múltiplo das sequências evidencia que as enzimas de friedelina sintase possuem um resíduo de glicina (G) nessa posição, enquanto as enzimas de β -amirina sintase e lupeol sintase possuem um resíduo de alanina (A) e as enzimas de cicloartenol sintase possuem um resíduo de treonina (T).

Devido a relatos da literatura, sabemos que a posição vizinha possui resíduos aromáticos conservados de triptofano (W) e tirosina (Y). Hoshino e colaboradores (2017) descrevem que o resíduo W534 da β -amirina sintase de *Euphorbia tirucalli* está relacionado com a ciclização do anel E. Wu e colaboradores (2020) descrevem para a oxidoesqualeno ciclase de *Iris tectorum* o resíduo de tirosina (Y534) nessa posição como essencial para manter o cátion lupenila na orientação ideal para a desprotonação e formação de triterpenos pentacíclicos.

Para a posição 533, o único estudo presente na literatura é para a cicloartenol sintase de *Pisum sativum*, que possui uma treonina (T). Takase e colaboradores (2015) supõem que esse resíduo está envolvido na etapa final de desprotonação. Porém, não existem até o momento relatos para as mutações na posição 533 para enzimas que ciclizam o oxidoesqualeno na conformação cadeira-cadeira-cadeira.

Além disso, a partir do modelo feito para *MiFRS*, é possível verificar que o resíduo de glicina (G) está a menos de 5Å de distância da carbonila C-3 (Figura 20).

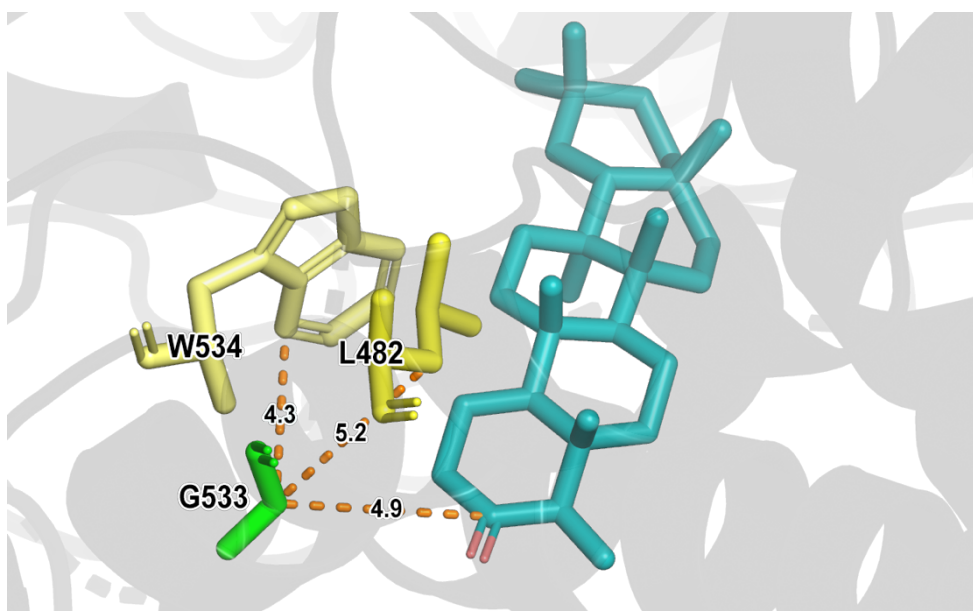


Figura 20 – Interação entre os resíduos G533, L482, W534 e a posição C-3. Fonte: Imagem fornecida pelo autor.

Devido a essas interações e ao alinhamento apresentado, optou-se por realizar nas enzimas *MiFRS* e *PdFRS* a troca do resíduo de glicina (G) por um resíduo de alanina (A) (Figura 21).

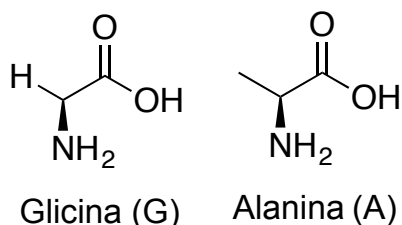


Figura 21 – Aminoácidos utilizados nos estudos de mutação sítio-dirigida para a posição 533. Fonte: Imagem fornecida pelo autor.

Desse modo, pode-se avaliar como essas enzimas se comportam com um aminoácido que possui as mesmas propriedades estéricas e físico-químicas, não alterando significativamente a região próxima do sítio ativo, o que poderia causar uma perda da função e da estabilidade da enzima (Tabela 16).

Tabela 16 – Mutações propostas para a posição 533.

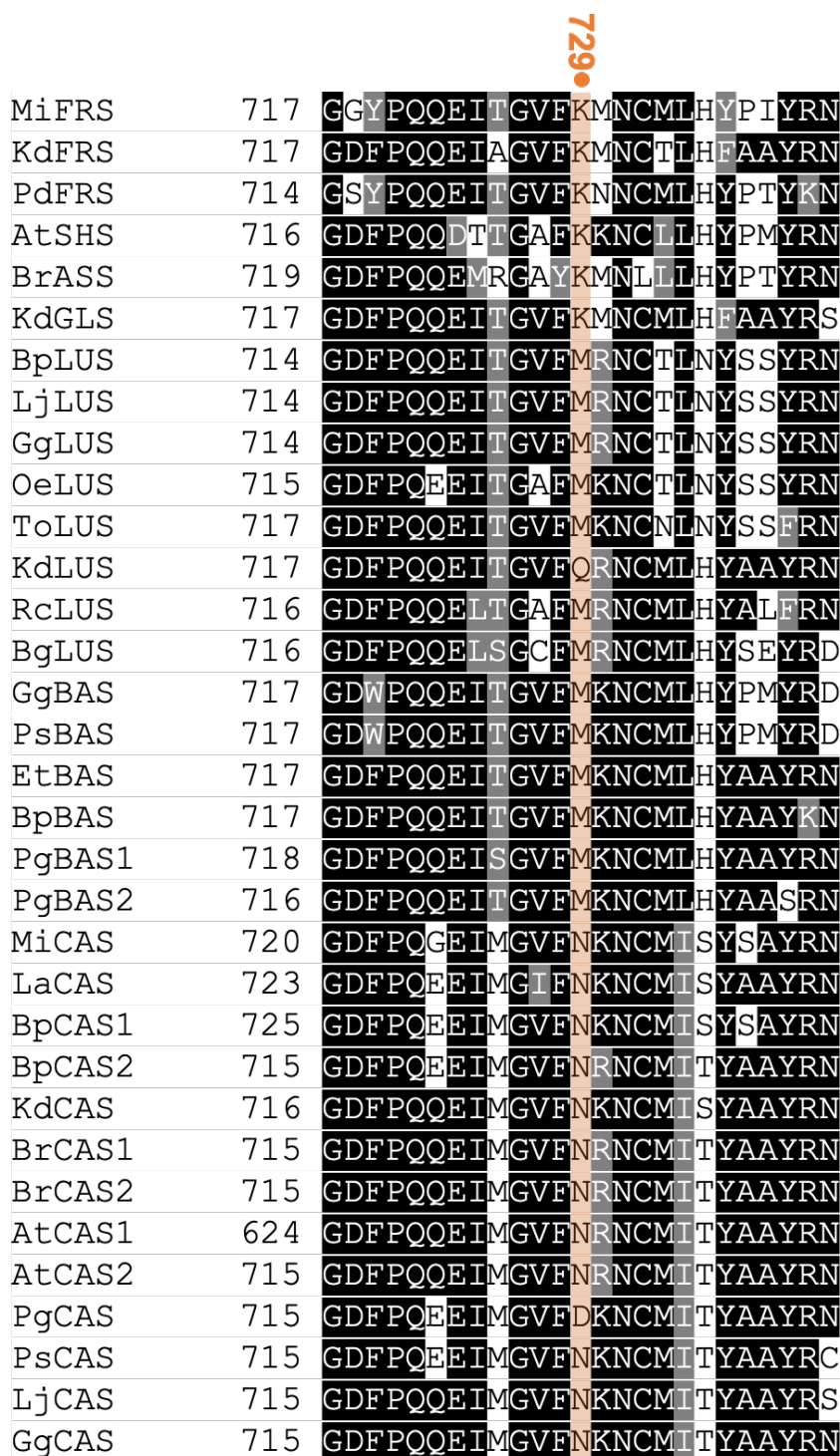
Aminoácido	Mutação	Característica físico-química	Justificativa
G533	A	Apolar alifático	Adicionar o resíduo presente nas demais OSCs.

As reações de mutação sítio-dirigidas foram realizadas (item 4.2.2) para as enzimas *MiFRS* e *PdFRS*, sendo os nucleotídeos alterados apresentados na Tabela 17.

Tabela 17 – Alterações realizadas para a obtenção dos mutantes da posição 533.

Gene	Alteração
<i>MiFRS</i> – G533A	G → A GGT por GCT
<i>PdFRS</i> – G530A	G → A GGT por GCT

Posição 729: Outros resíduos relacionados com a atividade das OSCs são os aminoácidos da posição C-terminal, como o resíduo da posição 729 (Figura 22).



MiFRS	717	GGYPQQEITGVFKMNCMLHYPIYRN
KdFRS	717	GDFPQQEITAGVFKMNC TLHFAAYRN
PdFRS	714	GSYPQQEITGVFKNNCMLHYPTYKN
AtSHS	716	GDFPQQDITGAFKKNC LLHYPMYRN
BrASS	719	GDFPQQEMRGAYKMNLL LHYP TYRN
KdGLS	717	GDFPQQEITGVFKMNCMLHFAAYRS
BpLUS	714	GDFPQQEITGVFMRNCTLNYS SYRN
LjLUS	714	GDFPQQEITGVFMRNCTLNYS SYRN
GgLUS	714	GDFPQQEITGVFMRNCTLNYS SYRN
OeLUS	715	GDFPQQEITGAFMKNC TLNYS SYRN
ToLUS	717	GDFPQQEITGVFMKNCNLNYS SFRN
KdLUS	717	GDFPQQEITGVFORNCMLHYAAYRN
RcLUS	716	GDFPQQELTGAFMRNCMLHYALFRN
BgLUS	716	GDFPQQELSGCFMRNCMLHYSEYRD
GgBAS	717	GDWPQQEITGVFMKNCMLHYPMYRD
PsBAS	717	GDWPQQEITGVFMKNCMLHYPMYRD
EtBAS	717	GDFPQQEITGVFMKNCMLHYAAYRN
BpBAS	717	GDFPQQEITGVFMKNCMLHYAAYKN
PgBAS1	718	GDFPQQEISGVFMKNCMLHYAAYRN
PgBAS2	716	GDFPQQEITGVFMKNCMLHYAASRN
MiCAS	720	GDFPQGEIMGVFNKNCMISYSAYRN
LaCAS	723	GDFPQEEIMGIENKNCMISYAAYRN
BpCAS1	725	GDFPQEEIMGVFNKNCMISYSAYRN
BpCAS2	715	GDFPQEEIMGVFNRCMITYAAYRN
KdCAS	716	GDFPQQEIMGVFNKNCMISYAAYRN
BrCAS1	715	GDFPQQEIMGVFNRCMITYAAYRN
BrCAS2	715	GDFPQQEIMGVFNRCMITYAAYRN
AtCAS1	624	GDFPQQEIMGVFNRCMITYAAYRN
AtCAS2	715	GDFPQQEIMGVFNRCMITYAAYRN
PgCAS	715	GDFPQEEIMGVFDKNCMITYAAYRN
PsCAS	715	GDFPQEEIMGVFNKNCMITYAAYRC
LjCAS	715	GDFPQQEIMGVFNKNCMITYAAYRS
GgCAS	715	GDFPQQEIMGVFNKNCMITYAAYRN

Figura 22 – Representação do MSA com destaque para a posição 729. Imagem fornecida pelo autor.

A partir do alinhamento, verifica-se que as 3-ceto OSCs possuem um resíduo de lisina (K) conservado nessa posição, em contraste com a metionina (M) presente

nas enzimas de β -amirina sintase e lupeol sintase e da asparagina (N) presente nas enzimas de cicloartenol sintase.

Acredita-se que esse resíduo (e os vizinhos) estejam relacionados com o tamanho da cavidade do sítio ativo. Hoshino e colaboradores (2017) relatam que o resíduo de metionina (M729) na β -amirina sintase de *Euphorbia tirucalli* está posicionado próximo do anel E e seu volume estérico é responsável por manter o substrato na orientação correta permitindo a expansão desse anel.

Chen e colaboradores (2023) verificaram que mutações na oxidoesqualeno ciclase de *Astragalus membranaceus*, adicionando os resíduos de valina-fenilalanina-metionina (VFM), resultaram em um aumento significativo da produção de β -amirina.

Por fim, é relatado na literatura que o resíduo adjacente de fenilalanina (F728) promove interações cátion- π que estabilizam o cátion oleanila (ITO et al., 2013).

A partir do modelo construído para MiFRS, foi possível verificar a que a lisina (K729) interage com resíduos aromáticos como Y259 (do *motif* conservado MWCYCR) e Y609 (Figura 23).

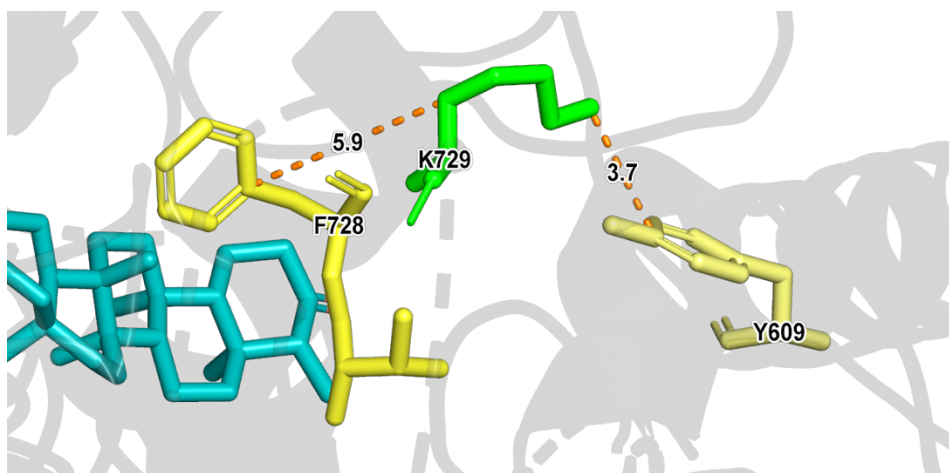


Figura 23 – Interação entre os resíduos K729 e Y609. Imagem fornecida pelo autor.

Como a lisina (K) é um resíduo que diferencia as 3-ceto OSCs das demais, optou-se por realizar mutações nessa posição. Propusemos a troca desse resíduo para os aminoácidos metionina (K729M), arginina (K729R), triptofano (K729W), alanina (K729A) e ácido glutâmico (K729E), explorando diferentes características físico-químicas e estéricas e avaliando a ação desses resíduos na capacidade produtiva da enzima (Figura 24).

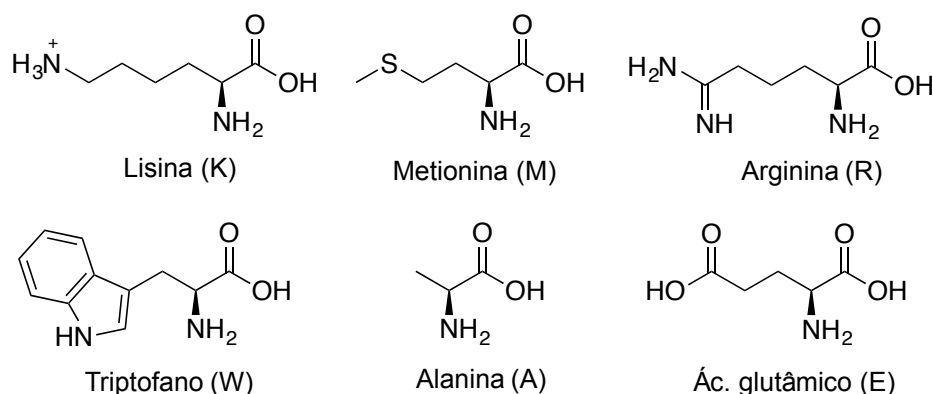


Figura 24 – Aminoácidos utilizados nos estudos de mutação sítio-dirigida da posição 729. Fonte: Imagem fornecida pelo autor.

A Tabela 18 resume as mutações realizadas para ambas as enzimas de friedelina sintase.

Tabela 18 – Mutações propostas para a posição 729.

Aminoácido	Mutação	Característica físico-química	Justificativa
K729	M	Polar não carregado	Adicionar o resíduo presente nas demais OSCs.
	R	Polar carregado positivamente	Adicionar um resíduo com a mesma capacidade de realizar interações cátion- π .
	W	Polar aromático	Adicionar um aminoácido volumoso.
	A	Apolar alifático	Eliminar a interação cátion- π e adicionar um grupo com menor impedimento estérico nessa região da cavidade do sítio-ativo.
	E	Polar carregado negativamente	Verificar como um aminoácido carregado negativamente altera a estabilidade da enzima.

As reações de PCR para a obtenção dos mutantes foram realizadas para ambas as enzimas e os nucleotídeos alterados são apresentados na Tabela 19.

Tabela 19 – Primers utilizados para a realização da PCR de mutação sítio-dirigida na posição 729.

Gene	Alteração
<i>MiFRS</i> – K729M	K → M AAG por ATG
<i>MiFRS</i> – K729R	K → R AAG por AGA
<i>MiFRS</i> – K729W	K → W AAG por TGG
<i>MiFRS</i> – K729A	K → A AAG por GCT
<i>MiFRS</i> – K729E	K → E AAG por GAA
<i>PdFRS</i> – K726M	K → M AAG por ATG
<i>PdFRS</i> – K726R	K → R AAG por AGA
<i>PdFRS</i> – K726W	K → W AAG por TGG
<i>PdFRS</i> – K726A	K → A AAG por GCT

Posição 609: A tirosina (Y) na posição 609 é altamente conservada nas OSCs de plantas (Figura 25).

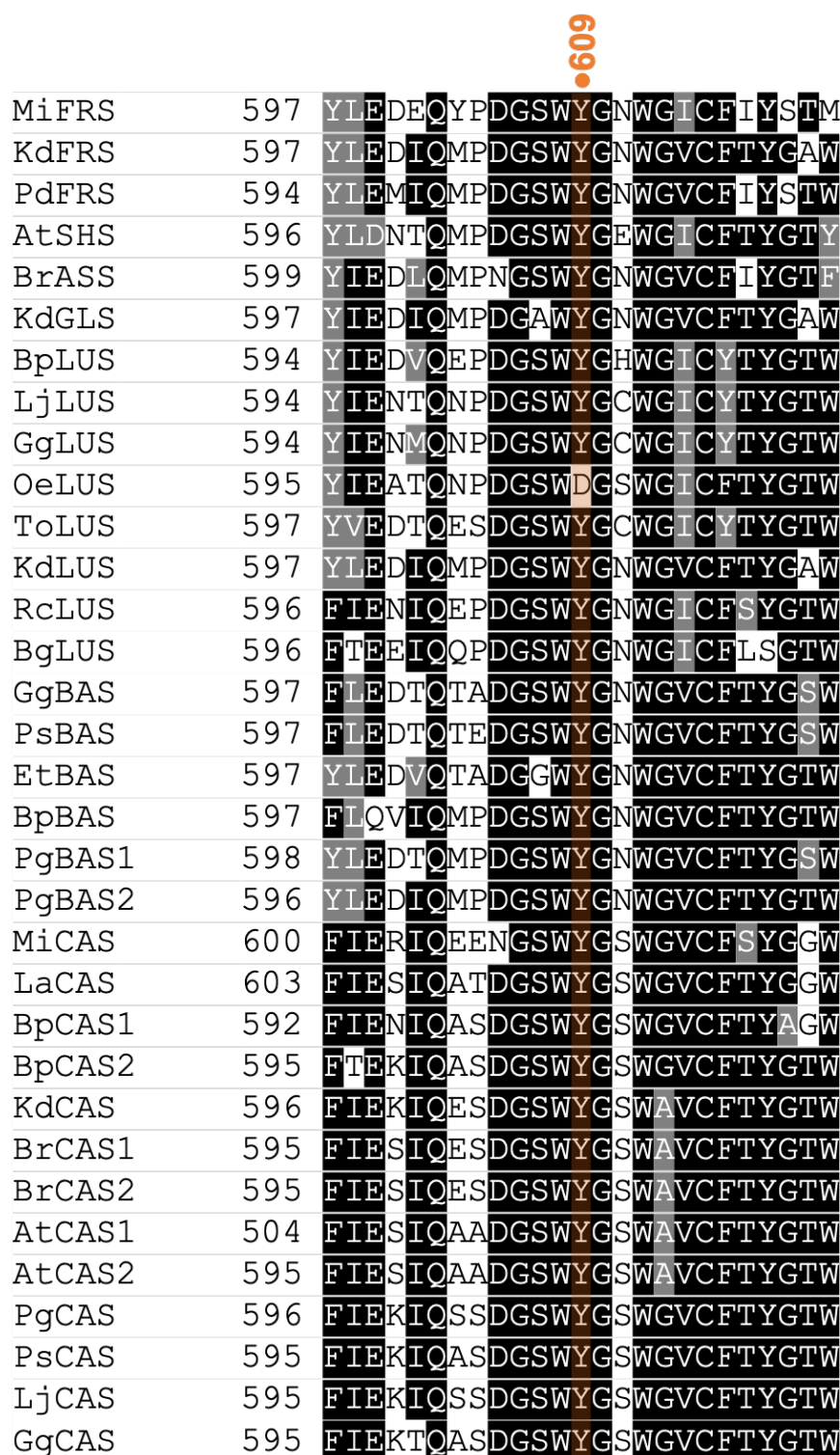


Figura 25 – Representação do MSA com destaque para a posição 609. Fonte: Imagem fornecida pelo autor.

Estudos na literatura descrevem a importância de aminoácidos aromáticos para a estabilização do carbocátion da cascata de rearranjos a partir de interações do tipo cátion- π (DITCHFIELD; SPENCER, 2016). A partir do modelo construído

para a enzima *MiFRS*, verifica-se que esse resíduo está a 3.7 Å de distância da lisina (K729), sendo localizado em uma região importante para a cavidade do sítio ativo (Figura 26).

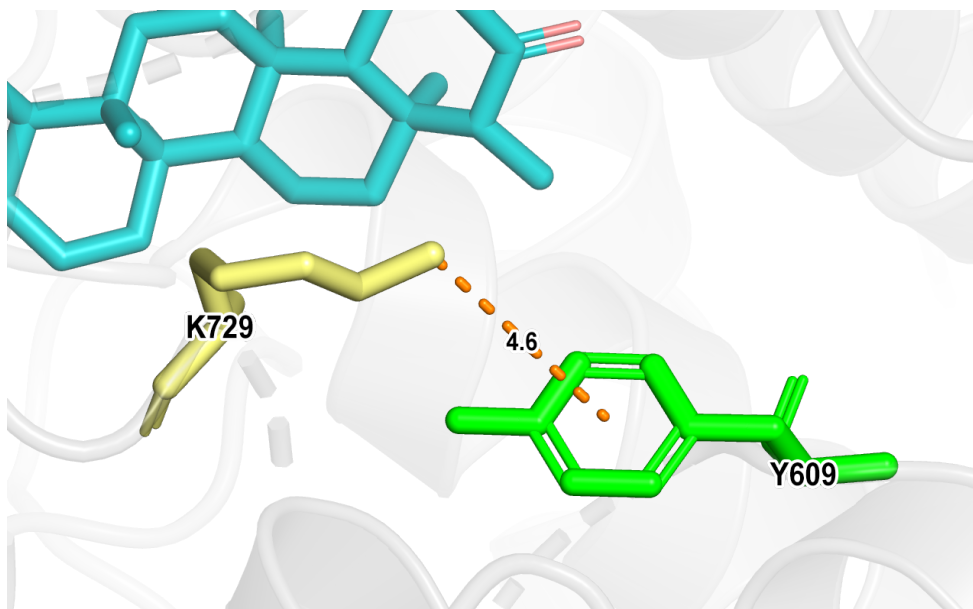


Figura 26 – Interação entre os resíduos Y609, W608 e K729. Fonte: Imagem fornecida pelo autor.

Por esse motivo, estima-se que esses resíduos interajam entre si e propôs-se a alteração da tirosina para alanina (Figura 27). O objetivo dessa mudança é eliminar uma possível interação cátion- π e avaliar como a presença de um aminoácido menos impedido estericamente afeta essa região (Tabela 20).

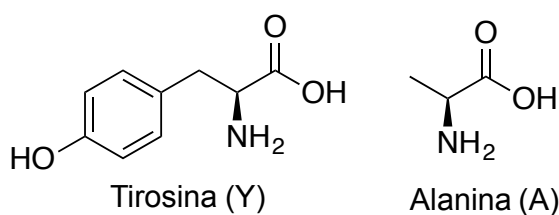


Figura 27 – Aminoácidos utilizados nos estudos de mutação sítio-dirigida para a posição 609. Fonte: Imagem fornecida pelo autor.

Tabela 20 – Mutações propostas para a posição 609.

Aminoácido	Mutação	Característica físico-química	Justificativa
Y609	A	Apolar alifático	Eliminar a interação cátion- π e adicionar um grupo com menor impedimento estérico nessa região da cavidade do sítio-ativo.

As reações de mutação sítio-dirigidas foram realizadas (item 4.2.2) e os nucleotídeos alterados para ambas as enzimas são apresentados na Tabela 21.

Tabela 21 – Primers utilizados para a realização da PCR de mutação sítio-dirigida na posição 609.

Gene	Alteração
<i>MiFRS</i> – Y609A	Y → A TAC por GCT
<i>PdFRS</i> – Y606A	Y → A TAT por GCT

As modificações sugeridas nesse trabalho foram realizadas para ambas as enzimas e as leveduras mutantes foram cultivadas e expressas. Os triterpenos produzidos foram avaliados de maneira qualitativa por GC-MS e a quantidade de friedelina produzida foi quantificada através da utilização de uma curva padrão.

5.3 Expressão dos mutantes em sistema heterólogo e quantificação da friedelina

No total, 17 alterações de resíduos foram realizadas nas enzimas *MiFRS* e *PdFRS*. Após as reações de PCR de mutação sítio-dirigida (item 4.2.2), os DNAs mutantes foram sequenciados (item 4.1.4).

Em seguida, os plasmídeos contendo o DNA mutante foram transformados em *S. cerevisiae* (item 4.1.5), sendo essa transformação confirmada a partir de PCR de colônia (item 4.1.6). Com esse experimento, foi possível verificar a presença do gene de *MiFRS* e do gene de *PdFRS* nas leveduras selvagens e mutantes a partir dos amplicons em 1208 bp e 1013 bp, respectivamente (Figura 28).

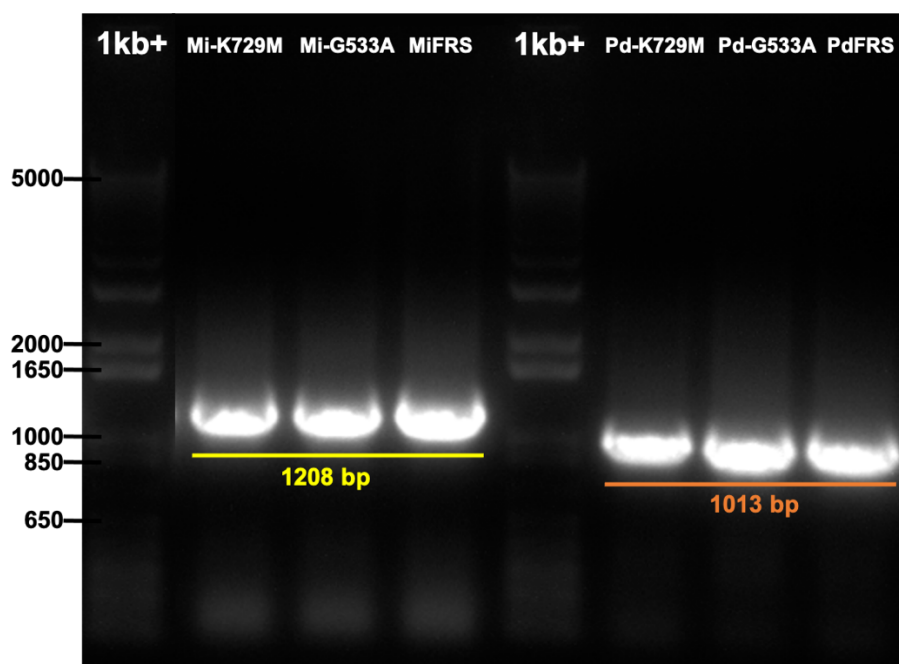


Figura 28 – PCR de colônia do DNA selvagem e mutante após a transformação em levedura. Fonte: Imagem fornecida pelo autor.

Após, as colônias foram expressas como descrito no item 4.2.3, sendo a friedelina extraída e quantificada por GC-MS (item 4.2.4). Desse modo, foi possível analisar a influência das mutações para a produção global de friedelina.

5.3.1. Mutações que não produziram friedelina

Posição 729: A Figura 29 mostra os cromatogramas obtidos por GC-MS para a enzima selvagem e os mutantes de *MiFRS*.

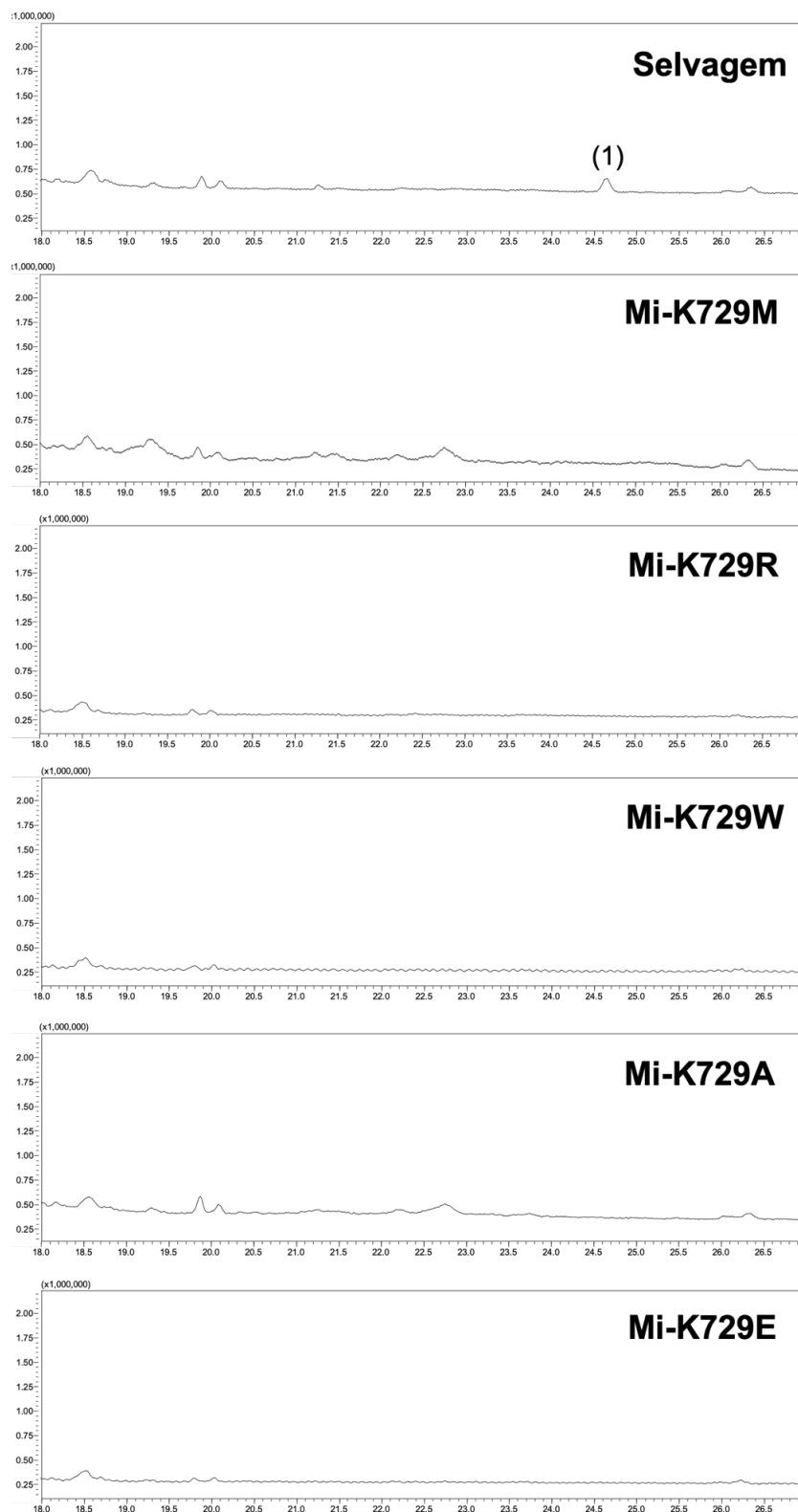


Figura 29 – Expressão da enzima *MiFRS* e dos mutantes da posição 729. A friedelina (1) foi identificada a partir da biblioteca NIST14. Fonte: Imagem fornecida pelo autor.

Em comparação com a enzima selvagem, foi possível verificar que os mutantes de *MiFRS* não resultaram na produção de friedelina (1). O mesmo foi observado para a enzima selvagem e os mutantes de *PdFRS* (Figura 30).

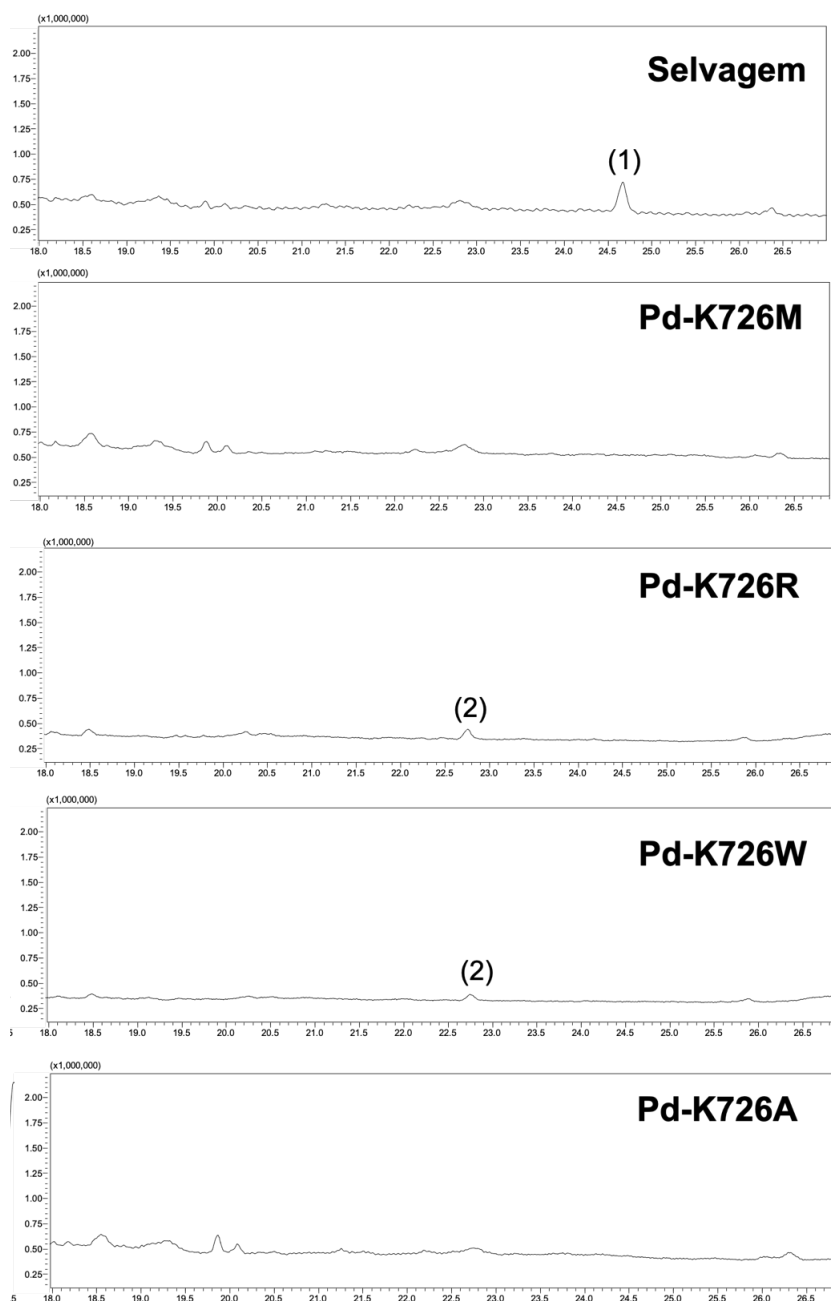


Figura 30 – Expressão da enzima *PdFRS* e dos mutantes da posição 729. Fonte: Imagem fornecida pelo autor.

Esses resultados demonstram que o tamanho e característica físico-química desse resíduo (K) são importantes para a capacidade da enzima em realizar a ciclização do 2,3-oxidoesqualeno. Como estudos relatam que os aminoácidos aromáticos vizinhos estão envolvidos nas últimas etapas de ciclização dos

triterpenos (expansão do anel E), acredita-se que a lisina atue na estabilização e orientação desses resíduos próximos através de interações cátion- π , garantindo que os mesmos estejam na distância e na conformação ideal para interagirem com o substrato e realizar a ciclização.

Apesar da friedelina não ser produzida para os mutantes, um sinal com $T_R = 22,7$ min foi observado para *Pd-K726R* e *Pd-K726W*. Através do espectro de massas desse pico (anexo D-2), pode-se propor que o mesmo é um triterpeno pentacíclico resultante de uma etapa anterior da cascata de rearranjos. Desse modo, pode-se inferir que a troca da lisina (K) pela arginina (R), aminoácido carregado positivamente, e pelo triptofano (W), aminoácido aromático polar, permitiu a estabilização de uma etapa anterior da cascata de rearranjos, alterando o produto resultante dessa ciclização.

5.3.2. Mutações que produziram friedelina

Posição 416: A Figura 31 mostra o cromatograma obtido por GC-MS para a enzima selvagem de *MiFRS* e os mutantes da posição 416.

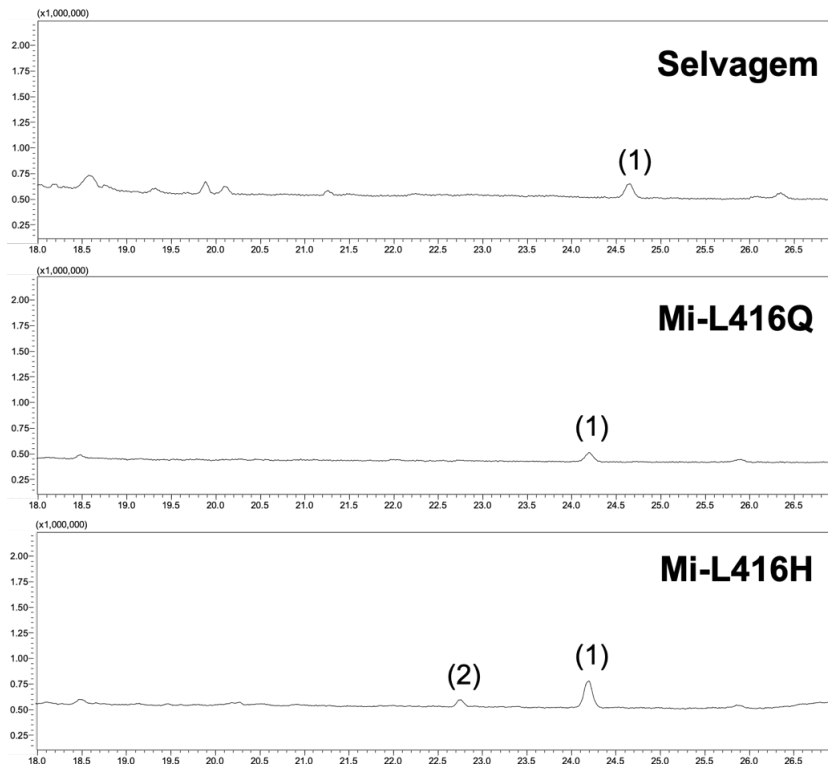


Figura 31 – Expressão da enzima *MiFRS* e os mutantes *Mi-L416Q* e *Mi-L416H* por GC-MS. Fonte: Imagem fornecida pelo autor.

É possível verificar que os mutantes de *MiFRS* produziram friedelina (1) pelo sinal observado em 24,2 min. O mutante *Mi-L416H* também apresentou um sinal com $T_R = 22,7$ min (2). Para os mutantes da enzima *PdFRS*, os resultados estão apresentados na Figura 32.

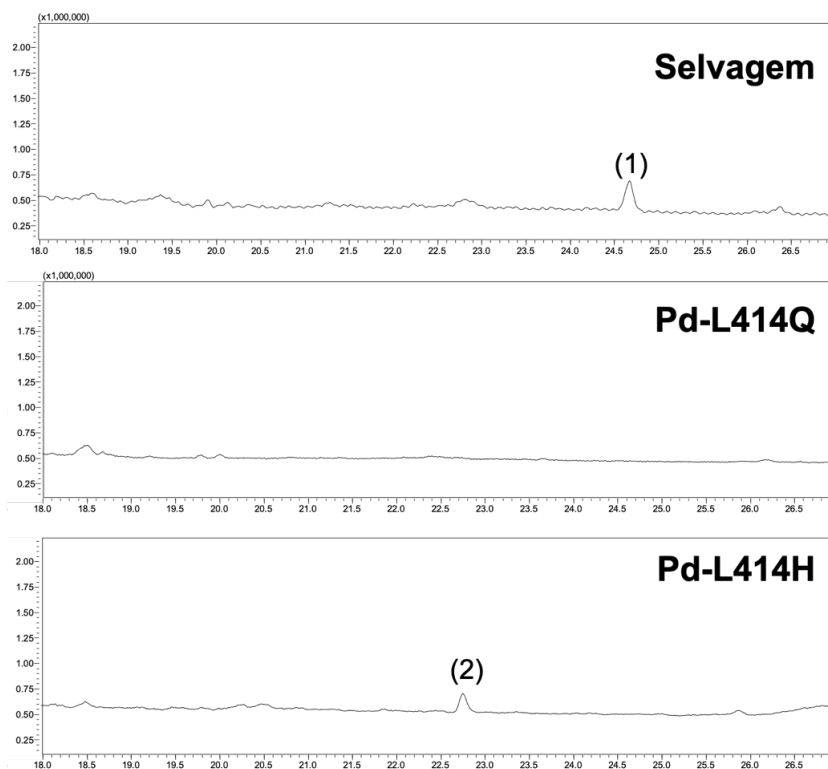


Figura 32 – Expressão da enzima *PdFRS* e os mutantes *Pd-L414Q* e *Pd-L414H* por GC-MS. Fonte: Imagem fornecida pelo autor.

Foi possível observar que os mutantes *Pd-L414Q* e *Pd-L414H* não produziram friedelina. Apesar disso, o mutante *Pd-L414H* apresentou um sinal em 22,7 min (2).

A partir dos dados apresentados, verifica-se que as mutações nessa posição levaram a resultados mistos. Quando a modificação foi realizada na enzima *MiFRS*, observou-se que a mesma continuou a produzir friedelina em níveis quantificáveis. Já para os mutantes de *PdFRS*, esse triterpeno não foi produzido.

Essa diferença de atividade entre as enzimas de friedelina sintase pode estar relacionada com o enovelamento das mesmas e com a influência de resíduos espacialmente próximos e a maneira como eles interagem com o resíduo da posição 416. Estudos mais aprimorados de modelagem molecular devem ser realizados para explicar essa diferença produtiva.

Em ambas as enzimas, a mutação para um resíduo positivamente carregado de histidina (H) levou a produção de um triterpeno não-identificado pela base de dados NIST14. Assim como visto nas mutações da posição 729, acredita-se que esse sinal seja referente a um triterpeno oriundo de etapas anteriores da ciclização. A adição da histidina, um resíduo carregado positivamente, possibilitou a presença de interações cátion- π e de empilhamento-T com os aminoácidos aromáticos vizinhos, o que pode explicar a perda da especificidade da enzima e a produção desse pico.

A quantidade de friedelina produzida nos mutantes foi posteriormente quantificada e comparada com o teor produzido pela enzima selvagem.

Posição 533: A Figura 33 mostra os cromatogramas obtido na análise da enzima selvagem *MiFRS* e do mutante *Mi-G533A*.

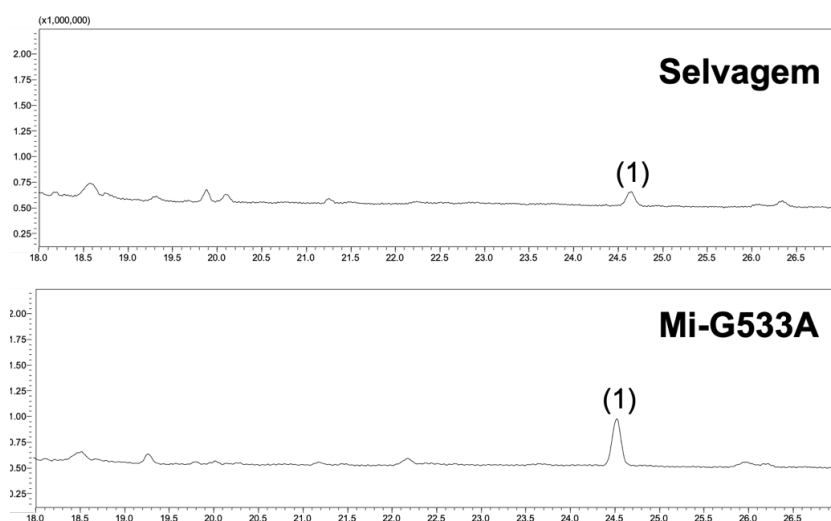


Figura 33 – Expressão da enzima *MiFRS* e o mutante *Mi-G533A*. Fonte: Imagem fornecida pelo autor.

Foi possível observar, de modo qualitativo, um aumento na quantidade de friedelina (1) produzida. O mesmo resultado pôde ser visto para a análise da expressão da enzima *PdFRS* e do mutante *Pd-G530A* (Figura 34).

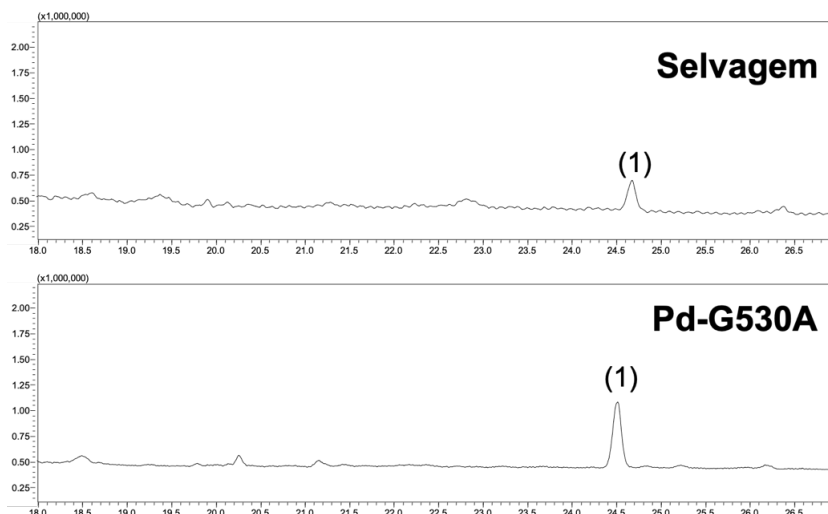


Figura 34 – Expressão da enzima *PdFRS* e o mutante *Pd-G530A*. Fonte: Imagem fornecida pelo autor.

Desse modo, verificou-se que a adição do resíduo de alanina possibilitou um aumento global na produção do composto de interesse. Como visto na Figura 20, esse resíduo apresenta-se espacialmente próximo da posição C-3 e da região do sítio-ativo.

Alterações da glicina por aminoácidos com diferentes volumes estéricos nessa região poderiam acarretar em uma desestabilização da estrutura da enzima e consequentemente na perda da atividade, como descrito na literatura para os resíduos vizinhos (L482 e W534). Portanto, a adição do resíduo de alanina não provocou grandes alterações conformacionais e apresentou uma maior especificidade entre a enzima e o substrato, oriunda da natureza doadora de elétrons das metilas da alanina.

Para confirmar que a quantidade produzida pelo mutante é significativamente diferente do obtido para a enzima selvagem, realizou-se a quantificação.

Posição 609: A Figura 35 apresenta os cromatogramas obtidos para a enzima selvagem *MiFRS* e para o mutante *Mi-Y609A*.

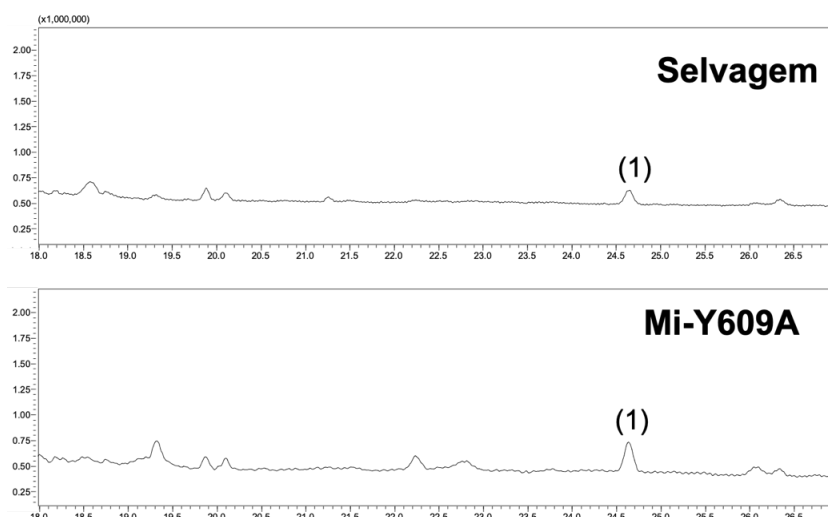


Figura 35 – Expressão da enzima *MiFRS* e o mutante *Mi-Y609A*. Fonte: Imagem fornecida pelo autor.

Foi possível verificar que o mesmo perfil cromatográfico. O mesmo resultado foi observado para o mutante *Pd-Y606A* em comparação com a enzima selvagem *PdFRS* (Figura 36).

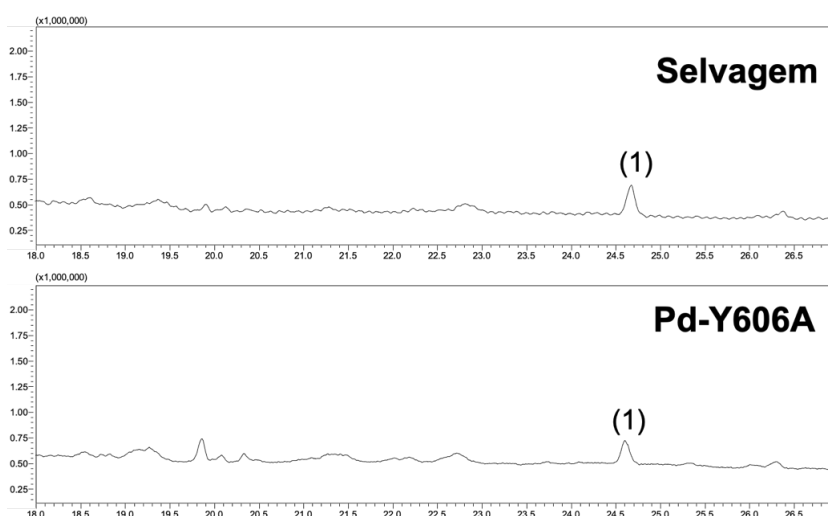


Figura 36 – Expressão da enzima *PdFRS* e o mutante *Pd-Y606A*. Fonte: Imagem fornecida pelo autor.

Para essa mutação, mesmo alterando a propriedade físico-química do aminoácido, não foi possível verificar mudanças significativas na expressão da enzima. Isso indica que esse resíduo, apesar de interagir com a lisina (K729), não possui influência direta na especificidade e na capacidade produtiva. Para confirmar essa informação, essa levedura mutante foi selecionada para quantificação.

5.4 Quantificação dos mutantes ativos

As enzimas que se apresentaram ativas foram selecionadas e a quantidade de friedelina produzida na enzima selvagem e nos mutantes foi quantificada por GC-MS. Com isso, buscou-se identificar o efeito das alterações na capacidade produtiva e realizar comparações intra e inter-enzimáticas.

5.4.1 Construção da curva analítica

Para a quantificação, uma curva analítica utilizando o padrão de friedelina foi construída a partir de diluições sucessivas de uma solução mãe com concentração de 1 mg/mL.

Conforme descrito no item 4.2.4, a análise foi realizada utilizando um padrão interno de colesterol na concentração final de 40 µg/mL. A escolha por esse padrão foi realizada pelo fato do mesmo ser estruturalmente semelhante aos triterpenos, sem apresentar co-eluição no método testado.

As análises dos pontos da curva analítica foram realizadas por GC-MS a partir do monitoramento de íon selecionado (SIM), sendo o eixo Y referente à razão entre as áreas de friedelina e do padrão interno, enquanto o eixo X corresponde à concentração de friedelina (Figura 37). A curva apresentou excelentes resultados de linearidade, o que é expresso pelo coeficiente de correlação (R^2) de 0,996.

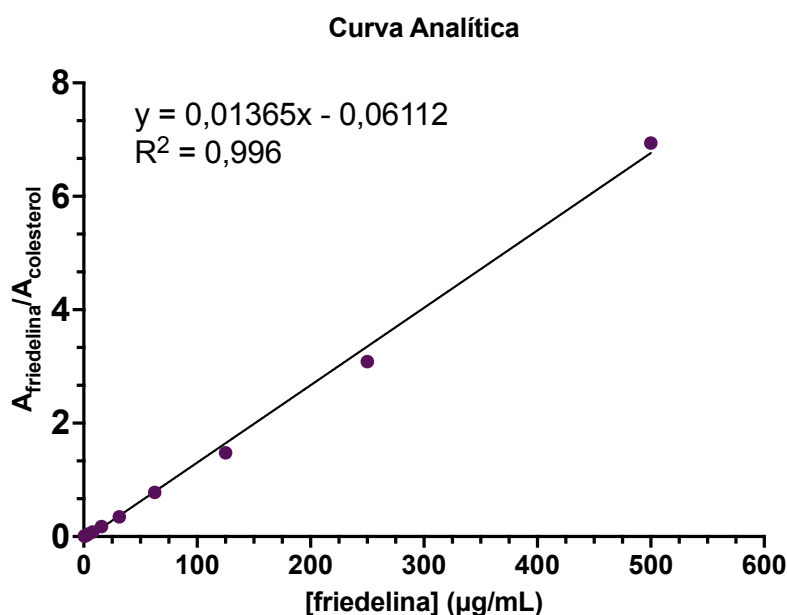


Figura 37 – Curva analítica obtida para o padrão de friedelina. Fonte: Imagem fornecida pelo autor.

Para a quantificação do teor de friedelina produzida nas leveduras selvagens e mutantes, as mesmas foram analisadas pela mesma metodologia e os dados interpolados na curva, sendo os resultados expressos em mg de friedelina por litro de meio de cultura.

5.4.2 Quantificação da friedelina produzida pelas leveduras

Para a determinação da produção de friedelina nas enzimas selvagens e mutantes, cada levedura foi cultivada em quadruplicata biológica e expressa seguindo as metodologias apresentadas em 4.2.3 e 4.2.4. Após a extração, as mesmas foram analisadas por GC-MS e a concentração da substância em cada amostra foi obtida a partir da equação da regressão linear apresentada.

A Tabela 22 resume os valores de concentração de friedelina em mg/L de meio de cultura obtidos para todos os mutantes ativos.

Tabela 22 – Concentração de friedelina em mg/L de meio de cultura obtida para as amostras após interpolação dos dados na curva analítica. n = 4.

Enzima	Média $A_{\text{friedelina}}/A_{\text{colesterol}}$	Concentração de friedelina (mg/L)	Desvio Padrão
<i>MiFRS</i> (selvagem)	67,473	2,024	0,169
<i>PdFRS</i> (selvagem)	34,149	1,024	0,072
<i>MiFRS</i> - L416Q	67,026	2,011	0,242
<i>MiFRS</i> - L416H	36,254	1,088	0,121
<i>MiFRS</i> - G533A	121,0860	3,633	0,203
<i>PdFRS</i> - G530A	51,161	1,535	0,169
<i>MiFRS</i> - Y609A	22,556	0,677	0,072
<i>PdFRS</i> - Y606A	40,160	1,205	0,070

O valor de friedelina em mg/L obtido para a enzima selvagem *MiFRS* (2,024 $\pm$ 0,169 mg/L) está de acordo com o valor recentemente descrito por Souza-Moreira e colaboradores (2018) para a mesma cepa (1,6 $\pm$ 0,6 mg/L). Além disso, os valores pequenos de desvio padrão obtidos indicam a eficiência da utilização do plasmídeo induzido por galactose pESC-URA para a reprodutibilidade da expressão.

A maior produção de friedelina foi obtida para a mutação *MiFRS* – G533A, com 3,633 ($\pm$ 0,203) mg/L, um valor 1,8x maior do que o da enzima selvagem e 2,4x maior que o obtido para o mesmo mutante em *PdFRS*. A partir do gráfico de barras

apresentado na Figura 38, é possível comparar os pares de mutantes e o comportamento que as alterações resultaram para ambas as enzimas.

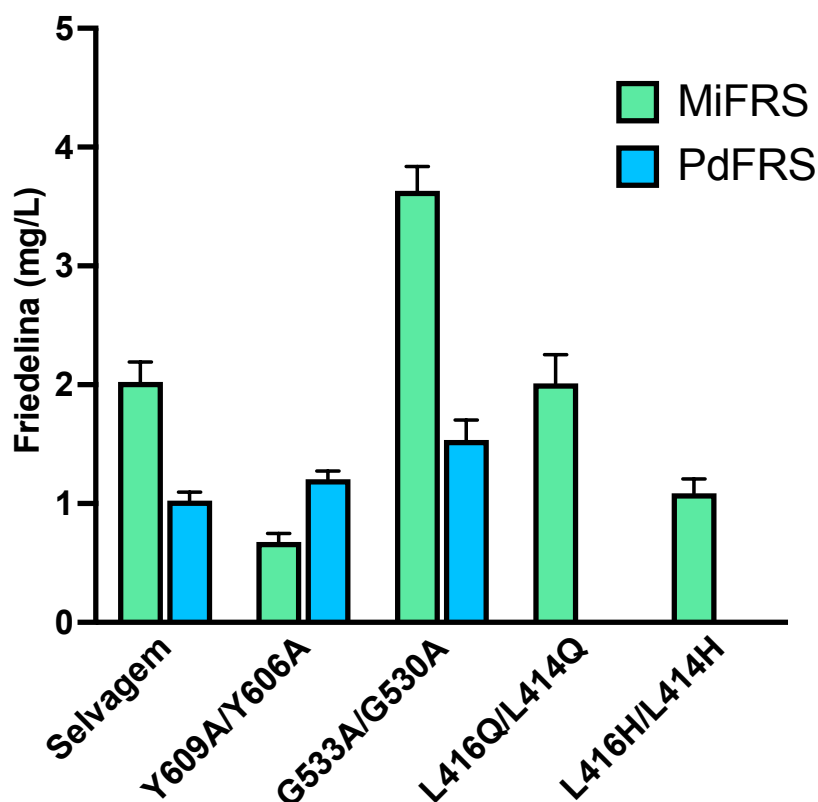


Figura 38 – Comparação da concentração de friedelina em mg/L para os pares de genes selvagens e mutantes. Fonte: Imagem fornecida pelo autor.

No geral, é possível observar que a enzima *MiFRS* é mais ativa que a enzima *PdFRS*, resultando em um maior teor de friedelina em todos os mutantes, com exceção do par *Mi-Y609A/Pd-Y606A*. Nesse caso, verifica-se que a presença da alanina interferiu de maneira negativa na produção enzimática de *MiFRS*, de modo que estudos posteriores devem ser realizados para explicar essa diferenciação entre as enzimas.

Os mutantes *Mi-G533A* e *Pd-G530A* resultaram em um aumento global na produção de friedelina, o que pode ser explicado pela capacidade do resíduo de alanina doar densidade eletrônica e estabilizar o carbocátion e os resíduos vizinhos sem grandes alterações estéricas.

Por fim, os mutantes *Mi-L416Q* e *Mi-L416H* apresentaram uma produção comparável à da enzima selvagem, mesmo com a perda da especificidade do mutante *Mi-L416H* que produziu um outro triterpeno. Essas mesmas modificações

levaram as enzimas correspondentes *Pd-L414Q* e *Pd-L414H* a não possuírem atividade.

Para determinar se existe diferença significativa entre os resultados, foi realizada uma análise da variância (ANOVA *one-way*) seguida de teste post-hoc de Tukey a 95% de confiança para a comparação das médias das replicatas. Os resultados obtidos para as enzimas e os mutantes de *MiFRS* são apresentados na Figura 39.

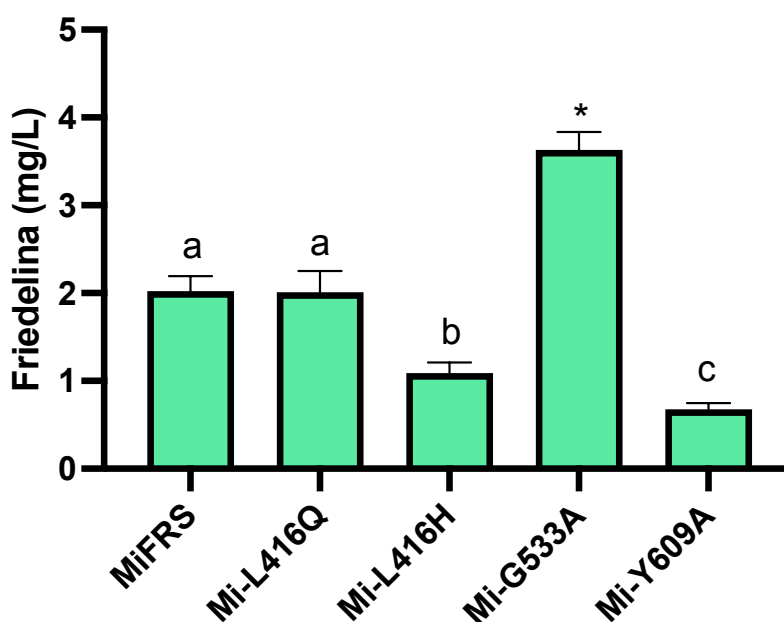


Figura 39 – Comparação da concentração de friedelina em mg/L para o gene selvagem e os mutantes de *MiFRS*. A presença de diferença significativa é indicada pela presença de um asterisco (*), enquanto barras com a mesma letra indicam a falta de diferença estatística ($p > 0,05$, $n = 4$).

Fonte: Imagem fornecida pelo autor.

A partir do resultado apresentado, é possível verificar a diferença significativa do mutante *Mi-G533A* em comparação com os demais. É comprovado estatisticamente que o mutante *Mi-L416Q* não altera a capacidade produtiva da enzima, e que os mutantes *Mi-L416H* e *Mi-Y609A* resultam em diminuição da quantidade de friedelina.

Essa análise estatística também foi realizada para a enzima selvagem e os mutantes de *PdFRS* (Figura 40), onde foi possível identificar diferença significativa somente para o mutante *Pd-G530A*.

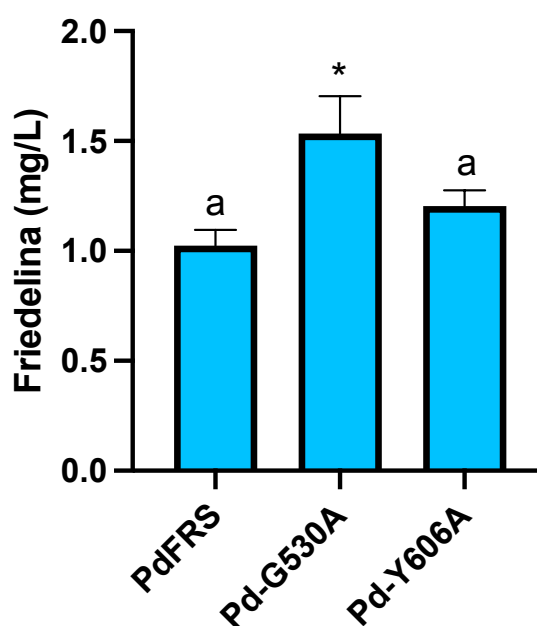


Figura 40 – Comparação da concentração de friedelina em mg/L para o gene selvagem e os mutantes de *PdFRS*. A presença de diferença significativa é indicada pela presença de um asterisco (*), enquanto barras com a mesma letra indicam a falta de diferença estatística ($p > 0,05$, $n = 4$).

Fonte: Imagem fornecida pelo autor.

Desse modo, a partir da quantificação e da análise estatística, foi possível determinar a interferência dos mutantes na produção das enzimas *MiFRS* e *PdFRS*, fornecendo informações sobre a importância desses resíduos para a atividade das enzimas.

6 CONCLUSÕES

Com o presente trabalho foi possível avaliar as enzimas friedelina sintase de *M. ilicifolia* e *P. davidiana* quanto à especificidade e atividade a partir da realização de reações de mutação sítio-dirigida.

A partir dos estudos aqui apresentados, foi possível verificar que o resíduo de lisina na posição 729 mostra-se essencial para a atividade da friedelina sintase. Também foi possível concluir que os resíduos 416 e 609 apresentam resultados mistos entre as enzimas *MiFRS* e *PdFRS*.

Foi possível descrever um mutante que resultou em um aumento significativo da capacidade produtiva da friedelina sintase, sendo a alteração da glicina pela alanina na posição 533 não descrita até o momento na literatura.

Esses resultados corroboram e complementam os dados apresentados até o momento, com a realização de mutações em posições ainda não analisadas e com um comparativo inter-enzimático.

Desta forma, esse trabalho mostrou-se importante por fornecer informações sobre a estabilidade e capacidade produtiva dessas enzimas, resultando identificação do mutante mais produtivo obtido até o momento por nosso grupo de pesquisa, com a produção de 3,63 mg de friedelina por litro de meio de cultura.

As informações aqui apresentadas sobre a enzima podem auxiliar estudos futuros sobre as próximas etapas da via biossintética e fornecer um ponto de partida para estudos focados em um aumento na capacidade produtiva.

7 REFERÊNCIAS

- AJIKUMAR, P. K.; XIAO, W.-H.; TYO, K. E. J.; WANG, Y.; SIMEON, F.; LEONARD, E.; MUCHA, O.; PHON, T. H.; PFEIFER, B.; STEPHANOPOULOS, G. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. **Science**, v. 330, n. 6000, p. 70–74, 2010.
- ALVES, T.; SOUZA-MOREIRA, T.; VALENTINI, S.; ZANELLI, C.; FURLAN, M. Friedelin in *Maytenus ilicifolia* Is Produced by Friedelin Synthase Isoforms. **Molecules**, v. 23, n. 3, p. 700, 2018.
- ANTONISAMY, P.; DURAI PANDIYAN, V.; IGNACIMUTHU, S. Anti-inflammatory, analgesic and antipyretic effects of friedelin isolated from *Azima tetraacantha* Lam. in mouse and rat models. **Journal of Pharmacy and Pharmacology**, v. 63, n. 8, p. 1070–1077, 2011.
- ANTUNES, E. R. M.; DUARTE, R. S.; MORITZ, T.; SAWAYA, A. C. H. F. Differentiation of two *Maytenus* species and their hybrid via untargeted metabolomics. **Industrial Crops and Products**, v. 158, n. 4, p. 1-9, 2020.
- BAGGIO, C. H.; FREITAS, C. S.; OTOFUJI, G. DE M.; CIPRIANI, T. R.; SOUZA, L. M. DE; SASSAKI, G. L.; IACOMINI, M.; MARQUES, M. C. A.; MESIA-VELA, S. Flavonoid-rich fraction of *Maytenus ilicifolia* Mart. ex. Reiss protects the gastric mucosa of rodents through inhibition of both H⁺,K⁺-ATPase activity and formation of nitric oxide. **Journal of Ethnopharmacology**, v. 113, n. 3, p. 433–440, 2007.
- BICALHO, K. U.; SANTONI, M. M.; ARENDT, P.; ZANELLI, C. F.; FURLAN, M.; GOOSSENS, A.; POLLIER, J. CYP712K4 Catalyzes the C-29 Oxidation of Friedelin in the *Maytenus ilicifolia* Quinone Methide Triterpenoid Biosynthesis Pathway. **Plant and Cell Physiology**, v. 60, n. 11, p. 2510–2522, 2019.
- BIRAL, L.; SIMMONS, M. P.; TEMBROCK, L. R.; BOLSON, M.; ARCHER, R. H.; LOMBARDI, J. A. Systematics of New World *Maytenus* (Celastraceae) and a New Delimitation of the Genus. **Systematic Botany**, v. 42, n. 4, p. 1–14, 2017.
- BIRCHFIELD, A. S.; MCINTOSH, C. A. Metabolic engineering and synthetic biology of plant natural products – A minireview. **Current Plant Biology**, v. 24, n. May, p. 100163, 2020.
- BRINKER, A. M.; MA, J.; LIPSKY, P. E.; RASKIN, I. Medicinal chemistry and pharmacology of genus *Tripterygium* (Celastraceae). **Phytochemistry**, v. 68, n. 6, p. 732–766, 2007.
- BRÖKER, J. N.; MÜLLER, B.; PRÜFER, D.; SCHULZE GRONOVER, C. Combinatorial Metabolic Engineering in *Saccharomyces cerevisiae* for the Enhanced Production of the FPP-Derived Sesquiterpene Germacrene. **Bioengineering**, v. 7, n. 4, p. 135, 2020.
- CALDART, E. T.; MATA, H.; CANAL, C. W.; RAVAZZOLO, A. P. Phylogenetic analysis: Basic concepts and its use as a tool for virology and molecular epidemiology. **Acta Scientiae Veterinariae**, v. 44, n. 1, p. 1–20, 2018.

- CAO, Y.-D.; HE, Y.-C.; LI, H.; KAI, G.-Y.; XU, J.-H.; YU, H.-L. Efficient biosynthesis of rare natural product scopolamine using *E. coli* cells expressing a S14P/K97A mutant of hyoscyamine 6 β -hydroxylase AaH6H. **Journal of Biotechnology**, v. 211, p. 123–129, 2015.
- CHANG, B.; JIN, C.; ZHANG, W.; KONG, L.; YANG, J.-H.; LIAN, F.-M.; LI, Q.-F.; YU, B.; LIU, W.-K.; YANG, L.-L.; ZHAO, P.; ZHEN, Z. *Euonymus alatus* in the Treatment of Diabetic Nephropathy in Rats. **The American Journal of Chinese Medicine**, v. 40, n. 06, p. 1177–1187, 2012.
- CHEN, K.; ZHANG, M.; YE, M.; QIAO, X. Site-directed mutagenesis and substrate compatibility to reveal the structure–function relationships of plant oxidosqualene cyclases. **Natural Product Reports**, v. 38, n. 12, p. 2261–2275, 2021.
- CHEN, K.; ZHANG, M.; XU, L.; YI, Y.; WANG, L.; WANG, H.; WANG, Z.; XING, J.; LI, P.; ZHANG, X.; SHI, X.; YE, M.; OSBOURN, A.; QIAO, X. Identification of oxidosqualene cyclases associated with saponin biosynthesis from *Astragalus membranaceus* reveals a conserved motif important for catalytic function. **Journal of Advanced Research**, v. 43, n. 1, p. 247–257, 2023.
- CORSINO, J.; CARVALHO, P. R. F.; KATO, M. J.; LATORRE, L. R.; OLIVEIRA, O. M. M. F.; ARAÚJO, A. R.; BOLZANI, V. S.; FRANÇA, S. C.; PEREIRA, A. M. S.; FURLAN, M. Biosynthesis of friedelane and quinomethide triterpenoids is compartmentalized in *Maytenus aquifolium* and *Salacia campestris*. **Phytochemistry**, v. 55, n. 7, p. 741–748, 2000.
- COURDAVAULT, V.; O’CONNOR, S. E.; JENSEN, M. K.; PAPON, N. Metabolic engineering for plant natural products biosynthesis: new procedures, concrete achievements and remaining limits. **Natural Product Reports**, v. 38, n. 12, p. 2145–2153, 2021.
- DENIS, J. N.; GREENE, A. E.; GUENARD, D.; GUERITTE-VOEGELEIN, F.; MANGATAL, L.; POTIER, P. Highly efficient, practical approach to natural taxol. **Journal of the American Chemical Society**, v. 110, n. 17, p. 5917–5919, 1988.
- DITCHFIELD, R.; SPENCER, T. A. Carbocation– π interaction: evaluation of the stabilization by phenylalanine of a biochemical carbocation intermediate. **Organic & Biomolecular Chemistry**, v. 14, n. 40, p. 9543–9548, 2016.
- EDELHEIT, O.; HANUKOGLU, A.; HANUKOGLU, I. Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. **BMC Biotechnology**, v. 9, p. 1–8, 2009.
- GIETZ, R. D.; SCHIESTL, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. **Nature Protocols**, v. 2, n. 1, p. 31–34, 2007.
- HAPPI, G. M.; NTABO, V. K.; TCHO, A. T.; WANSI, J. D. Naturally occurring dimeric triterpenoids: Occurrence, chemistry and bioactivities. **Phytochemistry**, v. 200, n. 1, p. 1–17, 2022.

HOSHINO, T. β -Amyrin biosynthesis: catalytic mechanism and substrate recognition. **Organic & Biomolecular Chemistry**, v. 15, n. 14, p. 2869–2891, 2017.

ITO, R.; HASHIMOTO, I.; MASUKAWA, Y.; HOSHINO, T. Effect of cation- π interactions and steric bulk on the catalytic action of oxidosqualene cyclase: A case study of phe728 of β -amyrin synthase from *Euphorbia tirucalli* L. **Chemistry - A European Journal**, v. 19, n. 50, p. 17150–17158, 2013.

ITO, R.; MASUKAWA, Y.; HOSHINO, T. Purification, kinetics, inhibitors and CD for recombinant β -amyrin synthase from *Euphorbia tirucalli* L. and functional analysis of the DCTA motif, which is highly conserved among oxidosqualene cyclases. **FEBS Journal**, v. 280, n. 5, p. 1267–1280, 2013.

ITO, R.; NAKADA, C.; HOSHINO, T. β -Amyrin synthase from *Euphorbia tirucalli* L. functional analyses of the highly conserved aromatic residues Phe413, Tyr259 and Trp257 disclose the importance of the appropriate steric bulk, and cation- π and CH- π interactions for the efficient. **Organic & Biomolecular Chemistry**, v. 15, n. 3, p. 717–717, 2017.

KUNTAL, B. K.; APAROY, P.; REDDANNA, P. EasyModeller: A graphical interface to MODELLER. **BMC Research Notes**, v. 3, p. 1–5, 2010.

KUSHIRO, T.; SHIBUYA, M.; EBIZUKA, Y. Chimeric Triterpene Synthase. A Possible Model for Multifunctional Triterpene Synthase. **Journal of the American Chemical Society**, v. 121, n. 6, p. 1208–1216, 1999.

KUSHIRO, T.; SHIBUYA, M.; MASUDA, K.; EBIZUKA, Y. Mutational Studies on Triterpene Synthases: Engineering Lupeol Synthase into β -Amyrin Synthase. **Journal of the American Chemical Society**, v. 122, n. 29, p. 6816–6824, 2000.

LI, W.; CUI, L.; MAI, J.; SHI, T. Q.; LIN, L.; ZHANG, Z. G.; LEDESMA-AMARO, R.; DONG, W.; JI, X. J. Advances in Metabolic Engineering Paving the Way for the Efficient Biosynthesis of Terpenes in Yeasts. **Journal of Agricultural and Food Chemistry**, v. 70, n. 30, p. 9246–9261, 2022.

MATSUMURA, E.; NAKAGAWA, A.; TOMABECHI, Y.; IKUSHIRO, S.; SAKAKI, T.; KATAYAMA, T.; YAMAMOTO, K.; KUMAGAI, H.; SATO, F.; MINAMI, H. Microbial production of novel sulphated alkaloids for drug discovery. **Scientific Reports**, v. 8, n. 1, p. 1–10, 2018.

MAZZEU, B. F.; SOUZA-MOREIRA, T. M.; OLIVEIRA, A. A.; REMLINGER, M.; FELIPPE, L. G.; VALENTINI, S. R.; GUIDO, R. V. C.; ZANELLI, C. F.; FURLAN, M. The Methionine 549 and Leucine 552 Residues of Friedelin Synthase from *Maytenus ilicifolia* Are Important for Substrate Binding Specificity. **Molecules**, v. 26, n. 22, p. 1–13, 2021.

MOSES, T.; MEHRSHAHI, P.; SMITH, A. G.; GOOSSENS, A. Synthetic biology approaches for the production of plant metabolites in unicellular organisms. **Journal of Experimental Botany**, v. 68, n. 15, p. 4057–4074, 2017.

NIERO, R.; ANDRADE, S. F.; FILHO, V. C. A Review of the Ethnopharmacology, Phytochemistry and Pharmacology of Plants of the *Maytenus* Genus. **Current Pharmaceutical Design**, v. 17, n. 18, p. 1851–1871, 2011.

- OHYAMA, K.; SUZUKI, M.; MASUDA, K.; YOSHIDA, S.; MURANAKA, T. Chemical Phenotypes of the hmg1 and hmg2 Mutants of *Arabidopsis* Demonstrate the In-plant Role of HMG-CoA Reductase in Triterpene Biosynthesis. **Chemical & Pharmaceutical Bulletin**, v. 55, n. 10, p. 1518–1521, 2007.
- OKABE, S.; OCHIAI, Y.; AIDA, M.; PARK, K.; KIM, S. J.; NOMURA, T.; SUGANUMA, M.; FUJIKI, H. Mechanistic aspects of green tea as a cancer preventive: Effect of components on human stomach cancer cell lines. **Japanese Journal of Cancer Research**, v. 90, n. 7, p. 733–739, 1999.
- OLIVEIRA, R.; CRUZ, L.; DIAS, L. Artemisinina e derivados: Descoberta, estratégias sintéticas e obtenção industrial. **Química Nova**, v. 45, n. 7, p. 831–846, 2022.
- PADDON, C. J.; KEASLING, J. D. Semi-synthetic artemisinin: A model for the use of synthetic biology in pharmaceutical development. **Nature Reviews Microbiology**, v. 12, n. 5, p. 355–367, 2014.
- PERESTELO, N. I. R. **Metabolitos secundarios aislados de *Prymnesium Maytenus jelskii* (Celastraceae). Estudio de relación estructura-actividad**. 387 f. Tese (Doutorado em Química) - Universidad de La Laguna, 2009.
- PICKENS, L. B.; TANG, Y.; CHOOI, Y.-H. Metabolic Engineering for the Production of Natural Products. **Annual Review of Chemical and Biomolecular Engineering**, v. 2, n. 1, p. 211–236, 2011.
- PYNE, M. E.; NARCROSS, L.; MARTIN, V. J. J. Engineering Plant Secondary Metabolism in Microbial Systems. **Plant Physiology**, v. 179, n. 3, p. 844–861, 2019.
- REECE-HOYES, J. S.; WALHOUT, A. J. M. Gateway Recombinational Cloning. **Cold Spring Harbor Protocols**, v. 2018, n. 1, p. 1-6, 2018.
- SAITOU, N.; NEI, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. **Molecular Biology and Evolution**, v. 4, n. 4, p. 406–425, jul. 1987.
- SALMON, M.; THIMMAPPA, R. B.; MINTO, R. E.; MELTON, R. E.; HUGHES, R. K.; O'MAILLE, P. E.; HEMMINGS, A. M.; OSBOURN, A. A conserved amino acid residue critical for product and substrate specificity in plant triterpene synthases. **Proceedings of the National Academy of Sciences**, v. 113, n. 30, p. E4407–E4414, 2016.
- SANTOS-OLIVEIRA, R.; COULAUD-CUNHA, S.; COLAÇO, W. Revisão da *Maytenus ilicifolia* Mart. ex Reissek, Celastraceae. Contribuição ao estudo das propriedades farmacológicas. **Brazilian Journal of Pharmacognosy**, v. 19, n. 2 B, p. 650–659, 2009.
- SANTOS, V. A. F. F. M.; REGASINI, L. O.; NOGUEIRA, C. R.; PASSERINI, G. D.; MARTINEZ, I.; BOLZANI, V. S.; GRAMINHA, M. A. S.; CICARELLI, R. M. B.; FURLAN, M. Antiprotozoal sesquiterpene pyridine alkaloids from *Maytenus ilicifolia*. **Journal of Natural Products**, v. 75, n. 5, p. 991–995, 2012.

SANTOS, V. A. F. F. M.; LEITE, K. M.; SIQUEIRA, M. DA C.; REGASINI, L. O.; MARTINEZ, I.; NOGUEIRA, C. T.; GALUPPO, M. K.; STOLF, B. S.; PEREIRA, A. M. S.; CICALLELLI, R. M. B.; FURLAN, M.; GRAMINHA, M. A. S. Antiprotozoal activity of quinonemethide triterpenes from *Maytenus ilicifolia* (Celastraceae). **Molecules**, v. 18, n. 1, p. 1053–1062, 2013.

SATO, T.; HOSHINO, T. Functional Analysis of the DXDDTA Motif in Squalene-Hopene Cyclase by Site-directed Mutagenesis Experiments: Initiation Site of the Polycyclization Reaction and Stabilization Site of the Carbocation Intermediate of the Initially Cyclized A-Ring. **Bioscience, Biotechnology, and Biochemistry**, v. 63, n. 12, p. 2189–2198, 1999.

SEO, H. W.; PARK, J.-H.; LEE, J. Y.; PARK, H.-J.; KIM, J.-K. Pristimerin, a Naturally Occurring Triterpenoid, Exerts Potent Anticancer Effect in Colon Cancer Cells. **Biomedical Science Letters**, v. 24, n. 1, p. 15–22, 2018.

SHEN, Y.; CHEN, B. LIAN; ZHANG, Q. XIU; ZHENG, Y. ZHONG; FU, Q. Traditional uses, secondary metabolites, and pharmacology of *Celastrus* species - a review: **Journal of Ethnopharmacology**, v. 241, n. 2025, p. 111934, 2019.

SOUZA-MOREIRA, T. M.; ALVES, T. B.; PINHEIRO, K. A.; FELIPPE, L. G.; DE LIMA, G. M. A.; WATANABE, T. F.; BARBOSA, C. C.; SANTOS, V. A. F. F. M.; LOPES, N. P.; VALENTINI, S. R.; GUIDO, R. V. C.; FURLAN, M.; ZANELLI, C. F. Friedelin Synthase from *Maytenus ilicifolia*: Leucine 482 Plays an Essential Role in the Production of the Most Rearranged Pentacyclic Triterpene. **Scientific Reports**, v. 6, n. May, p. 1–13, 2016.

SOUZA-MOREIRA, T. M.; NAVARRETE, C.; CHEN, X.; ZANELLI, C. F.; VALENTINI, S. R.; FURLAN, M.; NIELSEN, J.; KRIVORUCHKO, A. Screening of 2A peptides for polycistronic gene expression in yeast. **FEMS Yeast Research**, v. 5, n. 18, p. 1–26, 2018.

SUBASH-BABU, P.; LI, D. K.; ALSHATWI, A. A. In vitro cytotoxic potential of friedelin in human MCF-7 breast cancer cell: Regulate early expression of Cdkn2a and pRb1, neutralize mdm2-p53 amalgamation and functional stabilization of p53. **Experimental and Toxicologic Pathology**, v. 69, n. 8, p. 630–636, 2017.

TAKASE, S.; SAGA, Y.; KURIHARA, N.; NARAKI, S.; KUZE, K.; NAKATA, G.; ARAKI, T.; KUSHIRO, T. Control of the 1,2-rearrangement process by oxidosqualene cyclases during triterpene biosynthesis. **Organic & Biomolecular Chemistry**, v. 13, n. 26, p. 7331–7336, 2015.

TAMURA, K.; STECHER, G.; KUMAR, S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. **Molecular Biology and Evolution**, v. 38, n. 7, p. 3022–3027, 25 jun. 2021.

THIMMAPPA, R.; GEISLER, K.; LOUVEAU, T.; MAILLE, P. O.; OSBOURN, A. Triterpene Biosynthesis in Plants. **Annual Review of Plant Biology**, v. 65, n. January, p. 1–33, 2014.

THOMA, R.; SCHULZ-GASCH, T.; D'ARCY, B.; BENZ, J.; AEBI, J.; DEHMLow, H.; HENNIG, M.; STIHLE, M.; RUF, A. Insight into steroid scaffold formation from the structure of human oxidosqualene cyclase. **Nature**, v. 432, n. 7013, p. 118–122, 2004.

VELOSO, C. C.; SOARES, G. L.; PEREZ, A. C.; RODRIGUES, V. G.; SILVA, F. C. Pharmacological potential of *Maytenus* species and isolated constituents, especially tingenone, for treatment of painful inflammatory diseases. **Revista Brasileira de Farmacognosia**, v. 27, n. 4, p. 533–540, 2017.

WEBB, B.; SALI, A. Comparative protein structure modeling using MODELLER. **Current Protocols in Bioinformatics**, v. 5.6, n. 1, p. 1.37, 2016.

WESTFALL, P. J.; PITERA, D. J.; LENIHAN, J. R.; ENG, D.; WOOLARD, F. X.; REGENTIN, R.; HORNING, T.; TSURUTA, H.; MELIS, D. J.; OWENS, A.; FICKES, S.; DIOLA, D.; BENJAMIN, K. R.; KEASLING, J. D.; LEAVELL, M. D.; MCPHEE, D. J.; RENNINGER, N. S.; NEWMAN, J. D.; PADDON, C. J. Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. **Proceedings of the National Academy of Sciences**, v. 109, n. 3, p. 111–118, 2012.

WU, S.; ZHANG, F.; XIONG, W.; MOLNÁR, I.; LIANG, J.; JI, A.; LI, Y.; WANG, C.; WANG, S.; LIU, Z.; WU, R.; DUAN, L. An Unexpected Oxidosqualene Cyclase Active Site Architecture in the *Iris tectorum* Multifunctional α -Amyrin Synthase. **ACS Catalysis**, v. 10, n. 16, p. 9515–9520, 2020.

YANG, D.; EUN, H.; PRABOWO, C. P. S.; CHO, S.; LEE, S. Y. Metabolic and cellular engineering for the production of natural products. **Current Opinion in Biotechnology**, v. 77, p. 102760, 2022.

ZHENG, G. Q. Cytotoxic terpenoids and flavonoids from *Artemisia annua*. **Planta Medica**, v. 60, n. 1, p. 54–57, 1994.

ZHOU, J.; HU, T.; GAO, L.; SU, P.; ZHANG, Y.; ZHAO, Y.; CHEN, S.; TU, L.; SONG, Y.; WANG, X.; HUANG, L.; GAO, W. Friedelane-type triterpene cyclase in celastrol biosynthesis from *Tripterygium wilfordii* and its application for triterpenes biosynthesis in yeast. **New Phytologist**, v. 223, n. 2, p. 722–735, 2019.

ZHU, C.; COOK, S. P. A Concise Synthesis of (+)-Artemisinin. **Journal of the American Chemical Society**, v. 134, n. 33, p. 13577–13579, 2012.

ZHU, M.; WANG, C.; SUN, W.; ZHOU, A.; WANG, Y.; ZHANG, G.; ZHOU, X.; HUO, Y.; LI, C. Boosting 11-oxo- β -amyrin and glycyrrhetic acid synthesis in *Saccharomyces cerevisiae* via pairing novel oxidation and reduction system from legume plants. **Metabolic Engineering**, v. 45, p. 43–50, 2018.

ANEXOS

Anexo A - Protocolos

A-1 - Transformação de plasmídeo em bactéria

Para a realização da transformação do DNA plasmidial em bactéria, descongelou-se uma microtubo contendo a bactéria competente *Escherichia coli* DH5- α em banho de gelo e adicionou-se de 0,05-0,2 μ g de DNA purificado ao tubo.

Após, a amostra foi mantida em gelo por 30 minutos, sendo posteriormente levada a banho-maria para realização de um choque térmico (42 °C, 2 min). Realizou-se a adição de 1 mL de meio Luria-Bertani e incubação da amostra em modo dinâmico (37 °C, 1 h, 200 rpm).

O conteúdo do microtubo foi centrifugado, o sobrenadante removido e após a adição de 100 μ L de água ultrapura e ressuspensão do *pellet* celular realizou-se o plaqueamento da solução resultante em placas de Lauri-Bertani contendo o antibiótico ampicilina.

As placas foram incubadas por 16 horas em modo estático (37 °C) de modo que o aparecimento de colônias comprova o sucesso da transformação realizada.

A-2 – Drop-out utilizado para a produção de meio mínimo sem uracila (SC-URA)

O drop-out SC-URA foi produzido a partir da adição de quantidades definidas dos aminoácidos obtidos comercialmente em almofariz e trituração com auxílio de um pistilo.

- Adenina (Sigma-Aldrich®) – 18 mg;
- Ácido *p*-aminobenzoico (Sigma-Aldrich®) – 8 mg;
- Leucina (Sigma-Aldrich®) – 380 mg;
- Alanina (Sigma-Aldrich®) – 76 mg;
- Arginina (Sigma-Aldrich®) – 76 mg;
- Asparagina (Sigma-Aldrich®) – 76 mg;
- Ácido aspártico (Sigma-Aldrich®) – 76 mg;
- Cisteína (Sigma-Aldrich®) – 76 mg;
- Ácido glutâmico (Sigma-Aldrich®) – 76 mg;
- Glutamina (Sigma-Aldrich®) – 76 mg;
- Glicina (Sigma-Aldrich®) – 76 mg;
- Histidina (Sigma-Aldrich®) – 76 mg;
- *myo*-Inositol (Sigma-Aldrich®) – 76 mg;
- Isoleucina (Sigma-Aldrich®) – 76 mg;
- Lisina (Sigma-Aldrich®) – 76 mg;
- Metionina (Sigma-Aldrich®) – 76 mg;
- Fenilalanina (Sigma-Aldrich®) – 76 mg;
- Prolina (Sigma-Aldrich®) – 76 mg;
- Serina (Sigma-Aldrich®) – 76 mg;
- Treonina (Sigma-Aldrich®) – 76 mg;
- Triptofano (Sigma-Aldrich®) – 76 mg;
- Tirosina (Sigma-Aldrich®) – 76 mg;
- Valina (Sigma-Aldrich®) – 76 mg.

Anexo B – Primers utilizados nas reações de PCR

B-1 - Primers para o sequenciamento dos genes selvagens e mutantes.

Tabela 23 - Primers utilizados para o sequenciamento de toda cds dos genes selvagens e mutantes.

Gene	Primer	Sequências
MiFRS	<i>Forward (400 nctds)</i>	AAAAGCAGGCTTAATGTGG
	<i>Forward (400-800 nctds)</i>	AATCACTCCTGAACACAGAAAG
	<i>Forward (800-1200 nctds)</i>	GAGACATAGATGTGCCGAAG
	<i>Forward (1200-1600 ncts)</i>	GACAATCCTTCTGGTAACTTCG
	<i>Forward (1600-2000 nctds)</i>	ATACCCTGACGGTTCTTGG
PdFRS	<i>Forward (400 nctds)</i>	AAAAGCAGGCTTAATGTGG
	<i>Forward (400-800 nctds)</i>	TTTCCACCTGAATACAGGAAAG
	<i>Forward (800-1200 nctds)</i>	CCCTATGAGTCTGTCAAGTGG
	<i>Forward (1200-1600 ncts)</i>	AGAGTCAGGTGGTAGACAATCC
	<i>Forward (1600-2000 nctds)</i>	ATGATTCAGATGCCTGACG

B-2 - Primers utilizados para a PCR de colônia.

Tabela 24 - Primers utilizados para a PCR de colônia dos genes selvagens e mutantes.

Mutantes	Primers	Sequências
MiFRS	<i>forward</i>	GGGTCTGAAGATCCAAACGGTG
	<i>reverse</i>	GGGGACCACTTTGTACAAGAAAGCTGGGTATCAAGCACTGTTTATCTTCATAGCC
PdFRS	<i>forward</i>	CCAGCTGCTGCTAAACCTAC
	<i>reverse</i>	CCCTATGAGTCTGTCAAGTGG

B-3 - Primers para a produção de mutantes da posição 416.

Tabela 25 - Primers utilizados para a produção dos mutantes na posição 416 com destaque para os nucleotídeos alterados em relação ao gene selvagem (violeta).

Mutantes	Primers	Sequências
MiFRS - L416Q	<i>forward</i>	CTCTTTCGGTTCACAA CA ATGGGATGCTACTTTTG
	<i>reverse</i>	CAAAAGTAGCATCCC ATTG TTGTGAACCGAAAGAG
MiFRS - L416H	<i>forward</i>	CTCTTTCGGTTCACAA CATT GGGATGCTACTTTTG
	<i>reverse</i>	CAAAAGTAGCATCCCA ATG TTGTGAACCGAAAGAG
PdFRS - L414Q	<i>forward</i>	CAATCATTTGGTTCTCAG CA ATGGGATGCCACCTTCGG
	<i>reverse</i>	CCGAAGGTGGCATCCC ATTG TGAGAACCAAATGATTG
PdFRS - L414H	<i>forward</i>	CAATCATTTGGTTCTCAG CATT GGGATGCCACCTTCGGC
	<i>reverse</i>	GCCGAAGGTGGCATCCCA ATG CTGAGAACCAAATGATTG

B-4 - Primers para a produção de mutantes da posição 533.

Tabela 26 - Primers utilizados para a produção dos mutantes na posição 533 com destaque para os nucleotídeos alterados em relação ao gene selvagem (violeta).

Mutantes	Primers	Sequências
MiFRS – G533A	<i>forward</i>	CTGGTGGTTTGGCCGCTTGGGAACCTGTTAG
	<i>reverse</i>	CTAACAGGTTCCCAAGCGGCCAAACCACCAG
PdFRS – G530A	<i>forward</i>	CAAGAGGAGGATTAGCTGCTTGGGAACCAATACGTGG
	<i>reverse</i>	CCACGTATTGGTTCCCAAGCAGCTAATCCTCCTCTTG

B-5 - Primers para a produção de mutantes da posição 609.

Tabela 27 - Primers utilizados para a produção dos mutantes na posição 609 com destaque para os nucleotídeos alterados em relação ao gene selvagem (violeta).

Mutantes	Primers	Sequências
MiFRS – Y609A	<i>forward</i>	CCTGACGGTTCTTGGGCTGGTAACTGGGGTATC
	<i>reverse</i>	GATACCCCAGTTACCAGCCCAAGAACCGTCAGG
PdFRS – Y606A	<i>forward</i>	GCCTGACGGTTCCTGGGCTGGAACTGGGGTGTC
	<i>reverse</i>	GACACCCCAGTTTCCAGCCAGGAACCGTCAGGC

B-6 - Primers para a produção de mutantes da posição 729.

Tabela 28 - Primers utilizados para a produção dos mutantes na posição 729 com destaque para os nucleotídeos alterados em relação ao gene selvagem (violeta).

Gene	Primers	Sequências
MiFRS – K729M	<i>forward</i>	GAAATCACAGGTGTTTTTATGATGAACTGCATGTTGCAC
	<i>reverse</i>	GTGCAACATGCAGTTCATCATATAAAACACCTGTGATTTC
MiFRS – K729R	<i>forward</i>	CAAGAAATCACAGGTGTTTTTAGAATGAACTGCATGTTGCACTATC
	<i>reverse</i>	GATAGTGCAACATGCAGTTCATTCTATAAAACACCTGTGATTTCCTTG
MiFRS – K729W	<i>forward</i>	CAACAAGAAATCACAGGTGTTTTTTGGATGAACTGCATGTTGCACTATCC
	<i>reverse</i>	GGATAGTGCAACATGCAGTTCATCCAAAAACACCTGTGATTTCCTTGTTG
MiFRS – K729A	<i>forward</i>	GAAATCACAGGTGTTTTTTGCTATGAACTGCATGTTGCAC
	<i>reverse</i>	GTGCAACATGCAGTTCATAGCAAAAACACCTGTGATTTC
MiFRS – K729E	<i>forward</i>	GAAATCACAGGTGTTTTTTGAAATGAACTGCATGTTGCAC
	<i>reverse</i>	GTGCAACATGCAGTTCATTTCATAAAACACCTGTGATTTC
PdFRS – K726M	<i>forward</i>	GAAATCACAGGCGTCTTTATGAACAACTGTATGTTACAC
	<i>reverse</i>	GTGTAACATACAGTTGTTCTATAAAGACGCCTGTGATTTC
PdFRS – K726R	<i>forward</i>	CAAGAAATCACAGGCGTCTTTAGAACAACCTGTATGTTACACTATC
	<i>reverse</i>	GATAGTGTAACATACAGTTGTTCTCTAAAGACGCCTGTGATTTCCTTG
PdFRS – K726W	<i>forward</i>	CAACAAGAAATCACAGGCGTCTTTTGGAAACAACCTGTATGTTACACTATCC
	<i>reverse</i>	GGATAGTGTAACATACAGTTGTTCCAAAAGACGCCTGTGATTTCCTTGTTG
PdFRS – K726A	<i>forward</i>	GAAATCACAGGCGTCTTTGCTAACAACCTGTATGTTACAC
	<i>reverse</i>	GTGTAACATACAGTTGTTAGCAAAGACGCCTGTGATTTC

Anexo C – Sequência primária dos genes selvagens de friedelina sintase

C-1 – Sequência primária de aminoácidos da enzima *MiFRS*.

MWKIKIADRGNCOPYNEYLYTTNDFVGRQIWEFDPNSTPEELAEIEEARRKFTENRYE
VKPASDLLWMMQFLRKNNFKQTIPPLRIGEKEQVTYEDVTTALRRASSFFSALQASDG
HWPANAGVSFFLPPFIFCLYITGHLNSIITPEHRKEILRFIYNHQNEDGGWGIHIEGHS
TVFATAFTYVCMRILGVGPDEDACARARKWILDRGGITYMASWGKTWFSVLGIFDWHY
GCNPMPEFWILPSYLPHPAKMWCYCRMVYMPMSYLYGKRFVAPITPLILQLREELH
TQPYHEIEWRKMHRCAEEDLYFPHSLIQNFLWDSLYVASEPLLTRWPFSKIRERALE
KAMEHIHYEDENSRYITIGCVEKALCMLCCWVEDPNGEYFKKHLARIPDYLWVAEDG
MKVVSFGSQLWDATFGFQALVASNLTDVAPTTLVKAYDFIKKCQVRDNPSTGNFEKMF
RHISKGSWTFSDQDHGWQLSDCTAEALKCCLLAATMPEELVGEKLDPQWIFESVNIIL
SLQEPKTGGLAGWEPVRAGQWIEMLNPMEFLENIVIEHTYIECTGSSIIAFITLKKLFP
HRTKDIDNFIVNAIRYLEDEQYPDGSWYGNWGICFIYSTMFALGGLAAAGRTYKNCQA
VRRGVDFLLINQSDGGWGESYISCPKRYTPLEGRRSNVQTAWAMLGLLYAGQA
ERDPTPLHRGAKLLINYQMEEGGYPPQEITGVFKMNCMLHYPIYRNAFPIWALGEYRK
RVPLPSKGNMAMKINSA

C-2 – Sequência primária de aminoácidos da enzima *PdFRS*.

MWRLKIAEKGNNPYIFTTNEYAGRQIWEYDPNAGTPEEREQVEEARRNFTKNRSKV
KPSSDLLWQYQILREKNFKQTIPAVRVEEGEEVTEKTTTAMKRSASFYSALQASD
GHWPANSGVLFFLPPFVFCFYITGHLNTMFPPEYRKEIFRYIYNHQNEDGGWGLHI
ESHSNMFCTTFSYICLRMLGVGPDEEACARGRKWILDRGGVTSIPSWGKTWLSILG
LFDWSGCNPMPEFWILPTALPVHPAKMWCYCRLVYMPMSYLYGKRFVGPITPLIL
SLREELYLQPYESVKWKQVRHLCAEEDLYPHSLIQDFLWDSLYLMSEPLLTRWPF
NQLVRKKALEVTMKHIHYEDENSRYITIGCVEKVLCLMLSCWVEDPDGVAFKRHLAR
VPDYLWVGEDGMKVQSFGSQLWDATFGFQALYTSELGEEIKPTLAKAFDFIKKSQV
VDNPAGDFVGMYRHISKGSWTFSDQDHGWQLSDCTAEALKCVLFAQMLPTEYIGE
KMDPQMIFEAVNIILSLQGPRGGLAGWEPPIRGEMWLEKLNPMFLENIVIEHDYVEC
TSSAIHGFVFMFMKMPGHRKKEIETFIARAVDYLEMIQMPDGSWYGNWGVCFIYST
WFALVGLAAAGKTYNNQAMRRGVDFLLRAQSPDGGWGESYLSCPNKIYTPLEEK
RSTYVQTGWAMLGLIHSGQADRDPTPLHRGAKLLINSQAEDGSYPQQEITGVFKNN
CMLHYPTYKNVFPLWALAEYRKNVPLPSKKL

Anexo D – Espectro de massas dos picos identificados por GC-MS

D-1 – Espectro de massas obtido para a friedelina.

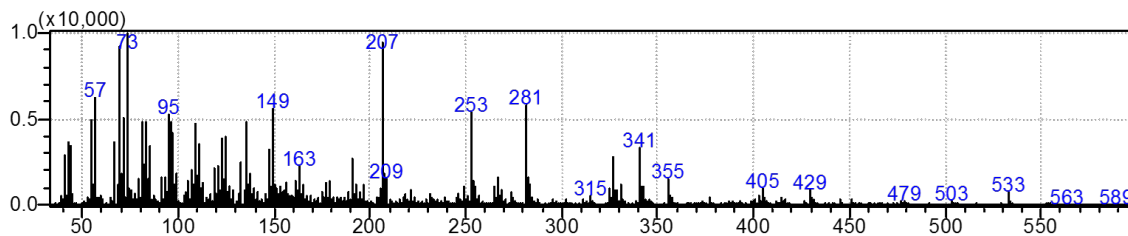


Figura 41 – Espectro de massas obtido para o sinal com $T_R = 24,6$ min e identificado como friedelina através da biblioteca NIST14.

D-2 – Espectro de massas do sinal não-identificado em $T_R = 22,7$ min.

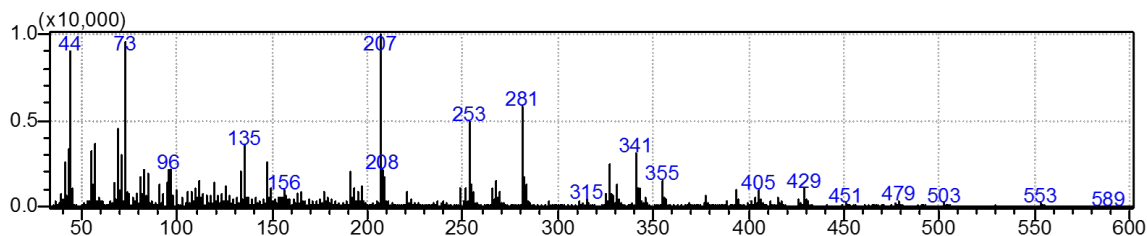


Figura 42 – Espectro de massas obtido para o sinal com $T_R = 22,7$ não-identificado.

Protein	Position	Sequence	Consensus
MiFRS	1	MWKLKIAER-----G-NCPYNEYLYTTNDFVGRQIWEFDPNSG--TPEELAEIEEARRKF	
KdFRS	1	MWKLKIAEG-----G-SDP-----YIYTTNNFVGRQIWEFDPQAT-DPQQLAKVEAARLNF	
PdFRS	1	MWRLKIAEK-----G-NNP-----YIFTTNEYAGRQIWEYDPNAG--TPEEREQVEEARRNF	
AtSHS	1	MWRLKIAEG-----G-NNP-----YLYSTNNFIGRQTWEFDPNYG--TPEERDEVEQARLHF	
BrASS	1	MWKLRIKAE-----AGQDPYLFSTNNYLGRQIWEFDADAG--SAEELAEVEEARRNF	
KdGLS	1	MWKLKIAEG-----G-SNP-----YIFTTNNFVGRQIWEFDPQAT-DPQQLAKVEAARLDF	
BpLUS	1	MWKLKIAEG-----G-PG-----LVSGNDFIGRQHWEFDPDAG--TPQERAEVEKVVREF	
LjLUS	1	MWKLKIAEG-----G-KG-----LVSVSNFIGRQHWVFDPNAG--TPQEHHEEIERMRQEF	
GgLUS	1	MWKLKIEGEG-----G-AG-----LISVNNFIGRQHWEFDPNAG--TPQEHAEIERLRREF	
OeLUS	1	MWKLKIAEG-----TGP-----WLTTTNNHIGRQHWEFDPDAG--TPDERVEVERLREEF	
ToLUS	1	MWKLKIAEG-----G-DDE-----WLTTTNNHVGROHWQFDPDAG--TEEERAEIEKIRLNF	
KdLUS	1	MWKLKIAEG-----G-SNP-----YIFTTNNFVGRQIWEFDPQAT-DPQQLAKVEAARLDF	
RcLUS	1	MWRLKIAEG-----G-NNP-----YIYSTNNFQGRQIWFDPNAG--TPEEQAEVEEARQNF	
BgLUS	1	MWRLKIAEG-----G-NNP-----YIYSTNNFVGRQTWEFDPDAG--TPEERAQVEEARENF	
GgBAS	1	MWRLKIAEG-----G-KDP-----YIYSTNNFVGRQTWEYDPDGG--TPEERAQVDAARLHF	
PsBAS	1	MWRLKIAEG-----G-NDP-----YLFSTNNFVGRQTWEYDPEAG--SEEEARQVEEARRNF	
EtBAS	1	MWKLKIAEG-----G-NDE-----YLYSTNNYVGRQTWVFDPP--TPQELAQVQCARLNF	
BpBAS	1	MWRLKIAEG-----G-SDP-----YIYSTNNFVGRQTWEFDPQAG--SPQERAEVEEARRNF	
PgBAS1	1	MWKLKIAEG-----NKNPD-----YLYSTNNFVGRQTWEFDPDYVASPGELEEEVEQVRRQF	
PgBAS2	1	MWRLMTAKG-----G-NDP-----YLYSTNNFIGRQTWEFDPDYG--TPAERAEVEEARLHF	
MiCAS	1	MWRLKFC AETIGDDG-----SWLRSMNGHVGRQVWEFCPELG--SPEDLSLVENARRGF	
LaCAS	1	MWQLKICADTVPADPSNAG---GWLSSLNNHVGRQVWHFHPELG--TPEDLQOIQHARQRF	
BpCAS1	1	MWKLKIC AETARGDGGGGGSETWLRSLNNHIGRQIWEFHHPELG--TQELQOIDDARRRF	
BpCAS2	1	MWKLKIAE-----GGSPWLRITLNNHVGRQVWEFDPKLG--SPEELAEIERARETS	
KdCAS	1	MWKLKIAE-----AGGSQWLRSVNNHIGRQIWFDFPALG--SPEELAQIEDARDNF	
BrCAS1	1	MWKLKIAE-----GGSPWLRITLNNHVGRQVWEFDPNLG--TPEELAAVEEARKSF	
BrCAS2	1	MWKLKIAE-----GGSPWLRITLNNHVGRQVWEFDPNLG--TPEELAAVEEARKSF	
AtCAS1	1		
AtCAS2	1	MWKLKIAE-----GGSPWLRITLNNHVGRQVWEFDPNLG--TPEDLAAVEEARKSF	
PgCAS	1	MWKLKIAE-----GGNPWLRITLNDHVGRQIWEFDPNIG--SPEELAEVEKVVREF	
PsCAS	1	MWKLKIAE-----GGTPWLRITLNNHVGRQVWEFDPHSG--SPQDLDIETARNF	
LjCAS	1	MWKLKIAE-----GGNPWLRSTNSHVGRQVWEFDPKLG--SPQDLAEIETARNF	
GgCAS	1	MWKLKIAE-----GGSPWLRITVNNHVGRQVWEFDPKLG--SPEDLLEIEKARQNF	
consensus	1		

Tabela 30 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	53	TENRMEVKPASDLLWMQFLREKNNK---	QTIPPLRIGKEQVTYEDVTTALRRASSFFS
KdFRS	50	YNHRHKIKPSSDLLWRMQFLREKDER---	QNTAQVKVEDGEEVSYEAATAALRGVHFYS
PdFRS	50	TKNRSKVKPSSDLLWQYQILREKNEK---	QTIPAVRVEEGEEVITYEKTITAMKRSASFYS
AtSHS	50	WNHRHEIKPSGDTLWRMQFLREKKK---	QTIPQVKTEDDEEISYDKVTATMRRSVHLL
BrASS	51	AANRSHYKASADLLWRMQFLREKKKE---	QNTPRVVIDDAEKTAYEDAELALRRGILYMA
KdGLS	50	YHNRKIKPNSDLLWRMQFLREKDER---	QNTIPQVKVEDGEEVSYEAVTAALRRGVHLYS
BpLUS	48	TKNREQMKQSADLLMRMQLRKENPCQ---	PIPPPVKVKETEVTITEEAVITTLRRSLSFYS
LjLUS	48	TKNRESIKQSADLLMRMQLRKENPCG---	PIPPAVKLRDVEKVTAEALITTLRRSITFYS
GgLUS	48	TKNRESIKQSADLLMRMQLRKENHGTNNN	IPAAVKLSDAENITVEALVTTITRAISFYS
OeLUS	49	KKNRERTKQSADLLMRMQLVKENQRV---	QIPPAIKIKETEGITEEAVITTLRRAISFYS
ToLUS	50	KLNREQFKQSADLLMRMQLRKENPIN---	KIPDAIKLNETEETVINDAVTTTLKRAISFYS
KdLUS	50	YHNRKIKPNSDLLWRMQFLREKAT---	QTIPQVKVEDGEEVSYEAVTAALRRGVHLYS
RcLUS	50	WKNREQVKPNSDLLWRMQFLREKNEK---	CKIPKVKVEDGEEITSEIAAAALRRSVHLFS
BgLUS	50	WRDRELKPSDLLWRMQFLSEKKK---	CRIPQVKVQDGEETTREIATTALRRSVHLVS
GgBAS	50	YNNREQVKPCGDLLWRMQFLRENNEK---	QTIASVKIGDGEETITYEKATTAVRRAHHLS
PsBAS	50	YNNRENVKPCGDLLWRMQFLRENNEK---	QTIGGVKIEDEEETITYEKTITTLRRGTHHLA
EtBAS	50	YNNRMHVKPSDLLWRMQFLREKNEK---	QTIPQAKINEGEDITIEKATTALRRAVHFFS
BpBAS	50	YDNRMQVKPSGDLLWRMQFLREKNEK---	QTIPPVKVEDGEEITYEKTSTAALRRAVHFYS
PgBAS1	52	WDNRMQVKPSGDLLWRMQFLREKNER---	QTIPQVKVGDDAEVITYEAATTLRRAVHFFS
PgBAS2	50	WNNRMQVKPSGDLLWRMQFLREKNEK---	QIIPQVKVEDGEEITYEAATTLRRAVHYFS
MiCAS	54	TNTHLEKKHSADLLMRMQFAKENPYF---	VNIPQIKVKASEDVTDEVTTTLRRAINFYS
LaCAS	57	SDHREKKHSADLLMRMQFAKNNSSF---	VNIPQIKVKDKEDVTEEAVSRTLRRAINFYS
BpCAS1	60	WERREERHSSDLLMRMQFAKENPSS---	ANIPQVKIKDTEEVREEAVGMTLRRAINFYS
BpCAS2	49	LKVREKKHSSDLLMRMQFAKENPRG---	AVLPQVKVNETEDVTEEMVTRMLRRAISFHS
KdCAS	50	ARHREDKKHSADLLMRMQFLTKENPQS---	DLLPKVNIKGIEDITEAVTNTLRRAINFHS
BrCAS1	49	RENREAKKHSSDLLMRMQFSRESLSR---	PVLPQVNIKGDVVTEKMMVETTLKRGVDFYS
BrCAS2	49	RENREAKKHSSDLLMRMQFSRESLSR---	PVLPQVNIKGDVVTEKMMVETTLKRGVDFYS
AtCAS1	1	-----	-----MVETTLKRGIDFYS
AtCAS2	49	SDNREVQKHSADLLMRMQFSRENLS---	PVLPQVKIEDTDDVTEEMVETTLKRGIDFYS
PgCAS	49	RNHREKKHSADLLMRMQFANENPGS---	VVLPQVKVNDGEDISEDKVTVTTLKRAMSFYS
PsCAS	49	HDNRETHKHSADLLMRMQFAKENPMN---	EVLPPKVKVKDVEDVTEEAVATTTLRRGINFYS
LjCAS	49	HDNRESHKHSADLLMRMQFSKENPIG---	EVLPPKVKVKDVEDVTEEAVTTTLRRAISFHS
GgCAS	49	HDNRETHKHSADLLMRMQFAKENPMN---	EVLPPKVRVKDIEDVTEETVKTTLRRAINFHS
consensus	61		*

Tabela 31 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	110	ALQASDGHWPAENAGV	SFFLPPEF	FC	LYITGHLNSIITPEHRKEILRF	ITYNHQ	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
KdFRS	107	ALQASDGHWPAENAG	PMFF	SPLVM	CLYITGHLNTIFTEEHRRET	IRYIY	YHQNEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
PdFRS	107	ALQASDGHWPAENSG	VFFLPPEVFC	FYITGHLNTMFPPEYRKEIF	FRYIYNH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
AtSHS	107	ALLADDGHWPAENSG	SFF	QPLVM	CLYITGHLNSVFP	AEHRKEILRYVY	SHQNKDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
BrASS	108	ALQADDGHWPAENSG	CMFF	APFVIC	LYITGHLDKTF	SQEHKREMLRYMYNH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
KdGLS	107	ALQASDGHWPAENAG	PMFF	PPVM	CLYITGHLNAIFTEEHRSET	IRYIY	YHQNEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
BpLUS	105	SIQAH	DGHWP	ESAGPIFF	LQPEV	ALYITGDLNTIFS	PAHQKEILRYLYNH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
LjLUS	105	SIQAH	DGHWP	ESAGPIFF	QPLVM	ALYITGSLDDVL	GPQHKKEILRYLYNH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
GgLUS	108	SIQAH	DGHWP	ESAGPIFF	LQPLVM	ALYITGSLDDVL	GP	EHKKEIVRYLYNH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
OeLUS	106	TIQAH	DGHWP	ESAGPIFF	LPLVM	ALYITGAINV	VL	SREHQKEILRYIYNH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
ToLUS	107	TIQAH	DGHWP	ESAGPIFF	LPLVM	ALYITGAIN	NDILTPAHQLEIKRYIYNH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
KdLUS	107	ALQASDGHWPAENAG	PMFF	PPVM	CLYITGHLNAIFTEEHRSET	IRYIY	YHQNEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
RcLUS	107	ALQASDGHWCAENAG	LFFLPPLVFA	YITGHLNTVFS	PEHRKEILRYTYCH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
BgLUS	107	ALQASDGHWCAENSG	PMFF	PPVM	FSLYITGHLNAVFSAEHCKEILRYTYCH	P	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
GgBAS	107	ALQTS	DGHWPAQIAGPIFF	LPLVFC	MYITGHLDSVP	PEEYRKEILRYTYYH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
PsBAS	107	TLQTS	DGHWPAQIAGPIFF	MPPLVFC	YITGHLDSVP	PEHRKEILRYTYCH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
EtBAS	107	ALQASDGHWPAENAG	PIFFLPPLVM	CLYITGHLDTVFP	APHRLEILRYTYCH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
BpBAS	107	ALQASDGHWPAENAG	PIFFLPPLVM	CMYITGHLNTVFP	AEHQKEILRYTYYH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
PgBAS1	109	ALQASDGHWPAENSG	PIFFLPPLVM	CMYITGHLDTVFP	AEHRKEILRYTYCH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
PgBAS2	107	ALQADDGHWPAENAG	PIFFLPPLVM	CLYITGHLNTVFP	AEHRIEILRYTYCH	Q	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
MiCAS	111	TIQADDGHWP	GDYGGPMFLPGLIITL	HITGALNAVL	SREHQREMC	RYLYNHQ	NSDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
LaCAS	114	TIQADDGHWP	GDYGGPMFLPGLVITL	SITGALNAVL	STEHQREIC	RYLYNHQ	NKDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
BpCAS1	117	TIQADDGHWP	GDYGGPMFLPGLVITL	SITGTLNAFL	SKEHQCEIC	RYLYNHQ	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
BpCAS2	106	TIQAH	DGHWAGDYGGPMFLMPGLVITL	SITGALNTVL	SEEHKKEMC	RYLYNHQ	NKDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
KdCAS	107	TIQAH	DGHWP	GDYGGPIFF	MPGLVITL	SITGALNAVL	SKEHKKEMC	RYLYNHQ	NEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
BrCAS1	106	TIQASDGHWAGDYGGPMFLPGLIITL	SITGALNTVL	SEQHKAEMRRYL	HNHQNEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
BrCAS2	106	TIQASDGHWAGDYGGPMFLPGLIITL	SITGALNTVL	SEQHKAEMRRYL	HNHQNEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
AtCAS1	15	TIQAH	DGHWP	GDYGGPMFLPGLIITL	SITGALNTVL	SEQHKAEMRRYL	YNHQNEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
AtCAS2	106	TIQAH	DGHWP	GDYGGPMFLPGLIITL	SITGALNTVL	SEQHKAEMRRYL	YNHQNEDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
PgCAS	106	TLQAH	DGHWP	GDYGGPMFLMPGLVITL	SITGVLNV	VL	SKEHKREICRYLYNHQ	NRDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
PsCAS	106	TIQSH	DGHWP	GDYGGPMFLMPGLVITL	SVTGALNAVL	IDEHRKEMRRYL	YNHQ	NKDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
LjCAS	106	TLQSH	DGHWP	GDYGGPMFLMPDLVITL	SITGALNAVL	IDEHRKEMC	RYLYNHQ	NKDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
GgCAS	106	TLQSH	DGHWP	GDYGGPMFLMPGLVITL	SITGALNAVL	TEEHRKEICRYLYNHQ	NKDGGWG																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
consensus	121	****	*	*	**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*</

Tabela 32 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	170	THIEGHSTVFATAFTYVCMRILGVGPD----	EDACARARKWILD	RGGITYMASW	GKGTWFS
KdFRS	167	FHIEGQSTMFGTVLNYICMRLLEGEP-	GGQDNAVS	RGRKWILD	HGGATAIP
PdFRS	167	LHIESHSNMFCTTFSYICLRMLGVGPD----	FEACARGR	KWILD	RGGVTSIP
AtSHS	167	LHMEGHSIMFGTTLSYICMRLLEGEPD-	GGLNGACT	RARKWILD	HGGAIANP
BrASS	168	THVESHSFMFSTVLNYICLRIFGEEADHDG	INSACAMAR	KWILD	HGGATYTP
KdGLS	167	FHIEGHSTMFGTVLNYICMRLLEGEP-	GGQDNAVS	RGRKWILD	HGGATSIP
BpLUS	165	FHIEGHSTMFGSALSYSIALRLLEGLE-	DGEDGAMA	KSRKWILD	HGGGLVAIP
LjLUS	165	FHIEGHSTMFGSALSYSIALRVLGQSLE-	DGEDMAVA	RGRKWILD	HGGGLVAIP
GgLUS	168	FHIEGHSTMFGSALSYSIALRLLEGEP----	QDKAMAKR	KWILD	HGGGLVAIP
OeLUS	166	THIEGHSTMFGSVLSYITLRLLEGQE-	DGEDKAVA	RGRKWILD	HGGAVGIP
ToLUS	167	LHIEGHSTIFGSVLSYITLRLLEGEEAD	SVAEEDMA	LVKR	KWILD
KdLUS	167	FHIEGHSTMFGTVLNYICMRLLEGEP-	GGQDNAVS	RGRKWILD	HGGATSIP
RcLUS	167	THIEGHSTMFC TVLNYICMRLGEARD-	GGIENACER	GRKWILD	HGGATGIS
BgLUS	167	LHIEGHSAMFSTVLNYNWLGLLEGGRD-	GCKDNACER	ARRK	WILD
GgBAS	167	LHIEGHSTMFC TALNYICMRLLEGEPD-	GGQDNACAR	ARKWILD	HGGVTHIP
PsBAS	167	LHIEGHSTMFC TALNYICMRLLEGEPD-	GGEDNACV	RARNWI	RQHGGVTHIP
EtBAS	167	LHIEGHSTMFC TVLSYICMRLLEGEPN-	GGQDNACSR	ARKWILD	HGGATYIP
BpBAS	167	LHIEGHSTMFC TALSYICMRLLEGEPD-	GGQDNACAR	ARKWILD	HGGVTHIP
PgBAS1	169	LHIEGHSTMFC TTLSYICMRLLEGEPD-	GGVNNACAR	GRKWILD	HGSVTAIP
PgBAS2	167	LHIEGHSTMFC TALSYICMRLLEGGRD-	GGENNACAR	ARKWILD	HGSVTAIP
MiCAS	171	LHIEGPSTMFGTALNYVTLRLLEGTEG--	GEGAIENG	RKWILD	HGGATAIT
LaCAS	174	LHIEGPSTMFGSVLNYVSLRLLGEEAED--	GQGAVDKA	RKWILD	HGGASAIT
BpCAS1	177	LHIEGPSTMFGTALNYITLRLLEPED--	GMGAVEKA	RKWILD	HGGATAIT
BpCAS2	166	LHIEGPSTMFGTVLSYVTLRLLEGAND--	GQGAIERGR	KWILD	HGSAIT
KdCAS	167	LHIEGPSTMFGSVLNYVTLRLLEGDVNG--	GDGEIERA	RKWILD	HGGATAIT
BrCAS1	166	LHIEGPSTMFGSVLNYVTLRLLEGPNND--	GDGAMEIG	RDWILN	HGGATNIT
BrCAS2	166	LHIEGPSTMFGSVLNYVTLRLLEGPNND--	GDGAMEIG	RDWILN	HGGATNIT
AtCAS1	75	LHIEGPSTMFGSVLNYVTLRLLEGPNND--	GDDGMEKGR	RDWILN	HGGATNIT
AtCAS2	166	LHIEGPSTMFGSVLNYVTLRLLEGPNND--	GDDGMEKGR	RDWILN	HGGATNIT
PgCAS	166	LHIEGPSTMFGTVLNYVTLRLLEGAND--	GQGAMEKG	RQWILD	HGSATAIT
PsCAS	166	LHIEGPSTMFGSVLCYVTLRLLEGPNND--	GEGDMEGR	RDWILN	HGGATYIT
LjCAS	166	LHIEGPSTMFGSVLNYVTLRLLEGPNND--	GQGDMEKA	RDWILN	HGGATYIT
GgCAS	166	LHIEGPSTMFGSVLNYVALRLLLEGPNND--	RQGE MEKGR	DWILN	HGGATFIT
consensus	181	* * * * *			

Tabela 33 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	226	VLGIFDWFYGNPMPPEFWLLPSYLPHPAKMWCYCRMVYMPMSYLYGKRFVAPITPLILQ
KdFRS	226	IMGLCDWSGGNPMPPEFWLLPSYLPMPHPAKMWCYCRMVYMPMSYLYGKRFTTHITPLILQ
PdFRS	223	ILGLFDWSGGNPMPPEFWLLPTALPVHPAKMWCYCRMVYMPMSYLYGKRFGVGPITPLILS
AtSHS	226	ILGVHEWFGGNPMPPEFWLLPSPSLPMSPGKMWSYCRMVYMPMSYLYGKRFGVGPITPLVLQ
BrASS	228	VLGVYDWSGCKPMPPEFWLLPYSPSPNGGTWIIYLRDVFMTSYLYGKKFVATPTPLILQ
KdGLS	226	IMGLCDWSGGNPMPPEFWLLPSYLPMPHPGKMWCYCRMVYMPMSYLYGKRFTARITPLILQ
BpLUS	224	VLGLYEWSGGNPMPPEFWLLPDIFPIHPGKMLCYCRMVYMPMSYLYGKRFGVGPITGLIQS
LjLUS	224	VLGVYEWSGGNPMPPEFWLLPKIEPIHPGKMLCYCRMVYMPMSYLYGKRFGVGPITALVRS
GgLUS	224	VLGAEWWSGNPMPPEIWLPLPKIAPFHPGKMLCYCRMVYMPMSYLYGKKFVGPITALIRS
OeLUS	225	VLGVYEWDCGNPMPPEFWLLPNSPIHPGKMLCYCRMVYMPMSYLYGKRFGVGPITGLVLS
ToLUS	227	ILGVYEWGGGNPMPPEFWLLPKIEPIHPGKMLCYCRMVYMPMSYLYGKRFGVKITELVRD
KdLUS	226	IMGLCDWSGGNPMPPEFWLLPSYLPMPHPGKMWCYCRMVYMPMSYLYGKRFTARITPLILQ
RcLUS	226	ILGVYEWDCGNPMPPEFWAFPSSEPIHPAKMFCYCRITYMPMSYLYGKRFGVGPITPLILQ
BgLUS	226	ILGVYEWDCGNPMPPEFWAFPTIEPIHPAKMFCYCRITYMAMSYLYGKKFVGPITPLILQ
GgBAS	226	ILGVFDWCGSNPMPPEFWLLPSYLPMPHPAKMWCYCRMVYMPMSYLYGKRFGVGPITPLILQ
PsBAS	226	ILGVFDWLCGNPMPPEFWLLPSYLPMPHPAKMWCYCRMVYMPMSYLYGKRFGVGPITPLILQ
EtBAS	226	ILGVYEWSCGNPMPPEFWLLPTSLPMHPAKMWCYCRMVYMPMSYLYGKRFGVGPITPLILQ
BpBAS	226	ILGIFEWICGNPMPPEFWLLPSYLPMPHPAKMWCYCRMVYMPMSYLYGKRFGVGPITPLILQ
PgBAS1	228	ILGVYEWICGNPMPPEFWLLPSYLPMPHPAKMWCYCRMVYMPMSYLYGKRFGVGPITPLILQ
PgBAS2	226	ILGLFDWSGSNPMPPEFWLLPPELPMHPAKMWCYCRMVYMPMSYLYGKRFGVGPITPLILQ
MiCAS	229	VLGVYEWICGNPMPPEFWLLCPYLLPMPHPGRMWCHCRMVYLPMSYLYGKRFGVGPITATVLS
LaCAS	232	VLGVYEWAGNPNPPEIWLPLPYLLPPIHPGRMWCHCRMVYLPMSYLYGKRFGVGPITPTIRS
BpCAS1	234	VLGVYEWSCGNPMPPEVWLCPYLLPCHPGRMWCHCRMVYLPMSYLYGKRFGVGPITSTIQS
BpCAS2	224	VLGAEWWSGNPMPPEIWLPLPYMLPVHPGRMWCHCRMVYLPMSYLYGKRFGVGPITPTVMS
KdCAS	225	VLGVFEWCGNPNPPEMWLEPYLLPVHPGRMWCHCRMVYLPMSYLYGKRFGVGPITPTVLS
BrCAS1	224	VLGAEWWSGNPMPPEIWLPLPYILPIHPGRMWCHCRMVYLPMSYLYGKRFGVGPITSTVLS
BrCAS2	224	VLGAEWWSGNPMPPEIWLPLPYILPIHPGRMWCHCRMVYLPMSYLYGKRFGVGPITSTVLS
AtCAS1	133	VLGAEWWSGNPMPPEIWLPLPYILPIHPGRMWCHCRMVYLPMSYLYGKRFGVGPITSTVLS
AtCAS2	224	VLGAEWWSGNPMPPEIWLPLPYILPIHPGRMWCHCRMVYLPMSYLYGKRFGVGPITSTVLS
PgCAS	224	VLGVFEWWSGNPMPPEIWLPLPYILPIHPGRMWCHCRMVYLPMSYLYGKRFGVGPITPTVLS
PsCAS	224	VLGVFEWWSGNPMPPEIWLPLPYALPVHPGRMWCHCRMVYLPMSYLYGKRFGVGPITPTVLS
LjCAS	224	VLGVFEWWSGNPMPPEIWLPLPYALPIHPGRMWCHCRMVYLPMSYLYGKRFGVGPITPTILS
GgCAS	224	VLGVYEWWSGNPMPPEIWLPLPYVLPPIHPGRMWCHCRMVYLPMSYLYGKRFGVGPITPTILS
consensus	241	..*...*.*.*.***.*..*..*.....*.....****.*.....*

Tabela 34 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	286	LRRELHTQPYHEIEWRMRHRC	AEDLYFPHS	LIQNFLWDS	SLYVASEP	LLTRWPF	FSKI	-R
KdFRS	286	LRRELHTQPYDQINWKVRH	VCKEDTYYPHPI	IQDLIWD	TLYLTTEP	LLTRWPL	NKLI	R
PdFRS	283	LRRELYLQPYESVKKWKQVRH	LCAEDLYYPHSL	IQDFLWD	SLYLMSEP	LLTRWPF	NQLVR	
AtSHS	286	LRKELYAQPYNDIKWKSSRH	VCAKEDTYYPHPI	IQDLMWDS	SLYITTEP	LLTRWPF	NKLI	-R
BrASS	288	VREELYDQPYKEICWSQARN	LCAKEDLYFP	QSFAQDW	FWNI	VHMFSEN	IFNL	WPFNKLIR
KdGLS	286	LRRELTHIQPYDQIDWKVRH	VCKEDTYYPHPI	IQDLIWD	TLYLTTEP	LLTRWPL	NKLI	R
BpLUS	284	LRQELYNEPYHQINWNKAR	STVAKEDLYYPHPI	IQDLWG	FLHHVAEP	VLTRWPF	SML	-R
LjLUS	284	LRKELYNEPYDRVDWNKAR	NTVAKEDLYYPHPI	IQDMLW	GFLHHVGER	VINTWPF	SML	-R
GgLUS	284	LRRELYNEPYNQINWNKAR	NTVAKEDLYYPHPI	IQDMLW	GFLYHVG	ERFLNCW	PF	SML-R
OeLUS	285	LRQELYTEPYHGINWNKAR	NTCAKEDLYYPHPI	AQDMLW	GFLHHFAEP	VLTRWPF	FSKI	-R
ToLUS	287	LRQELYTDPYDINWNKAR	NTCAKEDLYYPHPI	FVQDMW	GVLHNVEP	VLT	SRPI	STL-R
KdLUS	286	LRRELTHIQPYDQIDWKVRH	VCKEDTYYPHPI	IQDLIWD	TLYLTTEP	LLTRWPL	NKLI	R
RcLUS	286	LRRELYNEPYNKIKWNSVRH	LCAKEDNYTPHPT	IQDLWD	ALYTFSEP	LF	SRWPF	NKLI-R
BgLUS	286	LRRELYNEPYDQINWSEMRH	LCAKEDNYAHTLT	QITLWD	ATYMGEP	LLKRWPF	NKLI	-R
GgBAS	286	LRRELHTEPYEKVNWKARH	QCAKEDLYYPHPI	IQDLIWD	SLYLFTEP	LLTRWPF	NKLI	VR
PsBAS	286	LRRELHTEPYEKINWTRH	LCAKEDTYYPHPI	IQDLIWD	SLYIFTEP	LLTRWPF	NKLI	VR
EtBAS	286	LRQELHTQPYHHINWTRH	LCAHEDVYYPHPI	IQDLMWDS	SLYIFTEP	LLTRWPF	NKLI	IR
BpBAS	286	LRRELYTQPYHQVNWKVRH	LCAKEDTYYPHPI	IQDLIWD	SLYIFTEP	LLTRWPF	NKLI	VR
PgBAS1	288	LRRELYGQPYNEINWRTTR	VCAKEDTYYPHPI	IQDLIWD	SLYITEP	LLTRWPF	NKLI	-R
PgBAS2	286	LRRELYAQAYDEINWVRH	NCAKEDLYYPHPI	IQDLMWDS	SLYIFTEP	FLTRWPF	NKLI	-R
MiCAS	289	LRKELYTVPYHEVDWNRKAR	NTCAKEDLYYPHPI	VQDVLW	SSLYAYEPI	FMRWP	ARR	-R
LaCAS	292	LRKELYLVPYHEVDWNRKAR	NECAKEDLYYPHPI	VQDVLW	ASLHHVEPI	FMRWP	AKRI	-R
BpCAS1	294	LRKELYTVPYHEVDWNRKAR	NCAKEDLYYPHPI	VQDILW	ASLYAYEPI	FMYWP	AKRI	-R
BpCAS2	284	LRKELYSVPYHEVDWNRKAR	NCAKEDLYYPHPI	VQDILW	ASLHKIVEP	VLMRW	PGKRI	-R
KdCAS	285	LRKELETVPYHEVDWNEARS	LCAKEDLYYPHPI	VQDILW	ATLHKIVEP	VLLNW	PGKRI	-R
BrCAS1	284	LRKELETVPYHEVDWNEARN	LCAKEDLYYPHPI	VQDILW	ASLHKIVEP	VLTRW	PGAN	-R
BrCAS2	284	LRKELETVPYHEVDWNEARN	LCAKEDLYYPHPI	VQDILW	ASLHKIVEP	VLTRW	PGAN	-R
AtCAS1	193	LRKELETVPYHEVDWNEARN	LCAKEDLYYPHPI	VQDILW	ASLHKIVEP	VLMRW	PGAN	-R
AtCAS2	284	LRKELETVPYHEVDWNEARN	LCAKEDLYYPHPI	VQDILW	ASLHKIVEP	VLMRW	PGAN	-R
PgCAS	284	LRKEVFSVPYHEVDWNRKAR	NCAKEDLYYPHPI	IQDILW	ASLDKVEPI	FMHWP	AKRI	-R
PsCAS	284	LRKELETVPYHEVDWNRKAR	NCAKEDLYYPHPI	VQDILW	ATLHKIVEP	VFMNW	PGKRI	-R
LjCAS	284	LRKELETVPYHEVDWNRKAR	NCAKEDLYYPHPI	VQDILW	ASLHKIVEP	VLMQW	PGKRI	-R
GgCAS	284	LRKELYTVPYHEVDWNRKAR	NCAKEDLYYPHPI	VQDILW	ASLHKIFLEP	ILMHWP	PGKRI	-R
consensus	301	.* *... *... *... *... *... *... *... *... *... *... *... *						*... *... *

Tabela 35 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	345	EEALKEKAMEHHIHYEDENSRYITIGCVEKALCMLCCWVEDPNGEYFKKHLARIPDYLWVAE
KdFRS	346	EEALKKTMKHHIHYEDENSRYIVIGAVEKVLCLACWVEDPNGDYFKKHLARVPDYFWVAE
PdFRS	343	KKALEVTMKHHIHYEDENSRYITIGCVEKVLCLACWVEDPDGVAFKRHLARVPDYLWVGE
AtSHS	345	KKALATTMRHHIHYEDENSRYITIGSVEKILCLACWDEDPNGVCFKKHLARIPDYIWVAE
BrASS	348	HKALQKTMEIHHIHYEDENATQYITGGATPKPFHMLACWVEDPKSDYFKKHLARVPGFLWIGE
KdGLS	346	QKALQKTMKHHIHYEDENSRYITIGTVEKVLCLACWVEDPNGDYFKKHLARVPDYFWVAE
BpLUS	343	EKALKAAIGHVHYEDENSRYTCIGSVEKVLCLACWAEDPNGEAYKLHLGRIPDNYWVAE
LjLUS	343	QKALEVATNHVRYEDETTRYTCIGSVEKVLCLARWVEDPNSEAYKLHLARIPDYFWLAE
GgLUS	343	RKALEIATNHVHYEDENSRYTCIGSVEKVLCLARWVEDPNSEAYKLHLARIPDYFWLAE
OeLUS	344	EKALKVAMEHHVHYEDMNSRYTCIGCVEKVLCLACWVEDPNSEAYKRHLARIPDYFWVAE
ToLUS	346	EKALKVAMDHHHYEDKSSRYTCIGCVEKVLCLATWVEDPNGDAYKRHLARIPDYFWVAE
KdLUS	346	KKALQTTMKHHIHYEDENSRYITIGCVEKVLCLACWVEDPNGDYFKKHLARIPDYLWIAE
RcLUS	345	EKALKITMDHHIHYEDHNSRYITIGCVEKPLCLACWIEDPHGEAFKKHLARIADYIWVGE
BgLUS	345	EKALKITMDHHIHYEDENSQYITIGSVEKPLMLACWHEDPNGDAFKKHLARIPDYVWLGE
GgBAS	346	EKALQVTMKHHIHYEDETTRYITIGCVEKVLCLACWVEDPNGDAFKKHLARVPDYLWVSE
PsBAS	346	KKALEVTMKHHIHYEDENSRYITIGCVEKVLCLACWVEDPNGDAFKKHLARVPDYLWVSE
EtBAS	346	KKALEVTMKHHIHYEDENSRYITIGCVEKVLCLACWAEDPNGVFPFKKHLARIPDYMVVAE
BpBAS	346	EKALQVTMKHHIHYEDENSRYITIGCVEKVLCLACWVEDPNGDYFKKHLARIPDYIWVAE
PgBAS1	347	EKALQTTMKHHIHYEDENSRYITIGCVEKVLCLACWVEDPNGDYFRKHLARIPDYIWVAE
PgBAS2	345	EKALQTTMKHHIHYEDENSRYITIGCVEKVLCLACWVEDPNGDYFKQHLARIPDYIWVAE
MiCAS	348	EKALQTMQHHIHYEDENRYICIGPVNKVLNMLCCWVEDPQSEAFRLHITPRVYDYLWIAE
LaCAS	351	EKALQCVMQHHIHYEDENRYICIGPVNKVLNMLCCWVEDPHSEAFKLHITPRIFDYLWIAE
BpCAS1	353	EKALDTVMQHHIHYEDENRYICIGPVNKVLNMLCCWAEDPNSEAFKLHLPRILDYLWIAE
BpCAS2	343	EKALRTVLEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRINDYLWIAE
KdCAS	344	EKALCSATEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIDYLWIAE
BrCAS1	343	EKALRTTEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIHDYLWVAE
BrCAS2	343	EKALRTTEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIHDYLWVAE
AtCAS1	252	EKALRTATEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIHDFLWIAE
AtCAS2	343	EKALRTATEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIHDFLWIAE
PgCAS	343	EKSLRTVMEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIHDFLWIAE
PsCAS	343	EKALKTATEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIYDYLWVAE
LjCAS	343	EKALNSVMEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIYDYLWIAE
GgCAS	343	EMALKTATEHHIHYEDENRYICIGPVNKVLNMLCCWVEDPNSEAFKLHLPRIYDYLWIAE
consensus	361	 * . * * . . . * * * * *

Tabela 36 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	405	DGMKLVSE	GSQIWDATF	GFQALVASNL	TDEVAPTLV	KAYDFIKK	COVRDNPS	GNFEKMF	
KdFRS	406	DGMKIQSE	GSQHWDTA	FFVQALLAS	DMTDEIRTT	LAKAHDCI	KKSQVKDN	PSGDFRSMY	
PdFRS	403	DGMKMQSE	GSQIWDATF	GFQALYTSEL	GEEIKPTL	AKAFDFIK	KSQVVDNP	AGDFVGMY	
AtSHS	405	DGMKMQTE	GSQVWDAS	IGIQALLA	TELTHDIAP	ILKKGHEF	IKASQVRN	PSGDFKSMY	
BrASS	408	DGLIKTQTE	GSQIWDTA	FTLQMLAAD	VNDIEIRST	TLIKGYSY	LRKSQITEN	PSGDFIKMF	
KdGLS	406	DGMKIQSE	GSQHWDTV	FSQAQALL	ASDMADEIG	TTLAKAHY	CIKESQVK	DNPSGDFRSMY	
BpLUS	403	DGLIKTQSE	GCOMWDAG	FAIQALLS	CNINEEW	PPTLRKAH	EFVKASQV	PENPSGDFKAMY	
LjLUS	403	DGLIKTQSE	GCOMWDAA	FAIQALLS	GNVSEEW	GPTLKAH	HFVKASQV	RENPSGDFKAMY	
GgLUS	403	DGLIKTQSE	GCOMWDAA	FAIQALLA	CNVSEEW	GPTLKAH	HFVKASQV	RENPSGDFNAMY	
OeLUS	404	DGLIKMQSE	GCOMWDAA	FAIQALLS	SNLAEEW	GPTLMKAH	NFVKASQV	QENPSGDFNEMY	
ToLUS	406	DGMKMQSE	GCOMWDAA	FAIQALLF	SSNLTEEW	GPTLKAH	EFVKASQV	RNPFGDFSKMY	
KdLUS	406	DGMKMQSE	GSQHWDTA	FSIQALLA	SNMAEEIG	ITLAKGH	DFIKKSQV	KDNPSGDFKGMY	
RcLUS	405	DGLIKMQSE	GSQIWDTS	IAIQALLA	SDLSHEIG	PPTLKQGH	VFTKNSQAT	ENPSGDFRKM	
BgLUS	405	DGLIKTQSE	GSQVWDTS	IFVQALLA	SNLPSETG	PPTLEK	GHNFIKNS	QVTQNP	SGDFRRMF
GgBAS	406	DGMTMQSE	GSQEWDA	GFQALLA	TNLVEEI	APT	LAKGHDFIK	KSQVRDNPSGDFKSMY	
PsBAS	406	DGMTMQSE	GSQEWDA	GFQALLA	TNLIEEI	KPALAK	GHDFIKKSQV	TENPSGDFKSMH	
EtBAS	406	DGMKMQSE	GSQOWDT	GFQALLA	SNLTEEI	GQVLK	KGHD	FIKKSQVKENPSGDFKSMH	
BpBAS	406	DGLIKMQSE	GSQEWDT	GFQALLA	SNLTDEI	GPTLARG	HDFIKKSQV	KDNPSGDFESMH	
PgBAS1	407	DGMKMQSE	GSQEWDT	GFSTQALL	DSDLTHEI	GPTLMK	GHDFIKKSQV	KDNPSGDFKSMY	
PgBAS2	405	DGMKMQSE	GSQEWDT	GFQALLA	SDLI	DEIRPTLM	KGHDFIKKSQV	KENPSGDFKSMH	
MiCAS	408	DGMKMQG	NGSQIWD	SAFAQAIT	STDLVEE	WGLCLK	RAHNYIK	KDSQVLED	CPGDLSEFWY
LaCAS	411	DGMKMQG	NGSQIWD	TAFQAIMS	KLAEEW	GTTLRKA	HKYIK	KDSQVLED	CPGDLQSWY
BpCAS1	413	DGMKMQG	NGSQIWD	TTFQAIT	STNLAE	EWGQTLR	KAHEYI	KDSQVLED	CPGDLNFWY
BpCAS2	403	DGMKMQG	NGSQIWD	TAFQAIT	STNLFE	EWGPTLE	KAHMYI	KKSQVRED	CPGDLDFWY
KdCAS	404	DGMKMQG	NGSQIWD	TAFSTQAI	VATKLVEE	FSSTLS	KAHEFM	KNSQVLED	YPGDLSEWY
BrCAS1	403	DGMKMQG	NGSQIWD	TSTFAQ	AVLATNL	VEEWGP	VLKKAHSY	VKNSQV	SEDCPGDLSEWY
BrCAS2	403	DGMKMQG	NGSQIWD	TSTFAQ	AVLATNL	VEEWGP	VLKKAHSY	VKNSQV	SEDCPGDLSEWY
AtCAS1	312	DGMKMQG	NGSQIWD	TSTFAQ	AVLATNL	VEEWGP	VLEKAHSF	VKNSQV	LEDCPGDLNFWY
AtCAS2	403	DGMKMQG	NGSQIWD	TSTFAQ	AVLATNL	VEEWGP	VLEKAHSF	VKNSQV	LEDCPGDLNFWY
PgCAS	403	DGMKMQG	NGSQIWD	TAFQAIT	STNLAE	EWGPTLR	KAHTFM	KNSQVLE	DCPGDLDAWY
PsCAS	403	DGMKMQG	NGSQIWD	TAFQAIT	STNLID	EWGPTLR	KAHAFI	KNSQVSE	DCPGDLSEWY
LjCAS	403	DGMKMQG	NGSQIWD	TAFQAIT	STNLIE	EWGPTLR	KAHTFI	KNSQVLE	DCPGDLNFWY
GgCAS	403	DGMKMQG	NGSQIWD	TAFQAIT	STNLIE	EWGPTLR	KAHTFI	KNSQVLE	DCPGDLSEWY
consensus	421	**		*.*.*		*			

Tabela 37 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	464	RHISKGSWTFSDQDHGWQISDCTAEALKCCLLAATMPPELVGEKLDQWIFESVNIILSL
KdFRS	465	RHISKGAWTFSDQDHGWQISDCTAEGLKCCLLFSLMQPEVVGEMPPERLYDSVNVILYL
PdFRS	462	RHISKGSWTFSDQDHGWQISDCTAEALKCVLFAQMIPTETVGEKMDPQMIFEAVNIILSL
AtSHS	464	RHISKGSWTFSDQDHGWQISDCTTIGLTCCLLLSTMPPETVGEKMDPEQLKDAVNVILSL
BrASS	467	RDISKGSWTFSDKDGWQISDCTLSLKKCLLFESMPSEFTGEKMDVERLYDAINIIILYL
KdGLS	465	RHISKGSWTFSDQDHGWQISDCTAEGLKCCLLFSLMQPEVVGEMPPERLFDVSNVILYL
BpLUS	462	RHINKGAWTFSMQDHGWQISDCTAEGLKVAIILFSQMPPELVGEKIEKERLYDAVNVILSL
LjLUS	462	RHISKGAWTFSMQDHGWQISDCTAEGLKVAIILFSQMPPELVGEKIEKERLYDAVNVILSL
GgLUS	462	RHISKGAWTFSMQDHGWQISDCTAEGLKVAIILFSQMPPELVGEKIEKERLYDAVNVILSL
OeLUS	463	RHISKGAWTFSMQDHGWQISDCTAEGLKVAIILFSQMPPELVGEKIEKERLYDAVNVILSL
ToLUS	465	RHISKGAWTFSDQDHGWQISDCTAEGLKVSLLYSQMPPELVGEKIEKERLYDAVNVILSL
KdLUS	465	RHISKGAWTFSDQDHGWQISDCTAEGLKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
RcLUS	464	RHISKGAWTFSDKDGWQISDCTAELSKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
BgLUS	464	RHISKGSWTFSDKDHGWQISDCTAELSKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
GgBAS	465	RHISKGSWTFSDQDHGWQISDCTAEGLKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
PsBAS	465	RHISKGSWTFSDQDHGWQISDCTAEGLKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
EtBAS	465	RHISKGSWTFSDQDHGWQISDCTAEGLKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
BpBAS	465	RHISKGSWTFSDQDHGWQISDCTAEGLKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
PgBAS1	466	RHISKGSWTFSDQDHGWQISDCTAEGLKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
PgBAS2	464	RHISKGSWTFSDQDHGWQISDCTAELKCCLLFSLMQPEVVGESMAPESLYNSVNVILSL
MiCAS	468	RHISKGAWPFSTADHGWPISDCTAEGLKAAALLSTFPSEIVGEPLDAERLYDAVNVILSL
LaCAS	471	RHISKGAWPFSTADHGWPISDCTAEGLKAVLLLSKIPSEIVGKSIDEEQIYDAVNVILSL
BpCAS1	473	RHISKGAWPFSTADHGWPISDCTAEGLKAVLLLSKIPSEIVGKSIDEEQIYDAVNVILSL
BpCAS2	463	RHISKGAWPFSTADHGWPISDCTAEGLKAAALLSKIPSEIVGKSIDEEQIYDAVNVILSL
KdCAS	464	RHISKGAWPFSTADHGWPISDCTAEGLKAVLLLSKIPSEIVGKSIDEEQIYDAVNVILSL
BrCAS1	463	RHISKGAWPFSTADHGWPISDCTAEGLKAAALLSKIPSEIVGKSIDEEQIYDAVNVILSL
BrCAS2	463	RHISKGAWPFSTADHGWPISDCTAEGLKAAALLSKIPSEIVGKSIDEEQIYDAVNVILSL
AtCAS1	372	RHISKGAWPFSTADHGWPISDCTAEGLKAAALLSKIPSEIVGKSIDEEQIYDAVNVILSL
AtCAS2	463	RHISKGAWPFSTADHGWPISDCTAEGLKAAALLSKIPSEIVGKSIDEEQIYDAVNVILSL
PgCAS	463	RHISKGAWPFSTADHGWPISDCTAEGLKAVLLLSKIPSEIVGKSIDEEQIYDAVNVILSL
PsCAS	463	RHISKGAWPFSTADHGWPISDCTAEGLKAVLLLSKIPSEIVGKSIDEEQIYDAVNVILSL
LjCAS	463	RHISKGAWPFSTADHGWPISDCTAEGLKAVLLLSKIPSEIVGKSIDEEQIYDAVNVILSL
GgCAS	463	RHISKGAWPFSTADHGWPISDCTAEGLKAVLLLSKIPSEIVGKSIDEEQIYDAVNVILSL
consensus	481	*...*...*...*...*...*...*...*...*...*...*...*...*

Tabela 38 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	524	QEPKTTGGIAGWEPVRAGQWIEMLNPMEFLENIVIEHTYIJECTGSSIIIFAFITLKKLFPGHR
KdFRS	525	QSKN-GGMPGWEPAGESEWLELLNPTEFFENIVIEHEYVECTSSAVQALVLFKKLYPLHR
PdFRS	522	QGPR-GGIAGWEPPIRCWLEKLNPMFLNIVIEHDYVECTSSAIGHFVMFMKMPGHR
AtSHS	524	QSEN-GGLAAWEPAGSSNWLEMLNPIEFIEDIVIEHEYVECTGTGMEALVLFKKLYPKHR
BrASS	527	QGNN-GGLAIWEPAPCKTWLEWLNPFVEFVEESTVEHSELECTGSATVALLOFMKQFPDHR
KdGLS	525	QSKN-GGMPGWEPAGESEWLELLNPTEFFENIVIEHEYVECTSSAVQALVLFKKLHPGHR
BpLUS	522	QSSN-GGFPWEPQRAYGWLEKFNPTTEFFEDTLIEREYVECTSPAVHGLALFRKLYPRHR
LjLUS	522	QSSN-GGFPWEPQRAYQWLEKFNPTTEFFEDTLIEREYVECTGSAMQALALFRKLYPKHR
GgLUS	522	QSSN-GGFPWEPQKAYRWLEKFNPTTEFFEDTMIEREYVECTGSAMQGLALFRKQFPQHR
OeLUS	523	QSAS-GGFPWEPQKAYRWLEKLNPTTEFFEDVLIERYVECTSSAVQALKLFKLHPGHR
ToLUS	525	QSEN-GGFPWEPQRAYAWLEKFNPTTEFFEDVLIERYVECTSSAIQCLTLFKKLHPGHR
KdLUS	525	QSQN-GGLPAWEPAGAPEWLELLNPTEFFENIVIEHEYVECTSSAVQALVLFKKLYPLHR
RcLUS	524	QSQN-GGFTAWEPPARAGSWMEWLNPFVEMEDIVIEHEYVECTSSAIQALVLFKKLYPRHR
BgLUS	524	QSKN-GGCSAWEPAGAGSWMEWLNPFVFLADIVIEHEYVECTSSSIQALVLFKKLYPEHR
GgBAS	525	QSKK-GGLSAWEPAGAQEWLELLNPTEFFADIVIEHEYVECTGSAIQALVLFKKLYPGHR
PsBAS	525	QSKK-GGLAAWEPAGAQEWLELLNPTEFFADIVIEHEYVECTGSAIQALVLFKKLYPGHR
EtBAS	525	QSKN-GGLAAWEPAGAQQWLEMLNPTEFFADIVIEHEYVECTASAIHALIMFKKLYPGHR
BpBAS	525	QSKN-GGLAAWEPAGAQEWLELLNSTEFFADIVIEHEYIJECTASAVQTLVLFKKLYPGHR
PgBAS1	526	QRKN-GGLSAWEPAGAQEWLELLNPTEFFADIVIEHEYVECTSSAIQALVLFKKLYPGHR
PgBAS2	524	QSKN-GGLAAWEPAGSSEWLELLNPTEFFEDIVIEHEYVECTSSAIQAMVMFKKLYPGHR
MiCAS	528	QNAD-GGFATYELTRSYQWLEFINPAETFGDIVIDYPYVECTSAAIQALTFLKRFYPTHR
LaCAS	531	QNTD-GGFATYELTRSYPWLELINPAETFGDIVIDYTYVECTSAAIQALVAFKKLYPGHR
BpCAS1	533	QNTD-GGFATYELTRSYHWLELINPAETFGDIVIDYPYVECTSAAIQALTFLKKLHPGHR
BpCAS2	523	QNAD-GGFATYELTRSYPWLELINPAETFGDIVIDYNVVECTSAAIQALTSFKKSYPKHR
KdCAS	524	QNID-GGFATYELTRSYSWLELLNPAETFGDIVIDYPYVECTSAAIQSLVLFKKLHPHR
BrCAS1	523	QNAD-GGFATYELTRSYPWLELINPAETFGDIVIDYPYVECTSAAIQALVAFRKLYPGHR
BrCAS2	523	QNAD-GGFATYELTRSYPWLELINPAETFGDIVIDYPYVECTSAAIQALVAFRKLYPGHR
AtCAS1	432	QNAD-GGLATYELTRSYPWLELINPAETFGDIVIDYPYVECTSAAIQALISFRKLYPGHR
AtCAS2	523	QNAD-GGLATYELTRSYPWLELINPAETFGDIVIDYPYVECTSAAIQALISFRKLYPGHR
PgCAS	523	QNSD-GGYATYELTRSYSWLELINPAETFGDIVIDYPYVECTSAAIQALTAFKKLFPGHR
PsCAS	523	QNEH-GGLATYELTRSYTWLELINPAETFGDIVIDCPYVECTSAAIQALATFGKLYPGHR
LjCAS	523	QNEH-GGLATYELTRSYWLELINPAETFGDIVIDYPYVECTSAAIQALTSFRKLYPGHR
GgCAS	523	QNEH-GGFATYELTRSYTWLELINPAETFGDIVIDYPYVECTSAAIQALTSFKKLYPGHR
consensus	541	* ** * * * * *

Tabela 39 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	584	T	K	D	I	D	N	F	I	V	N	A	I	R	Y	L	E	D	E	Q	P	D	G	S	W	Y	G	N	W	G	I	C	F	I	Y	S	T	M	F	A	L	G	G	L	A	A	A	G	R	T	Y	K	N	C	Q	A	V	R	R	G																																																																																																																																																																																																																																																																																																																																																																											
KdFRS	584	R	K	E	V	E	R	F	I	T	N	C	A	K	Y	L	E	D	I	Q	M	P	D	G	S	W	Y	G	N	W	G	V	C	F	T	Y	G	A	W	F	A	L	E	G	L	S	A	A	G	K	T	Y	N	N	C	A	A	V	R	K	G																																																																																																																																																																																																																																																																																																																																																																										
PdFRS	581	K	K	E	I	E	T	F	I	A	R	A	V	D	Y	L	E	M	I	Q	M	P	D	G	S	W	Y	G	N	W	G	V	C	F	I	Y	S	T	M	F	A	L	V	G	L	A	A	A	G	K	T	Y	Y	N	N	Q	A	M	R	R	G																																																																																																																																																																																																																																																																																																																																																																										
AtSHS	583	T	K	E	V	E	S	F	L	T	N	A	A	R	Y	L	D	N	T	Q	M	P	D	G	S	W	Y	G	E	W	I	C	F	T	Y	G	T	Y	A	L	G	G	L	A	A	I	E	K	T	Y	E	N	C	S	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																													
BrASS	586	R	E	E	I	E	G	F	V	T	K	G	V	K	Y	I	E	D	I	O	M	P	N	G	S	W	Y	G	N	W	G	V	C	F	I	Y	G	T	F	A	V	R	G	L	V	A	A	G	K	T	Y	H	N	O	E	T	I	R	R	A																																																																																																																																																																																																																																																																																																																																																																											
KdGLS	584	R	K	E	V	E	R	F	I	T	N	C	A	K	Y	I	E	D	I	Q	M	P	D	G	A	W	Y	G	N	W	G	V	C	F	T	Y	G	A	W	F	A	L	G	G	L	A	A	A	G	K	T	Y	N	N	C	A	A	V	R	K																																																																																																																																																																																																																																																																																																																																																																											
BpLUS	581	G	T	E	I	D	S	S	I	Y	R	G	I	Q	Y	I	E	D	V	Q	E	P	D	G	S	W	Y	G	H	W	I	C	Y	T	Y	G	T	W	F	A	V	G	A	L	A	C	R	N	Y	K	N	C	P	A	L	R	K	S																																																																																																																																																																																																																																																																																																																																																																													
LjLUS	581	R	K	E	I	D	R	C	I	S	K	A	I	R	Y	I	E	N	T	Q	N	P	D	G	S	W	Y	G	C	W	G	I	C	Y	T	Y	G	T	W	F	A	V	E	G	L	T	A	C	G	K	N	F	Q	N	S	V	T	I	R	R	A																																																																																																																																																																																																																																																																																																																																																																										
GgLUS	581	S	K	E	I	D	R	C	I	A	K	A	I	R	Y	I	E	N	M	Q	N	P	D	G	S	W	Y	G	C	W	G	I	C	Y	T	Y	G	T	W	F	A	V	E	G	L	T	A	C	G	K	N	C	H	N	S	L	S	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																										
OeLUS	582	R	K	E	I	A	S	C	I	S	K	A	T	Q	Y	I	E	A	T	Q	N	P	D	G	S	W	D	G	S	W	I	C	F	T	Y	G	T	W	F	A	V	E	G	L	V	A	C	G	K	N	Y	H	N	S	P	T	L	R	R	A																																																																																																																																																																																																																																																																																																																																																																											
ToLUS	584	T	K	E	I	E	H	C	I	S	R	A	V	K	Y	V	E	D	T	Q	E	S	D	G	S	W	Y	G	C	W	G	I	C	Y	T	Y	G	T	W	F	A	V	D	A	L	V	A	C	G	K	N	Y	H	N	C	P	A	L	Q	K	A																																																																																																																																																																																																																																																																																																																																																																										
KdLUS	584	R	K	E	V	E	R	F	I	T	N	C	A	K	Y	L	E	D	I	Q	M	P	D	G	S	W	Y	G	N	W	G	V	C	F	T	Y	G	A	W	F	A	L	E	G	L	S	A	A	G	K	T	Y	N	N	C	A	A	V	R	K	G																																																																																																																																																																																																																																																																																																																																																																										
RcLUS	583	N	K	E	I	E	N	C	I	I	N	A	A	Q	F	I	E	N	I	Q	E	P	D	G	S	W	Y	G	N	W	G	I	C	F	S	Y	G	T	W	F	A	L	K	G	L	A	A	A	G	R	T	Y	E	N	C	S	A	I	R	K	G																																																																																																																																																																																																																																																																																																																																																																										
BgLUS	583	R	K	E	I	E	I	F	I	L	N	A	V	R	F	T	E	I	Q	O	P	D	G	S	W	Y	G	N	W	G	I	C	F	L	S	G	T	W	F	G	L	K	G	L	A	A	A	G	K	T	Y	N	C	T	A	V	R	K	G																																																																																																																																																																																																																																																																																																																																																																												
GgBAS	584	K	K	E	I	E	N	F	I	A	N	A	V	R	F	L	E	D	T	Q	T	A	D	G	S	W	Y	G	N	W	G	V	C	F	T	Y	G	S	W	F	A	L	G	L	A	A	A	G	K	T	F	A	N	C	A	A	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																											
PsBAS	584	K	K	E	I	E	N	F	I	A	N	A	V	R	F	L	E	D	T	Q	T	E	D	G	S	W	Y	G	N	W	G	V	C	F	T	Y	G	S	W	F	A	L	G	L	A	A	A	G	K	T	Y	N	C	A	A	I	R	K	G																																																																																																																																																																																																																																																																																																																																																																												
EtBAS	584	K	K	E	I	E	N	F	I	T	N	A	V	K	Y	L	E	D	V	Q	T	A	D	G	G	W	Y	G	N	W	G	V	C	F	T	Y	G	T	W	F	A	V	G	G	L	A	A	A	G	K	N	Y	N	N	C	A	A	M	R	K	A																																																																																																																																																																																																																																																																																																																																																																										
BpBAS	584	K	K	E	I	E	N	F	I	K	N	A	A	Q	F	I	Q	V	I	Q	M	P	D	G	S	W	Y	G	N	W	G	V	C	F	T	Y	G	T	W	F	A	L	G	G	L	A	A	V	G	K	T	Y	N	N	C	L	A	V	R	R	A																																																																																																																																																																																																																																																																																																																																																																										
PgBAS1	585	K	K	E	I	D	N	F	I	T	N	A	V	R	Y	L	E	D	T	Q	M	P	D	G	S	W	Y	G	N	W	G	V	C	F	T	Y	G	S	W	F	A	L	G	G	L	A	A	A	G	K	T	Y	N	N	C	A	A	V	R	K	A																																																																																																																																																																																																																																																																																																																																																																										
PgBAS2	583	K	K	E	I	E	V	S	I	T	N	A	V	Q	Y	L	E	D	I	Q	M	P	D	G	S	W	Y	G	N	W	G	V	C	F	T	Y	G	T	W	F	A	M	G	G	L	T	A	A	G	K	T	Y	N	N	C	Q	T	L	H	K	A																																																																																																																																																																																																																																																																																																																																																																										
MiCAS	587	R	E	E	V	D	N	C	I	T	K	A	G	E	F	I	E	R	I	Q	E	E	N	G	S	W	Y	G	S	W	G	V	C	F	S	Y	G	W	F	G	I	K	G	L	V	A	A	G	K	T	Y	S	N	O	H	S	L	R	K	A																																																																																																																																																																																																																																																																																																																																																																											
LaCAS	590	R	D	E	I	D	N	C	V	A	K	A	A	D	F	I	E	S	I	Q	A	T	D	G	S	W	Y	G	S	W	G	V	C	F	T	Y	G	W	F	G	I	R	G	L	V	A	A	G	R	R	Y	D	N	C	S	S	L	R	K	A																																																																																																																																																																																																																																																																																																																																																																											
BpCAS1	592	R	E	E	I	E	N	C	I	A	K	A	E	F	I	E	N	I	Q	A	S	D	G	S	W	Y	G	S	W	G	V	C	F	T	Y	A	G	W	F	G	I	K	G	L	V	A	A	G	R	T	Y	K	N	C	S	I	H	K	A																																																																																																																																																																																																																																																																																																																																																																												
BpCAS2	582	E	E	E	V	D	V	C	I	K	R	A	A	F	T	E	K	I	Q	A	S	D	G	S	W	Y	G	S	W	G	V	C	F	T	Y	G	T	W	F	G	V	K	G	L	V	A	A	G	K	N	F	N	D	C	F	G	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																											
KdCAS	583	K	E	E	V	E	L	C	I	K	K	A	A	F	I	E	K	I	Q	E	S	D	G	S	W	Y	G	S	W	A	V	C	F	T	Y	G	T	W	F	G	V	L	G	L	V	A	A	G	R	N	Y	K	N	S	P	S	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																											
BrCAS1	582	K	K	E	V	D	E	C	I	E	K	A	V	K	F	I	E	S	I	Q	E	S	D	G	S	W	Y	G	S	W	A	V	C	F	T	Y	G	T	W	F	G	V	K	G	L	E	A	A	G	K	T	L	K	N	S	P	T	V	A	K	A																																																																																																																																																																																																																																																																																																																																																																										
BrCAS2	582	K	K	E	V	D	E	C	I	E	K	A	V	K	F	I	E	S	I	Q	E	S	D	G	S	W	Y	G	S	W	A	V	C	F	T	Y	G	T	W	F	G	V	K	G	L	E	A	A	G	K	T	L	K	N	S	P	T	V	A	K	A																																																																																																																																																																																																																																																																																																																																																																										
AtCAS1	491	K	K	E	V	D	E	C	I	E	K	A	V	K	F	I	E	S	I	Q	A	A	D	G	S	W	Y	G	S	W	A	V	C	F	T	Y	G	T	W	F	G	V	K	G	L	V	A	V	G	K	T	L	K	N	S	P	H	V	A	K	A																																																																																																																																																																																																																																																																																																																																																																										
AtCAS2	582	K	K	E	V	D	E	C	I	E	K	A	V	K	F	I	E	S	I	Q	A	A	D	G	S	W	Y	G	S	W	A	V	C	F	T	Y	G	T	W	F	G	V	K	G	L	V	A	V	G	K	T	L	K	N	S	P	H	V	A	K	A																																																																																																																																																																																																																																																																																																																																																																										
PgCAS	582	R	E	E	I	Q	H	S	I	E	K	A	A	F	I	E	K	I	Q	S	S	D	G	S	W	Y	G	S	W	G	V	C	F	T	Y	G	T	W	F	G	I	K	G	L	V	T	A	G	R	T	F	S	S	C	A	S	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																											
PsCAS	582	R	E	E	I	Q	C	C	I	E	K	A	V	A	F	I	E	K	I	Q	A	S	D	G	S	W	Y	G	S	W	G	V	C	F	T	Y	G	T	W	F	G	I	K	G	L	I	A	A	G	K	N	F	S	N	C	L	S	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																										
LjCAS	582	R	E	E	I	Q	H	S	I	E	K	A	A	A	F	I	E	K	I	Q	S	S	D	G	S	W	Y	G	S	W	G	V	C	F	T	Y	G	T	W	F	G	V	K	G	L	I	A	A	G	K	S	F	S	N	C	S	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																											
GgCAS	582	R	E	E	I	Q	C	C	I	E	K	A	A	S	F	I	E	K	T	Q	A	S	D	G	S	W	Y	G	S	W	G	V	C	F	T	Y	G	T	W	F	G	V	K	G	L	I	A	A	G	K	S	F	N	N	C	S	I	R	K	A																																																																																																																																																																																																																																																																																																																																																																											
consensus	601	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Tabela 40 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	644	VD F LLINQSD D GGWGESY T SC P RRK K Y T PLEGR S SNV V QTAWAM L GL L IYAGQAERDPTPLH
KdFRS	644	VD F LLINI Q LEDGGWGESY Q SC P DK K Y V PLED N RSN L VQTSWAL M GL L IYAGQADRPTPLH
PdFRS	641	VD F LLIRAQSPDGGWGESY L SC P NK I Y T PLEE K RST V YVQTGWAM L GL L IHS G QADRPTPLH
AtSHS	643	VR F LLKTQGEDGGWGESY R SCA E K I Y I PLD G NRST V VHTAWAM L GL M HSK Q EERDPIPLH
BrASS	646	VR F LL E TQ N IEGGWGESY L SC P RRK K Y T PLE G NK T INV V NTGQAMMAL I MCGQMERDPLPVH
KdGLS	644	VD F LLRI Q LEDGGWGESY Q SC P DK K Y V PLED N RSN L VHTSWAL M GL L CSGQADRDPNPLH
BpLUS	641	CE F LLSKQLP N GGWGESY L SSQ N KV V TN I EG N RAN L VQTAWA L LS L IDARQAEIDPTPIH
LjLUS	641	CK F LLSKQLP N GGWGESY L SSQ D KV Y TN I EG K RAN L VQSSWA L LS L MRAGQAEIDPTPIH
GgLUS	641	CQ F LLSKQLP N AGWGESY L SSQ N KV Y TN L EG N RAN L VQSSWA L LS L THAGQAEIDPTPIH
OeLUS	642	CE F LLSKQLP D GGW S ESY L SS S NKV Y TN L EG N RSN L VQTSWA L LS L IKAGQVEIDPGPIH
ToLUS	644	CK F LLSKQLP D GGWGESY L SS S NKV Y TN L EG N RSN L VHTSWA L LS L IKAGQAEIDPTPIH
KdLUS	644	VD F LLINI Q LEDGGWGESY Q SC P DK K Y V PLED N RSN L VQTSWAL M GL L IYAGQADRPTPLH
RcLUS	643	VD F LLKSQ R DDGGWA E SY L SC P KK V Y V P E GNRSN L VQTAWAM M GL L IYGGQAKRDPMPH
BgLUS	643	VE F LLQ T ORDDGGWGESY L SC P KK I Y V P L EGNRSN L VQTALAM M GL L ILGGQ C ERDPTPLH
GgBAS	644	VK F LL I TQ R EDGGWGESY L SP P K I Y V P L EG S RSNV V HTAWAL M GL L IHAGQAE R DAPLH
PtBAS	644	VK F LL I TQ R EDGGWGESY L SP P K I Y V P L EG N RSNV V HTAWAL M GL L IHAGQ S ERDPTPLH
EtBAS	644	VD F LLRTQKQDGGWGESY L SC P HK K Y V P L ED N RSN L VHTSWAL M GL L ISAGQMDRDPPTPLH
BpBAS	644	VD F LLIRAQ R DNGGWGESY L SC P KK E Y V P L EG N K S N L VHTAWAM M GL L IHAGQAERDPTPLH
PgBAS1	645	VE F LLKSQ V DDGGWGESY L SC P KK V Y V P L EGNRSN L VHTGWAL M GL L IHSEQAERDPTPLH
PgBAS2	643	VD F LLKSQ R SDGGWGESY L SC P NK E Y T P L EGNRSN L VHTSWAM M GL L IHSRQAERDPTPLH
MiCAS	647	CAYLLSKEL R TGGWGESY L SCQ N KV Y TN L K G DRPH I VNTGWAM L ALIEAGQAERDPTPLH
LaCAS	650	CD F LLSKELASGGWGESY L SGQ N KV Y TN L K D DRPH I VNTGWAM L SLIDAGQ S ERDPTPLH
BpCAS1	652	CD V LLSKELASGGWGESY L SCQ D KV Y TN L K D NRPH I VNTGWAM L ALIDAGQAERDPTPLH
BpCAS2	642	CD F LLSKQLPSGGWGESY L SCQ N KV Y SH V EGNRS H VNTGWAM L ALIEAGQAERDPTPLH
KdCAS	643	CD F LLSKQLASGGWGESY L SCQ N KV Y TN T P G GRSH V VNTGWAM L ALIGAGQAERDPVPLH
BrCAS1	642	CE F LLSKQLPSGGWGESY L SCQ D KV Y SN L EGNRS H VNTAWA L LS L IGAGQVEVDQKPLH
BrCAS2	642	CE F LLSKQLPSGGWGESY L SCQ D KV Y SN L EGNRS H VNTAWA L LS L IGAGQVEVDQKPLH
AtCAS1	551	CE F LLSKQ Q PSGGWGESY L SCQ D KV Y SN L DGNRS H VNTAWAM L ALIGAGQAEVDRKPLH
AtCAS2	642	CE F LLSKQ Q PSGGWGESY L SCQ D KV Y SN L DGNRS H VNTAWAM L ALIGAGQAEVDRKPLH
PgCAS	642	CD F LLSKQ V ASGGWGESY L SCQ N KV Y TN L EGNRS H VNTGWAM L ALIDAGQAERDATPLH
PsCAS	642	CE F LLSKQLPSGGWA E SY L SCQ N KV Y SN L EGNRS H VNTGWAM L ALIEAEQAKRDPPTPLH
LjCAS	642	CE F LLSKQLPSGGWGESY L SCQ N KV Y SN L EGNR P HAVNTGWAM L ALIEAEQAKRDPPTPLH
GgCAS	642	CE F LLSKQLPSGGWGESY L SCQ N KV Y SN V ESNR S HVNTGWAM L ALIDAEQAKRDPPTPLH
consensus	661	*.....**.....*.....*.....*.....*.....*.....*.....*

Tabela 41 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

Protein	Position	Sequence
MiFRS	704	RAAKLLINYSQMEEGDGFPPQOEITGVFKNKNCMLHYPIYRNAFPFIWALGEYRKRVPLPSKSGNS
KdFRS	704	RAAKLLINSQIEDGDFPQOEITAGVFKMNCNLTHTFAAYRNIFPIWALAVYRRFCNPNSEAIS
PdFRS	701	RAAKLLINSQAEEDGSYPQOEITGVFKNKNCMLHYPTYKNVFPFIWALAEYRKRVPLPSKKL
AtSHS	703	RAAKLLINSQMEENGDFPQOEDTIGAFKKNCLLHYPMYRNITLWALAQYRKQVLRQPTGI
BrASS	706	RAAKVILINSQLENGDFPQOEMRGAYKMNLLHYPTYRNMFTLWALTYYTKALRRT----
KdGLS	704	RAAKLLINSQIEDGDFPQOEITGVFKNKNCMLHTFAAYRSIFPFIWALAEYKRFCNLSSEAIS
BpLUS	701	REVRVLINSQMEDGDFPQOEITGVFMRNCTLNYSYRNIFPIWALGEYRRRVILFA----
LjLUS	701	RGIRLLINSQMEDGDFPQOEITGVFMRNCTLNYSYRNIFPIWALGEYRRRVILCA----
GgLUS	701	RCMKLLINSQMEDGDFPQOEITGVFMRNCTLNYSYRNIFPIWALGEYRRQVILCAHSY--
OeLUS	702	RGIKLLINSQMEDGDFPQOEITGAFMKNCNLTNYSYRNIFPIWALGEYRRRIILHAQT---
ToLUS	704	NEVRLLLINSQMEEGDGFPPQOEITGVFKNKNCNLNYSYSRNIFPIWALGEYRRIVQNI-----
KdLUS	704	RAAQLLINSQIEDGDFPQOEITGVFQNRNCLMHYAAYRNIFPIWALAEYRRQIQQLHSEATK
RcLUS	703	RAAKLLINSQTDLGDFPQOELTGAFMRNCLMHYALFRNTFPIWALAEYRRHVLFPSAGFG
BgLUS	703	RAAKLLINSQTELGDFFPQOELSGCFMRNCLMHYSEYRDIFFPTWALAEYCKLFLPLPSKND
GgBAS	704	RAAKLLINSQLEEGDVPQOEITGVFMKNCMLHYPMYRDIYPMWALAEYRRRVPLPSTPVC
PsBAS	704	RAAKLLINSQLEGGDVPQOEITGVFMKNCMLHYPMYRDIYPLWALAEYRRRVPLP----
EtBAS	704	RAAKLLINSQIEDGDFPQOEITGVFMKNCMLHYAAYRNITPLWALAEYRNRVPLPSTTL
BpBAS	704	RAAKLLINSQIEDGDFPQOEITGVFMKNCMLHYAAYKNITPLWALAEYRKHVPLPLGKNL
PgBAS1	705	RAAKLLINSQMEDGDFPQOEITGVFMKNCMLHYAAYRNITPLWALAEYRRRVPLPSLGT-
PgBAS2	703	RAAKLLINSQMESGDFPQOEITGVFMKNCMLHYAASRNITPLWALAEYRKRVRLPSKSV-
MiCAS	707	RAAKLLINSQMEENGDFPQOETIMGVFNKNCMTISYAYRNIFPIWALGVYRCRVLQSL----
LaCAS	710	RAARILINSQMEDGDFPQOETIMGTFNKNCMTISYAAYRNIFPIWALGEYRCRVLQAP----
BpCAS1	712	RAARILINSQMEENGDFPQOETIMGVFNKNCMTISYAYRNIFPIWALGEYRCRVLQAL----
BpCAS2	702	RAARVLINSQMEENGDFPQOETIMGVFNKNCMTITYAAYRNIFPIWALGEYRCRVLQAP----
KdCAS	703	RAAKFLIESQTEENGDFPQOETIMGVFNKNCMTISYAAYRNIFPIWALGEYRCQVILNASRGQM
BrCAS1	702	RAARYLINAQMESGDFPQOETIMGVFNKNCMTITYAAYRNIFPIWALGEYRSKVLLQQGE--
BrCAS2	702	RAARYLINAQMESGDFPQOETIMGVFNKNCMTITYAAYRNIFPIWALGEYRSKVLLQQGE--
AtCAS1	611	RAARYLINAQMEENGDFPQOETIMGVFNKNCMTITYAAYRNIFPIWALGEYRCQVLLQQGE--
AtCAS2	702	RAARYLINAQMEENGDFPQOETIMGVFNKNCMTITYAAYRNIFPIWALGEYRCQVLLQQGE--
PgCAS	702	RAAKLLINSQMEENGDFPQOETIMGVFDKNCMTITYAAYRNIFPIWALGEYRCRVLQGPS---
PsCAS	702	RAAVCLINSQTEENGDFPQOETIMGVFNKNCMTITYAAYRCIFPIWALGEYRRVILQAC----
LjCAS	702	RAALYLINSQMEENGDFPQOETIMGVFNKNCMTITYAAYRSIFPIWALGEYRCRVILQAR----
GgCAS	702	RAAVYLINSQMEENGDFPQOETIMGVFNKNCMTITYAAYRNIFPIWALGEYRHRVILQSQ----
consensus	721	RAAKLLINSQMEENGDFPQOETIMGVFNKNCMTITYAAYRNIFPIWALGEYRHRVILQSQ----

Tabela 42 – Alinhamento múltiplo global de sequências. A última linha evidencia a presença de resíduos conservados. Fonte: ClustalW (EMBL-EBI) e BoxShade (Swiss institute of Bioinformatics).

MiFRS	764	MAMKINSA-----
KdFRS	764	KPSK-----
PdFRS		-----
AtSHS		-----
BrASS		-----
KdGLS	764	KPSK-----
BpLUS		-----
LjLUS		-----
GgLUS		-----
OeLUS		-----
ToLUS		-----
KdLUS	764	MV-----
RcLUS	763	FGFTNNL-----
BgLUS		-----
GgBAS	764	LT-----
PsBAS		-----
EtBAS		-----
BpBAS	764	NQVVNCIGQSLYKKYK
PgBAS1		-----
PgBAS2		-----
MiCAS		-----
LaCAS		-----
BpCAS1		-----
BpCAS2		-----
KdCAS	763	KT-----
BrCAS1		-----
BrCAS2		-----
AtCAS1		-----
AtCAS2		-----
PgCAS		-----
PsCAS		-----
LjCAS		-----
GgCAS		-----
consensus	781	