

UNIVERSIDADE FEDERAL DO PARANÁ

AHMAD KASSEM EL ZEIN

*Monteverdia ilicifolia*: REVISÃO DE ESCOPO DOS ASPECTOS FITOQUÍMICOS,  
FARMACOLÓGICOS E TECNOLÓGICOS E DESENVOLVIMENTO DE MÉTODO  
QUIMIOMÉTRICO PARA SUA IDENTIFICAÇÃO E DIFERENCIAÇÃO DE  
ESPÉCIES SIMILARES

CURITIBA

2024

AHMAD KASSEM EL ZEIN

*Monteverdia ilicifolia*: REVISÃO DE ESCOPO DOS ASPECTOS FITOQUÍMICOS,  
FARMACOLÓGICOS E TECNOLÓGICOS E DESENVOLVIMENTO DE MÉTODO  
QUIMIOMÉTRICO PARA SUA IDENTIFICAÇÃO E DIFERENCIAÇÃO DE  
ESPÉCIES SIMILARES

Dissertação apresentada como requisito para  
obtenção do título de Mestre, Curso de Ciências  
Farmacêuticas, Setor de Ciências da Saúde,  
Universidade Federal do Paraná.

Orientador: Prof. Dr. Roberto Pontarolo.

Coorientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Luana Mota Ferreira

CURITIBA

2024

El Zein, Ahmad Kassem

*Monteverdia ilicifolia* [recurso eletrônico]: revisão de escopo dos aspectos fitoquímicos, farmacológicos e tecnológicos e desenvolvimento de método quimiométrico para sua identificação e diferenciação de espécies similares / Ahmad Kassem El Zein – Curitiba, 2024.

1 recurso online : PDF

Dissertação (mestrado) – Programa de Pós-Graduação em Ciências Farmacêuticas. Setor de Ciências da Saúde, Universidade Federal do Paraná, 2024.

Orientador: Prof. Dr. Roberto Pontarolo

Coorientador: Profa. Dra. Luana Mota Ferreira

1. Celastraceae. 2. Fraude. 3. Quimiometria. 4. Aprendizado de máquina.  
I. Pontarolo, Roberto. II. Ferreira, Luana Mota. III. Universidade Federal do Paraná.  
IV. Título.

CDD 583.67



MINISTÉRIO DA EDUCAÇÃO  
SETOR DE CIÊNCIAS DA SAÚDE  
UNIVERSIDADE FEDERAL DO PARANÁ  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIAS  
FARMACÊUTICAS - 40001018042P8

## TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação CIÊNCIAS FARMACÊUTICAS da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de AHMAD KASSEM EL ZEIN intitulada: *Monteverdia ilicifolia* : REVISÃO DE ESCOPO DOS ASPECTOS FITOQUÍMICOS, FARMACOLÓGICOS E TECNOLÓGICOS E DESENVOLVIMENTO DE MÉTODO QUIMIOMÉTRICO PARA SUA IDENTIFICAÇÃO E DIFERENCIAÇÃO DE ESPÉCIES SIMILARES, sob orientação do Prof. Dr. ROBERTO PONTAROLO, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua aprovação para o rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 20 de Setembro de 2024.

ROBERTO PONTAROLO

Presidente da Banca Examinadora

GUSTAVO BERTOL

Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ)

JANE MANFRON BUDEL

Avaliador Externo (UNIVERSIDADE ESTADUAL DE PONTA GROSSA)

## AGRADECIMENTOS

Primeiramente agradeço ao meu orientador, professor doutor Roberto Pontarolo, por ter me dado a oportunidade de desenvolver o presente estudo, pela confiança, apoio, espaço para pôr em prática minhas ideias, ensinamentos valiosos ao longo desses anos e por seu enorme coração. À minha coorientadora, professora doutora Luana Mota Ferreira, que aceitou o desafio de coorientar este trabalho e me auxiliou neste processo de construção, com paciência e muito carinho.

À minha família, Mariam, Márcia e Kassem, pela vivência e por me ensinarem que o importante não é evoluir para vencer os outros, mas sim se tornar melhor que o “eu” de antes. Por sempre estarem ao meu lado, pela confiança em mim até nos piores momentos, pelo amor, carinho, compreensão e presença na minha vida. Também agradeço aos familiares mais distantes por todo amor, companheirismo e momentos de felicidade.

À minha namorada, Milena, que ao longo desses três anos sempre esteve disponível para me ajudar em tudo que fosse necessário, pela sobrecarga de amor, apoio, carinho, ideias e toda felicidade que é estar ao seu lado. Sem a tua presença tudo seria mais difícil e me sinto lisonjeado por tê-la ao meu lado.

Aos meus amigos, com os quais converso todos os dias, que sempre estão presentes e proporcionam deliciosas risadas, momentos descontraídos, ensinamentos, apoio, amor e muito companheirismo.

A todo o pessoal do CEB e GEATS pelo companheirismo, auxílio nas mais árduas tarefas, conselhos, ensinamentos, momentos de união e vivência. Um agradecimento especial ao Alexandre e ao Raul pelo incentivo, disponibilidade e capacidade de me ajudar ao longo das diversas atividades.

Menciono também todos aqueles mais distantes que, de alguma forma, fizeram-se presentes na minha vida, compartilharam de alguns momentos ao longo dos anos, e me auxiliaram ou não no desenvolvimento do presente trabalho.

*“A recompensa do jovem cientista é a emoção de ser a primeira pessoa na história do mundo a ver algo ou a entender algo. Nada se compara a essa experiência... A recompensa do velho cientista é a sensação de ter visto um esboço vago se transformar em uma paisagem magistral.”*

*(Cecilia Payne-Gaposchkin, 1977)*

## RESUMO

Compondo a extensa flora da América do Sul, *Monteverdia ilicifolia* (Mart. ex Reissek) Biral, popularmente conhecida como espinheira-santa (Celastraceae), é reconhecida por seus diversos efeitos terapêuticos. A popularização desta espécie favoreceu a adulteração/falsificação, mas pouco contribuiu para o estudo e comprovação de suas propriedades medicinais. O presente estudo teve como objetivo realizar uma revisão de escopo para mapear as evidências em relação aos aspectos fitoquímicos, farmacológicos e tecnológicos envolvendo *Monteverdia ilicifolia* (Mart. ex Reissek) Biral. Foi também desenvolvido um método quimiométrico e de aprendizado de máquina (*machine learning*), para sua identificação e discriminação de espécies morfológicamente similares. A revisão abordou estudos pré-clínicos, clínicos e perspectivas tecnológicas, visando responder à questão “Quais são as evidências científicas sobre *M. ilicifolia* com relação aos seus aspectos fitoquímicos, farmacológicos e tecnológicos?”. A pesquisa foi conduzida seguindo as diretrizes JBI (Instituto Joanna Briggs) e recomendações do PRISMA-ScR. Os artigos foram buscados nas plataformas PubMed, Scopus e *Web of Science*. As patentes foram extraídas das plataformas *Google Patents*, Espacenet, *Derwent Innovations Index*, e *The Lens*. Os estudos foram avaliados conforme os critérios de inclusão e exclusão estabelecidos. Artigos e patentes aprovados passaram por uma etapa de extração de informações, enquanto aqueles que não se adequaram foram excluídos, com os motivos de remoção justificados. Um total de 13 artigos e 3 patentes foram aprovados. Estes estudos indicam que *M. ilicifolia* pode ser utilizada na forma de extratos aquosos, alcoólicos e hidroalcoólicos, formulados como gel, creme, gel-creme e soluções parenterais, com doses de 0,1 a 500 mg/kg, uma a cinco vezes no dia, sendo segura até 3000 mg/kg. Suas propriedades terapêuticas incluem atividades antimicrobiana, antiulcerativa, espasmolítica, antidiarreica, anti-inflamatória, natriurética, diurética, hipotensiva e antimutagênica, sem sinais de toxicidade ou teratogenicidade. Esses efeitos foram atribuídos aos metabólitos triterpenos, taninos, flavonoides, ácido gálico e cafeico, e outros compostos polifenólicos. Este trabalho também abordou o desenvolvimento de um método quimiométrico realizado de forma exploratória, por meio da Análise de Componentes Principais (PCA), e discriminatória, pela Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA). O estudo foi baseado em dados na forma de espectros de infravermelho médio com transformada de Fourier das folhas, pós e extratos de *M. ilicifolia* e espécies semelhantes utilizadas como adulterantes. Todas as espécies vegetais foram individualmente identificadas pelo método PCA, e a discriminação pelo PLS-DA foi mais eficiente para amostras de extrato, demonstrando sensibilidade, especificidade e acurácia superiores (94, 100 e 99%, respectivamente) aos demais modelos; apesar disso, os métodos também foram satisfatórios para análise de pós e folhas. Complementando a quimiometria, um modelo *machine learning* foi desenvolvido pela metodologia *Random Forest* e validação cruzada 10 *Fold*. Os resultados foram promissores, com alta acurácia para todas as formas de apresentação, baixo número de falsos positivos e excelentes resultados para as métricas de acurácia, *Recall*, precisão, *F1 Score*, índice *Kappa* e Coeficiente de Correlação de Matthews (MCC). Os modelos criados apresentam potencial de utilização em amostras comerciais, identificando adulterações e falsificações de produtos à base de espinheira-santa.

**Palavras-chave:** *Monteverdia ilicifolia*; *Maytenus ilicifolia*; *Monteverdia aquifolium*; *Sorocea bonplandii*; falsificação botânica; autenticação por quimiometria; *machine learning*.

## ABSTRACT

Present on the extensive flora of South America, *Monteverdia ilicifolia* (Mart. ex Reissek) Biral, popularly known as espinheira-santa (Celastraceae), is recognized for its broad-spectrum healing effects. The popularization of this plant species has led to it becoming a target of adulteration/falsification, while contributing less to the study and validation of its medicinal properties. This study aimed to conduct a scoping review to map the evidence on the phytochemical, pharmacological, and technological aspects of *Monteverdia ilicifolia* (Mart. ex Reissek) Biral, as well as to develop a chemometric and machine learning method for its identification and differentiation from morphologically similar species. The review covered pre-clinical and clinical uses, and technological perspectives, with the aim of answering the question "What are the scientific evidences on *M. ilicifolia* in terms of its phytochemical, pharmacological, and technological aspects?". The research was conducted following the Joanna Briggs Institute guidelines and the recommendations of PRISMA-ScR. Articles were searched for in the PubMed, Scopus and Web of Science databases. Patents were extracted from the Google Patents, Espacenet, Derwent Innovations Index and The Lens databases. The studies were assessed according to the established inclusion and exclusion criteria. Approved articles and patents underwent an information extraction stage, while those that did not fit were excluded, with the reasons for their removal justified. 13 articles and 3 patents were approved and indicated that *M. ilicifolia* can be used in the form of aqueous, alcoholic and hydroalcoholic extracts, gel, cream, gel-cream and parenteral solutions, at doses of 0.1 to 500 mg/kg, one to five times a day, with up to 3000 mg/kg being safe. Its therapeutic properties include antimicrobial, antiulcer, spasmolytic, antidiarrheal, anti-inflammatory, natriuretic, diuretic, hypotensive and antimutagenic activities, with no signs of toxicity or teratogenicity. These effects were attributed to the metabolites triterpenes, tannins, flavonoids, gallic and caffeic acid, and other polyphenolic compounds. In a second approach, this work presents chemometric analysis, which carried out in an exploratory way, using Principal Component Analysis (PCA), and in a discriminatory way, using Partial Least Squares Discriminant Analysis (PLS-DA). The study was based on data in the form of Fourier transform mid-infrared spectra of the leaves, powders and extracts of *M. ilicifolia* and similar species used as adulterants. All plant species were individually identified by the PCA method, and the discrimination by PLS-DA was much more efficient for extract samples, showing higher sensitivity, specificity and accuracy (94, 100 and 99%, respectively) than the other models; despite this, the method were also satisfactory for powders and leaves. Complementing the chemometrics approach, the machine learning model was developed using the Random Forest methodology and 10 Fold cross-validation. The results were promising, with high accuracy for all forms of presentation, a low number of false positives and excellent results for the metrics of accuracy, recall, precision, F1 Score, *Kappa* index and Matthews Correlation Coefficient (MCC). The models created have the potential to be used in commercial samples, identifying adulteration and counterfeiting of products based on espinheira-santa.

**Key-words:** *Monteverdia ilicifolia*; *Maytenus ilicifolia*; *Monteverdia aquifolium*; *Sorocea bonplandii*; botanical adulteration; authentication by chemometrics; machine learning;

## LISTA DE SIGLAS E ABREVIATURAS

RDC	Resolução de Diretoria Colegiada
MIR-FTIR	Espectroscopia de infravermelho com transformada de Fourier
PCA	Análise de Componentes Principais
PLS-DA	Análise Discriminante por Mínimos Quadrados Parciais
RENAME	Relação Nacional de Medicamentos Essenciais
UV	Ultravioleta
ROS	Espécies reativas de oxigênio
IL-6	Interleucina-6
IL-1	Interleucina-1
TNF- $\alpha$	Fator de necrose tumoral-alfa
GM/MS	Ministério da Saúde Gabinete do Ministro
FB 6 <sup>a</sup> edição	Farmacopeia Brasileira 6 <sup>a</sup> edição
SUS	Sistema Único de Saúde
IN	Instrução Normativa
CCD	Cromatografia em camada delgada
IR	Radiação infravermelha
IR-A	Infravermelho próximo
IR-B	Infravermelho médio
IR-C	Infravermelho longo
LAEMIL	Extrato aquoso liofilizado das folhas de <i>M. ilicifolia</i>
SEI	Sobrenadante etanólico da infusão
Fr.H <sub>2</sub> O	Fração aquosa do sobrenadante etanólico da infusão
Fr.EtOAc	Fração orgânica do sobrenadante etanólico da infusão
MIE	Extrato de <i>M. ilicifolia</i>
CRIB	Células derivadas do modelo Mardin-Darby de fígado bovino
CER	Células relacionadas ao embrião e galinha
L929	Linha celular de fibroblastos de camundongos
CRFK	Linha celular de fígado felino
BHV-5	Herpesvírus bovino tipo 5
TLR2	Receptor <i>toll-like 2</i>

LTA	Ácido lipoteicoico
ULI	Índice de lesão ulcerativa
CF	Fração à base de clorofórmio
EAF	Fração à base acetato de etila
BuF	Fração à base de butanol
AqF	Fração à base de água
InsF	Fração insolúvel à base de acetato de etila
solF	Fração solúvel à base de acetato de etila
ES70	Extrato com espinheira-santa à 70%
ES95	Extrato com espinheira-santa à 95%
MBC	Concentração bactericida mínima
MIC	Concentração inibitória mínima
MHV-3	Betacoronavírus
BRSV	Vírus sincicial respiratório bovino
LTA	Ácido lipotecóico
TEER	Resistência elétrica transepitelial
iNOS	Óxido-nítrico sintase indutível
COX-2	Ciclooxigenase-2
MIA	<i>M. ilicifolia</i> no acesso A
MIB	<i>M. ilicifolia</i> no acesso B
TBARS	Substâncias reativas ao ácido tiobarbitúrico
4NQO	<i>4-oxide-1-nitroquinoline</i>
2NF	<i>2-nitrofluorene</i>
AZS	Azida de sódio
Fr.EtOAc	Infusões etanólicas
Fr.H2O	Infusões aquosas
HCTZ	Hidroclorotiazida
DDA	<i>2',5'-dideoxyadenosine</i>
LD50	Dose letal mediana
DPPH	Teste 2,2-diphenyl-1-picrylhydrazyl
ORACFL	Capacidade de Absorção do Radical de Oxigênio com fluoresceína.
CLAE/HPLC	Cromatografia Líquida de alta eficiência
RMSEC	<i>Root Mean Square Error of Calibration</i>

RMSECV	<i>Root Mean Square Error of Cross-Validation</i>
RMSEP	<i>Root Mean Square Error of Prediction</i>
MCC	Coeficiente de Correlação de Matthews
AUC ROC	Gráficos de Área sob Curva ROC
E1	<i>Monteverdia ilicifolia</i> (Mart. ex Reissek) Biral
E2	<i>Monteverdia aquifolium</i> (Mart.) Biral
E3	<i>Sorocea bonplandii</i> (Baill.) W.C. Burger et al.
E4	<i>Citronella gongonha</i> (Mart.) R.A. Howard
E5	<i>Jodina rhombifolia</i> (Hook & Arn.) Reissek
E6	<i>Zollernia ilicifolia</i> (Brongn.) Vogel
GLS W	<i>Generalized Least Squares Weighting</i>
VLs	Variáveis latentes
CV%	Coeficiente de variação

## LISTA DE FIGURAS

Figura 1 - Aparência das folhas de <i>Monteverdia ilicifolia</i> (Mart. ex Reissek) Biral (A, B) e sua distribuição geográfica (C) .....	26
Figura 2 – Estrutura química dos taninos hidrolisáveis e condensados .....	27
Figura 3 – Estrutura química do pirogalol.....	28
Figura 4 - Estrutura molecular básica dos terpenos .....	28
Figura 5 – Estrutura química básica de um flavonoide.....	29
Figura 6 – Estrutura química da epicatequina .....	30
Figura 7 – Características macro e microscópicas de folhas de <i>Monteverdia ilicifolia</i> (Mart. ex Reissek) Biral .....	37
Figura 8 - Região do espectro magnético contendo o infravermelho .....	39
Figura 9 – Características dos artigos incluídos no estudo .....	52
Figura 10 - Estrutura molecular dos principais metabólitos secundários presentes na espinheira-santa.....	67
Figura 11 – Efeito dos extratos de <i>M. ilicifolia</i> a base de hexano e acetato de etila sobre o edema induzido em cobaias.....	74
Figura 12 – Níveis de glucose sanguínea (mg.dl-1) em ratos hiperglicêmicos após tratamento com diferentes doses de extratos etanólicos de <i>M. ilicifolia</i> .....	76
Figura 13 – Resultados dos ensaios de antimutagenicidade em revertantes/placa a partir de <i>M. ilicifolia</i> e <i>P. peltatus</i> contra mutagênicos conhecidos sem ativação metabólica.....	78
Figura 14 - Características dos artigos incluídos no estudo .....	84
Figura 15 – Método de corte para análise completa da folha de espinheira-santa ...	98
Figura 16 – Padrão espectroscópico de folhas íntegras de <i>M. ilicifolia</i> .....	104
Figura 17 – Padrão espectroscópico dos pós de <i>M. ilicifolia</i> .....	105
Figura 18 – Padrão espectroscópico os extratos de <i>M. ilicifolia</i> .....	106
Figura 19 – Sobreposição do perfil espectral das formas de apresentação estudadas para <i>M. ilicifolia</i> .....	107
Figura 20 - Análise gráfica da escolha do número de componentes principais.....	109
Figura 21 - Representação tridimensional do modelo PCA desenvolvido para folhas íntegras .....	110
Figura 22 – Gráfico variáveis/ <i>Loading Plots</i> para as folhas íntegras.....	111

Figura 23 - Análise gráfica para a escolha do número de componentes principais para o modelo PCA dos pós.....	112
Figura 24 - Representação tridimensional do modelo PCA desenvolvido para os pós .....	113
Figura 25 - Gráfico variáveis/ <i>Loading Plots</i> para os pós .....	114
Figura 26 - Análise gráfica para escolha do número de componentes principais para o modelo PCA dos extratos.....	114
Figura 27 - Representação tridimensional do modelo PCA para os extratos .....	115
Figura 28 - Gráfico variáveis/ <i>Loading plots</i> para os extratos .....	116
Figura 29 – Gráfico de avaliação das variáveis latentes para amostras de folhas ..	117
Figura 30 – Gráfico <i>Student Residual</i> para amostras na forma de folhas .....	118
Figura 31 – Gráfico Y predito para avaliação da capacidade discriminatória do modelo para as folhas, com foco em <i>M. ilicifolia</i> .....	120
Figura 32 – Gráfico VIP para a identificação das variáveis de maior relevância para discriminação das espécies na forma de folhas .....	123
Figura 33 – Perfil VIP individual para folhas de <i>M. ilicifolia</i> .....	124
Figura 34 – Gráfico de avaliação das variáveis latentes para amostras em pó.....	125
Figura 35 – Gráfico <i>Student Residual</i> para identificação de outliers para as amostras na forma de pó .....	126
Figura 36 – Gráfico Y predito com foco em <i>M. ilicifolia</i> , avaliando capacidade preditiva do modelo para amostras em pó .....	127
Figura 37 – Gráfico VIP para a identificação das variáveis de maior relevância para discriminação das espécies na forma de pós.....	129
Figura 38 – Perfil VIP individual para <i>M. ilicifolia</i> em pós .....	130
Figura 39 – Gráfico de avaliação das variáveis latentes para amostras em extratos .....	131
Figura 40 - <i>Student Residual</i> para identificação de outliers para as amostras na forma de extratos .....	131
Figura 41 – Gráfico Y predito com foco em <i>M. ilicifolia</i> , avaliando capacidade preditiva do modelo para amostras em extrato .....	132
Figura 42 – Gráfico VIP para a identificação das variáveis de maior relevância para discriminação das espécies na forma de extratos.....	134
Figura 43 – Perfil VIP individual para <i>M. ilicifolia</i> em extratos .....	135

Figura 44 – Progressão da acurácia da calibração e validação do modelo <i>machine learning</i> para amostras na forma de folha .....	138
Figura 45 – Matriz de confusão do modelo <i>machine learning</i> para amostras na forma de folha .....	140
Figura 46 – Gráfico de erro de classificação para a amostra de folhas.....	141
Figura 47 – Área sob a curva ROC para a amostra de folhas.....	142
Figura 48 – Gráfico do relatório de classificação para a validação do modelo para amostras na forma de folha.....	144
Figura 49 – Números de onda mais importantes na classificação das diferentes espécies vegetais na forma de folha .....	145
Figura 50 – Progressão da acurácia da calibração e validação do modelo <i>machine learning</i> para amostras na forma de pó.....	147
Figura 51 – Matriz de confusão do modelo <i>machine learning</i> para amostras na forma de pó .....	148
Figura 52 – Gráfico de erro de classificação para a amostra de pó .....	149
Figura 53 – Área sob a curva ROC para a amostra de pó .....	150
Figura 54 – Gráfico do relatório de classificação para a validação do modelo para amostras na forma de pó .....	152
Figura 55 – Progressão da acurácia da calibração e validação do modelo <i>machine learning</i> para amostras na forma de extrato.....	153
Figura 56 – Matriz de confusão do modelo <i>machine learning</i> para amostras na forma de extrato .....	155
Figura 57 – Gráfico de erro de classificação para a amostra de extrato .....	156
Figura 58 – Área sob a curva ROC para a amostra de extrato .....	156
Figura 59 – Gráfico do relatório de classificação para a validação do modelo para amostras na forma de extrato.....	158
Figura 60 – Números de onda mais importantes na classificação das diferentes espécies vegetais na forma de extrato.....	159

## LISTA DE QUADROS

Quadro 1- Formas farmacêuticas reconhecidas pela RENAME para tratamentos com <i>Monteverdia ilicifolia</i> (Mart. ex Reissek) Biral .....	32
Quadro 2 - Informações extraídas dos registros na forma de artigos.....	53
Quadro 3 – Compostos encontrados em <i>M. ilicifolia</i> .....	62
Quadro 4 - Estratégia de análise das folhas e pós para plantas adquiridas em mais de uma localidade .....	99
Quadro 5 - Estratégia de análise dos extratos para plantas adquiridas em mais de uma localidade .....	100
Quadro 6 – Avaliação das amostras do gráfico Y predito do modelo PLS-DA para as folhas.....	121
Quadro 7 - Cálculo das métricas de qualidade do modelo para as folhas .....	121
Quadro 8 – Avaliação das amostras do gráfico Y predito do modelo PLS-DA para os pós .....	127
Quadro 9 – Avaliação das amostras do gráfico Y predito e o cálculo das métricas de qualidade do modelo PLS-DA de extratos .....	133
Quadro 10 – Métricas calculadas para o método <i>machine learning</i> das folhas.....	143
Quadro 11 – Métricas calculadas para o método <i>machine learning</i> dos pós .....	151
Quadro 12 – Métricas calculadas para o método <i>machine learning</i> dos extratos ...	157

## LISTA DE TABELAS

Tabela 1 - Espécies adulterantes da <i>M. ilicifolia</i> .....	34
Tabela 2 – Estratégia de busca para os artigos .....	45
Tabela 3 – Estratégia de busca para as patentes .....	46
Tabela 4 – Efeitos da administração de friedelan-3- $\beta$ -ol e friedelina.....	68
Tabela 5 – Efeito comparativo das frações solF e insF quanto a motilidade gastrointestinal das cobaias .....	70
Tabela 6 - Principais informações extraídas das patentes .....	85
Tabela 7 - Número de amostras para cada espécie analisada .....	95

## **LISTA DE FLUXOGRAMAS**

Fluxograma 1 - Fluxograma para revisão de escopo (artigos) .....	49
Fluxograma 2 - Fluxograma para revisão de escopo (patentes) .....	50

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	21
<b>2. OBJETIVOS</b> .....	24
2.1. OBJETIVOS ESPECÍFICOS .....	24
<b>3. REVISÃO BIBLIOGRÁFICA</b> .....	25
3.1. ESPINHEIRA-SANTA .....	25
<b>3.1.1. Características gerais</b> .....	25
<b>3.1.2. Composição fitoquímica</b> .....	27
3.1.2.1 Taninos .....	27
3.1.2.2 Triterpenos.....	28
3.1.2.3 Flavonoides.....	29
<b>3.1.3 Efeitos farmacológicos</b> .....	30
<b>3.1.4 Produtos disponíveis no mercado</b> .....	31
<b>3.1.5 Espécies vegetais usadas como adulterantes de <i>Monteverdia ilicifolia</i> (Mart. ex Reissek) Biral</b> .....	32
<b>3.1.6 Métodos de identificação</b> .....	37
3.2 ESPECTROSCOPIA DE INFRAVERMELHO.....	38
<b>4 CAPÍTULO 1</b> .....	42
4.1 DESCRIÇÃO GERAL .....	42
4.2 OBJETIVOS.....	44
<b>4.2.1 Objetivos específicos</b> .....	44
4.3 METODOLOGIA .....	45
<b>4.3.1 Critérios de elegibilidade</b> .....	47
<b>4.3.2 Critérios de inclusão</b> .....	47
<b>4.3.3 Critérios de exclusão</b> .....	47
<b>4.3.4 Gerenciamento dos resultados</b> .....	47
<b>4.3.5 Processo de seleção</b> .....	48
<b>4.3.6 Extração dos dados</b> .....	48

4.4	RESULTADOS E DISCUSSÕES .....	49
<b>4.4.1</b>	<b>Inovações científicas.....</b>	<b>51</b>
4.4.1.1	Efeitos gástricos.....	68
4.4.1.2	Atividade antimicrobiana e antiviral .....	71
4.4.1.3	Atividade anti-inflamatória .....	73
4.4.1.4	Efeito hipoglicemiante .....	76
4.4.1.5	Atividade antimutagênica .....	77
4.4.1.6	Efeito anti-hipertensivo.....	79
4.4.1.7	Fertilidade .....	80
4.4.1.8	Avaliação da segurança .....	81
<b>4.4.2</b>	<b>Prospecção tecnológica.....</b>	<b>83</b>
4.5	CONCLUSÃO .....	88
<b>5</b>	<b>CAPÍTULO 2 .....</b>	<b>90</b>
5.1	DESCRIÇÃO GERAL.....	90
<b>5.1.1</b>	<b>Quimiometria .....</b>	<b>90</b>
<b>5.1.2</b>	<b><i>Machine Learning</i> .....</b>	<b>92</b>
5.2	OBJETIVOS.....	94
<b>5.2.1</b>	<b>Objetivos específicos .....</b>	<b>94</b>
5.3	MATERIAIS E MÉTODOS.....	95
<b>5.3.1</b>	<b>Amostras para o desenvolvimento do modelo quimiométrico .....</b>	<b>95</b>
<b>5.3.2</b>	<b>Preparação das folhas.....</b>	<b>96</b>
<b>5.3.3</b>	<b>Preparação dos pós .....</b>	<b>97</b>
<b>5.3.4</b>	<b>Preparação dos extratos .....</b>	<b>97</b>
<b>5.3.5</b>	<b>Espectroscopia por infravermelho.....</b>	<b>97</b>
<b>5.3.6</b>	<b>Análise dos espectros obtidos .....</b>	<b>100</b>
<b>5.3.7</b>	<b>Desenvolvimento do método quimiométrico.....</b>	<b>100</b>
<b>5.3.8</b>	<b>Critérios de avaliação para os modelos.....</b>	<b>102</b>
<b>5.3.9</b>	<b><i>Machine learning</i> .....</b>	<b>102</b>
5.4	RESULTADOS E DISCUSSÃO.....	104

<b>5.4.1</b>	<b>Modelo PCA</b> .....	108
5.4.1.1	Folhas .....	108
5.4.1.2	Pós.....	112
5.4.1.3	Extratos.....	114
<b>5.4.2</b>	<b>Modelo PLS-DA</b> .....	116
5.4.2.1	Folhas .....	116
5.4.2.2	Pós.....	124
5.4.2.3	Extratos.....	130
<b>5.4.3</b>	<b><i>Machine Learning</i></b> .....	136
5.4.3.1	Folhas .....	137
5.4.3.2	Pós.....	146
5.4.3.3	Extratos.....	153
5.5	CONCLUSÃO .....	161
<b>6</b>	<b>REFERÊNCIAS</b> .....	163

## 1. INTRODUÇÃO

A espécie vegetal *Monteverdia ilicifolia* (Mart. ex Reissek) Biral, também conhecida como espinheira-santa (Celastraceae), é uma planta encontrada principalmente na região sul do Brasil e em países vizinhos como Paraguai, Uruguai e Argentina, sendo extremamente difundida para fins medicinais (Simões *et al.*, 2017). Seu nome popular é originado da presença de espinhos na borda de suas folhas aliado a seus efeitos curativos promissores. As partes usadas da espinheira-santa são as raízes, cascas e principalmente as folhas que estão descritas em monografia individual presente na Farmacopeia Brasileira 6ª Edição. É utilizada principalmente na forma de infusão. Esta espécie vegetal também pode ser encontrada na forma de cápsulas, tinturas, emulsões, entre outros (Brasil, 2019; Silva *et al.*, 2023). Os efeitos terapêuticos característicos da espinheira-santa são antiulcerativo, cicatrizante, anti-inflamatório e analgésico. Todas essas funções são atribuídas aos principais metabólitos encontrados na espécie vegetal, que consistem de taninos, triterpenos e flavonoides (Jorge *et al.*, 2004; Simões *et al.*, 2017).

Por conta de sua popularidade, a espinheira-santa se tornou alvo de falsificações e/ou adulterações visando vantagens econômicas e a sua extensa exploração compromete a preservação das espécies nativas. Como consequência, a espinheira-santa acaba sendo substituída por outras espécies sem eficácia terapêutica comprovada e/ou aparência semelhante, são exemplos de adulterantes *Monteverdia aquifolium* (Mart.) Biral, *Sorocea bonplandii* (Baill.) W.C. Burger *et al.*, *Citronella gongonha* (Mart.) R.A.Howard, *Jodina rhombifolia* (Hook. & Arn.) Reissek, *Zollernia ilicifolia* (Brongn.) Vogel (Castro; Ramos, 2003; Bochner *et al.*, 2012; Antunes *et al.*, 2023a).

A espinheira-santa é comercializada na forma de fitoterápicos e plantas medicinais. Os fitoterápicos obtidos a partir de plantas medicinais têm origem industrial ou magistral. As formulações industriais devem, obrigatoriamente, possuir registro na Agência Nacional de Vigilância Sanitária (ANVISA), enquanto os magistrais não têm essa exigência, mas seu local de produção deve cumprir com as boas práticas de manipulação de medicamentos. Atualmente, existem 24 produtos registrados à base de espinheira-santa, dos quais 6 encontram-se em situação regular, mas apenas 4 são comercializados. A Resolução de Diretoria Colegiada

(RDC) nº. 14, de 31 de março de 2010, regulamenta esta classe de medicamentos, estabelecendo seus requisitos mínimos de segurança, eficácia e comercialização. As plantas medicinais, na forma de droga vegetal, por sua vez, são de uso tradicional, mas não apresentam registro, padronização de dosagem, nem forma de administração estipulada. Em convergência com as dificuldades anteriormente mencionadas, nota-se que no Brasil há falta de rigor na fiscalização dos locais de aquisição destas plantas medicinais, os ervanários. A crescente comercialização de produtos sem eficácia comprovada e sua baixa qualidade resultam no insucesso de tratamentos à base da espinheira-santa e de outras plantas medicinais. Além disso, há uma escassez de literatura científica que comprove a eficácia e segurança da espinheira-santa, sendo seu uso frequentemente baseado na etnofarmacologia, proveniente dos povos originários do Brasil, e em evidências clínicas fracas. A falta de padronização dos medicamentos à base dessa espécie vegetal também amplia as dificuldades relacionadas à fiscalização dos diversos produtos comercialmente disponíveis. Tendo em vista o panorama atual é necessário um rigoroso controle sobre o processo de produção de fitoterápicos, seus respectivos registros e locais de comercialização das espécies vegetais (Brasil, 2010; Hasenclever, 2017; Brasil, 2022a).

Considerando a existência de falsificações e adulterações destaca-se a importância de métodos analíticos para a identificação e diferenciação da espinheira-santa de outras plantas utilizadas na falsificação. Este desafio pode ser contornado com a utilização da espectroscopia de infravermelho com transformada de Fourier (MIR-FTIR) aliada a quimiometria. Esta técnica caracteriza-se por ser uma metodologia rápida, eficiente, verde, capaz de diferenciar a composição química de diferentes amostras em função dos espectros obtidos quando estas são submetidas a radiação infravermelha nas diversas frequências (Biancolillo; Marini, 2018; Veerasingam *et al.*, 2020). A quimiometria é a ciência que relaciona a matemática e estatística à química para extrair informações relevantes a partir de dados gerados por diferentes métodos de análise. Estas informações são obtidas a partir da construção de modelos matemáticos de natureza empírica (Biancolillo; Marini, 2018; Ferreira, 2015; Kowalczyk *et al.*, 2015; Lopes; Neto, 2018).

A técnica de aprendizado de máquina (*machine learning*) tem como base a evolução constante de algoritmos matemáticos, alimentado continuamente por novos

dados. Os algoritmos de aprendizado de máquina podem ser aplicados para complementar à quimiometria na identificação de espécies vegetais buscando métodos ainda mais robustos. Esta metodologia facilita o entendimento de matrizes de resultados complexas, transformando-as em um modelo preciso e eficaz, permitindo previsões e associações de dados adicionados posteriormente à matriz (Janiesch; Zschech; Heinrich, 2020; Nichols; Chan; Baker, 2018; Sarker, 2021).

Com este trabalho buscou-se identificar e discriminar *M. ilicifolia* (espinheira-santa oficial) das espécies *Monteverdia aquifolium*, *Sorocea bonplandii*, *Citronella gongonha*, e *Jodina rhombifolia* e *Zollernia ilicifolia*, as quais são consideradas espécies vegetais adulterantes, não reconhecidas como espinheira-santa pela Farmacopeia Brasileira 6ª edição. Para isso, utilizou-se uma técnica considerada verde, a espectroscopia de infravermelho médio com transformada de Fourier (MIR-FTIR). Os espectros gerados passaram por estudo estatístico com aplicação da quimiometria a nível exploratório, utilizando a Análise de Componentes Principais (PCA) e discriminatório por meio da Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA), identificando e distinguindo as espécies a nível analítico. Algoritmos de aprendizado de máquina também foram aplicados visando complementar os modelos quimiométricos, e permitir a previsão e classificação de amostras comerciais.

## 2. OBJETIVOS

Mapear as evidências em relação aos aspectos fitoquímicos, farmacológicos e tecnológicos envolvendo *Monteverdia ilicifolia* (Mart. ex Reissek) Biral, e construir um método quimiométrico e de aprendizado de máquina para identificá-la e distingui-la de adulterantes.

### 2.1. OBJETIVOS ESPECÍFICOS

- Realizar uma revisão de escopo à cerca dos aspectos fitoquímicos, farmacológicos, pré-clínicos e clínicos de *M. ilicifolia*, por meio da busca de artigos e patentes;
- Buscar e registrar os atuais avanços tecnológicos que envolvem a espinheira-santa por meio de patentes, informando suas formas de apresentação e fins terapêuticos em desenvolvimento;
- Coletar e preparar amostras representativas das espécies de espinheira-santa e espécies contaminantes mais comuns para análise por espectroscopia MIR-FTIR;
- Desenvolver métodos de identificação e discriminação baseados na técnica quimiométrica (PCA e PLS-DA) e aprendizado de máquina;
- Validar os métodos propostos e aplicá-los para a análise de amostras comerciais de espinheira-santa.

### 3. REVISÃO BIBLIOGRÁFICA

#### 3.1. ESPINHEIRA-SANTA

##### 3.1.1. Características gerais

Celastraceae é o nome dado à família botânica que agrega cerca de 100 gêneros. Desses, 22 podem ser encontrados no Brasil, com destaque a *Monteverdia*. A família Celastraceae se destaca por agregar 1410 espécies que se encontram distribuídas em regiões com clima tropical e subtropical, com diversas dessas apresentando significância fitoterapêutica (Nunez *et al.*, 2024).

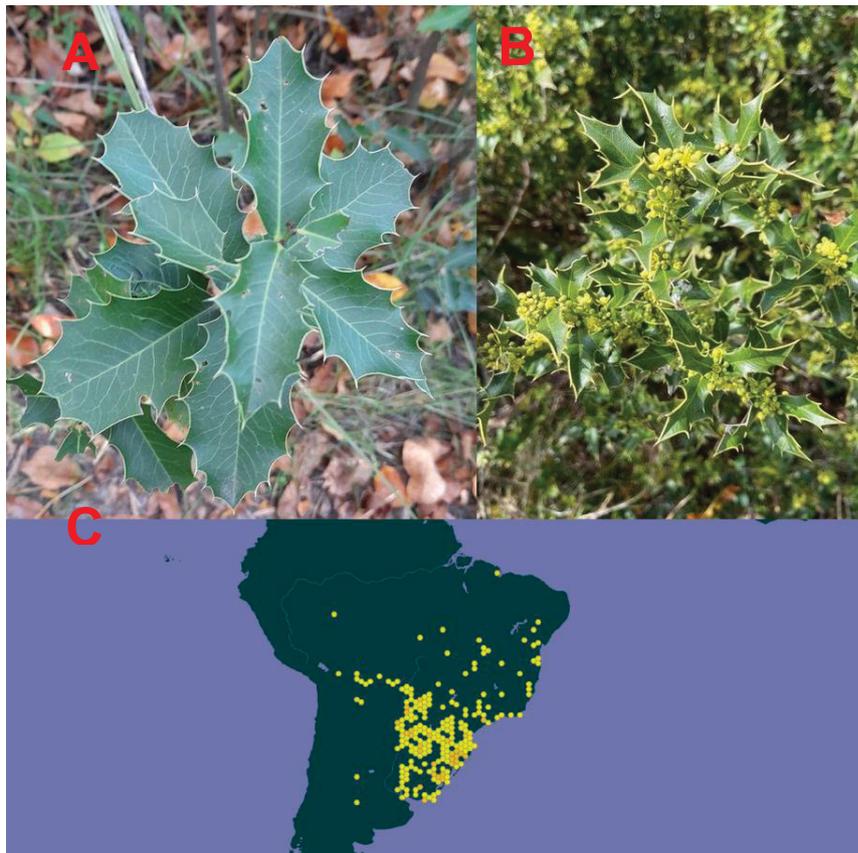
O gênero *Monteverdia* agrega cerca de 120 espécies, com algumas destas ainda em processo de transição de nomenclatura, passando do gênero *Maytenus* para este em questão. As espécies vegetais de *Monteverdia* são amplamente distribuídas ao longo de regiões com clima neotropical, como a Mata atlântica, Cerrado, Savana, entre outros. Dentro desse gênero, destaca-se a *Monteverdia ilicifolia* (Mart. ex Reissek) Biral, de sinonímia *Maytenus ilicifolia* Mart. ex Reissek, conhecida popularmente como espinheira-santa e cangorosa. Usada na medicina tradicional principalmente na forma de infusão a partir de suas folhas, é classificada como uma espécie nativa, mas não endêmica do Brasil. Pode ser encontrada nos biomas de Mata Atlântica, Pampa, Cerrado e Pantanal, habitando os estados de Goiás, Mato Grosso do Sul, Mato Grosso, São Paulo, Paraná, Rio Grande do Sul e Santa Catarina (Biral *et al.* 2017; Biral, Lombardi, 2024).

Morfologicamente, essa espécie vegetal apresenta ramos multicarenados e angulosos, com folhas pecioladas e em formato helicoidal e com 5 a 15 projeções pontiagudas distribuídas, geralmente, em ambos os lados da lâmina (**Figura 1**), característica responsável pela nomenclatura desta espécie vegetal junto aos seus efeitos curativos (Brasil, 2019; Simões *et al.*, 2017).

Microscopicamente, sua folha apresenta mesofilo dorsiventral, com estômatos laterocíticos e epiderme uniestratificada coberta por cutícula espessa. O parênquima paliçádico apresenta células alongadas, enquanto o parênquima esponjoso possui células mais compactadas. Na nervura principal é possível verificar um feixe vascular único, colateral e circundado por uma bainha parenquimática. As células epidérmicas

comuns contêm cristais de oxalato de cálcio, prismáticos retangulares e estiloides. Quanto ao pó dessa espécie são encontrados fragmentos da epiderme, com as paredes das células anticlinais retas cobertas por cutículas lisas e calibradas que apresentam estruturas como cristais piramidais, prismáticos e estiloides. Verifica-se também a presença de estômatos anomocíticos e fragmentos de parênquima paliádico de duas a três camadas. Esses fragmentos apresentam braquiesclereides, idioblastos fenólicos e cristais de oxalato de cálcio em diferentes formas, como estiloides na forma de diamante, cristais hexagonais e bipiramidais. Vale ressaltar que, mesmo permitindo a distinção da *Monteverdia ilicifolia* (Mart. ex Reissek) Biral de diversas plantas de morfologia macroscópica similar, a microscopia não é capaz de discriminar essa espécie do contaminante *Monteverdia aquifolium* (Mart.) Biral (Brasil, 2019; Antunes *et al.*, 2023b).

Figura 1 - Aparência das folhas de *Monteverdia ilicifolia* (Mart. ex Reissek) Biral (A, B) e sua distribuição geográfica (C)



Fonte: Trap Hers (2021); Daniel (2021); Catalogue of Life Checklist (2017)

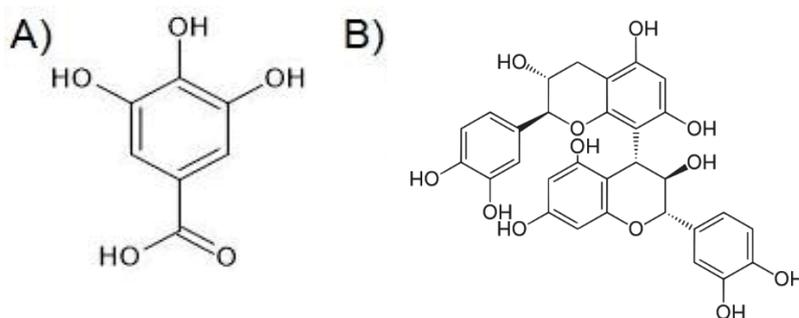
### 3.1.2. Composição fitoquímica

Rica em metabólitos secundários da classe dos taninos, triterpenos e flavonoides, a espinheira-santa serve como coadjuvante ou protagonista de diversos tratamentos por meio de seus efeitos anti-inflamatórios, antiulcerogênicos, antitumorais, entre outros (Araujo Junior *et al.*, 2013; Moura *et al.*, 2021; Duarte; Debur, 2005).

#### 3.1.2.1 Taninos

Com elevado peso molecular (entre 500 e 3000 Da), os taninos são os metabólitos mais abundantes nas folhas da espinheira-santa. Este grupo de moléculas pode ser dividido em hidrolisáveis e condensados, sendo estes últimos predominantes. Sua estrutura básica consiste de duas a três hidroxilas fenólicas ligadas a um fenol, como ilustrado na **Figura 2** (Okuda; Ito, 2011).

Figura 2 – Estrutura química dos taninos hidrolisáveis e condensados



Legenda: A) Tanino hidrolisável, ácido gálico; B) Tanino condensado, procianidina;

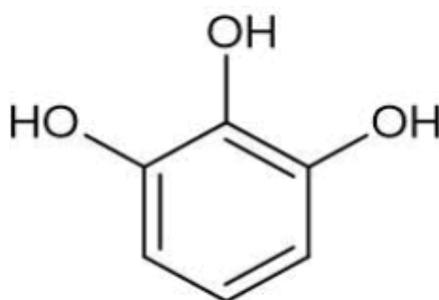
Fonte: O autor (2024)

Os taninos condensados prevalentes na espinheira-santa são principalmente monoméricos e diméricos. O mecanismo de ação desses metabólitos varia conforme seus efeitos benéficos. Como anticarcinogênico, os taninos atuam bloqueando ou retardando a formação de tumores, enquanto como antioxidantes sua ação vem do alto grau de hidroxilação presente no anel que os compõe. Além disso, esses compostos demonstram efeitos anti-helmínticos, antivirais, antimutagênicos, anti-inflamatórios, entre outros. A maioria de suas funções decorre da capacidade dessas

moléculas de se ligarem a proteínas, precipitá-las e complexar íons metálicos. (Monteiro; Brandelli, 2017; Sharma *et al.*, 2019).

A Relação Nacional de Medicamentos Essenciais (RENAME) estabelece que produtos fitoterápicos contendo espinheira-santa devem conter entre 60 e 90 mg de taninos totais em sua composição, calculados como pirogalol. Este composto fenólico, encontrado em diversas espécies vegetais, possui ação antioxidante, antipsoriática e inibitória de enzimas. Sua estrutura molecular consiste em 1,2,3-trihidroxibenzeno, conforme ilustrado na **Figura 3**. (Shin; Park; Lee, 2019; Sarikaya, 2015).

Figura 3 – Estrutura química do pirogalol

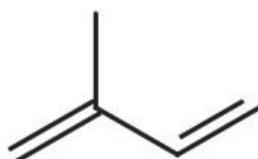


Fonte: O autor (2024)

### 3.1.2.2 Triterpenos

Incluídos no grupo dos terpenos, os triterpenos (**Figura 4**) são formados pela condensação de seis unidades de isopreno, uma molécula de cinco unidades de carbono e que serve como base para a formação dos diversos tipos de terpenos. Outras classificações deste grupo incluem monoterpenos, diterpenos, sesquiterpenos, entre outros (Jackson, 2008).

Figura 4 - Estrutura molecular básica dos terpenos



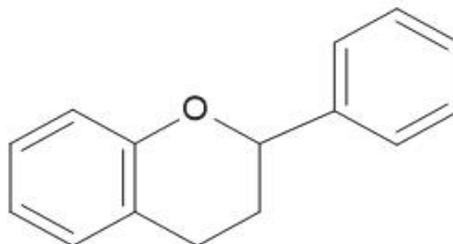
Fonte: O autor (2023)

Nas plantas, estes metabólitos inibem o crescimento das raízes, protegem contra herbívoros e favorecem a propagação da espécie (Monteiro; Brandelli, 2017). Estudos realizados por Jorge e colaboradores (2004) confirmam que, em seres humanos, os triterpenos auxiliam significativamente a epitelização e deposição de colágeno nos tecidos, ajudando no tratamento de úlceras e outros danos teciduais semelhantes.

### 3.1.2.3 Flavonoides

Os flavonoides constituem um enorme grupo de polifenóis com núcleo fundamental de 15 carbonos, composto por dois anéis aromáticos ligados por uma ponte de três carbonos a uma estrutura aromática central que contém um átomo de oxigênio. Uma representação gráfica desta estrutura pode ser encontrada na **Figura 5**.

Figura 5 – Estrutura química básica de um flavonoide



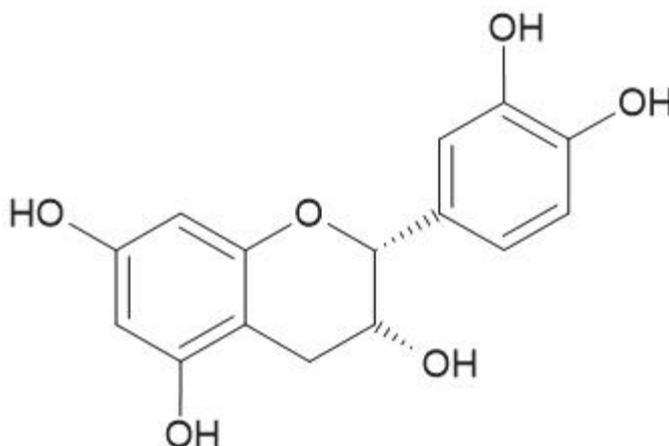
Fonte: O autor (2023)

Com mais de cinco mil polifenóis diferentes, os flavonoides são diversificados e oferecem benefícios tanto para as plantas quanto para seres humanos. Em espécies vegetais, esses compostos proporcionam proteção contra organismos agressores (insetos e vírus) e incidência de raios UV, além de atrair agentes polinizadores essenciais. Nos humanos, esses compostos atuam como anti-inflamatórios, antialérgicos, anticarcinogênicos, antivirais, antibacterianos, combatem doenças cardiovasculares e auxiliam na absorção de vitamina C. Essas propriedades provêm de sua capacidade de complexar metais, remover radicais livres e se ligar com especificidade a proteínas (Alzand; Mohamed; 2012; Simões *et al.*, 2017).

Inclusas neste grupo de compostos estão as catequinas, comumente encontradas em chás medicinais. Na espinheira-santa, a epicatequina é a principal representante deste grupo, com fórmula molecular  $C_{15}H_{14}O_6$  e crescente relevância. Esta molécula é reconhecida no método de identificação descrito na FB 6ª edição para garantia da procedência de *M. ilicifolia* (SIMÕES, *et al.*, 2017).

Classificada como uma catequina de configuração 2R, 3R, a epicatequina (**Figura 6**) é principalmente encontrada em vinhos e chás verdes. Produtos à base de cacau também contém quantidades relevantes dessas moléculas, um dos motivos pelos quais chocolates estão associados à saúde vascular. Este metabólito age como quelantes de íons metálicos e neutralizador de espécies reativas de oxigênio (ROS). Sua capacidade antioxidante o torna um potencial fitoterápico para diabetes do tipo 2, enfermidades neurodegenerativas, cardiovasculares e câncer. A conexão epicatequina-mitocôndria se mostrou bastante relevante devido à capacidade do metabólito de proteger contra a disfunção mitocondrial, reduzindo a produção excessiva de ROS, nocivos ao ser humano (Bernatoniene; Kopustinskiene, 2018; Daussin; Heyman; Burelle, 2020).

Figura 6 – Estrutura química da epicatequina



Fonte: O autor (2023)

### 3.1.3 Efeitos farmacológicos

Vários estudos etnofarmacológicos comprovaram o uso de folhas de *M. ilicifolia* para fins como anti-inflamatório, antiulcerogênico, antimutagênico, anticarcinogênico, entre outros. Os principais usos incluem o tratamento de toxicidade provocada pela

administração de indometacina e álcool, prevenção de tumores, entre outros. Além desta espécie vegetal ser bastante eficaz para esses tratamentos, estudos também não identificaram efeitos tóxicos agudos ou crônicos ao longo dos anos de uso. Pesquisas mostram que a eficácia da espinheira-santa é comparável àquelas alcançadas com o uso do fármaco cimetidina. *M. ilicifolia* é empregada em diversas partes do mundo, mas tem como característica ser uma planta medicinal típica da medicina tradicional do Brasil, (Filho; Zanchett, 2020).

Como citados anteriormente, os principais compostos bioativos responsáveis pelos efeitos farmacológicos de *M. ilicifolia* são os flavonoides, taninos e triterpenos. A combinação dos três compostos confere à infusão da planta uma gama de efeitos terapêuticos através de diversos mecanismos diferentes. Entre esses mecanismos estão a inibição da bomba  $H^+$ ,  $K^+$ -ATPase, a qual promove a redução de lesões gástricas; redução da pressão arterial e risco de eventos cardiovasculares, resultante da abertura dos canais de potássio, produção de óxido nítrico e ativação da guanilato-ciclase, ocasionando uma vasodilatação significativa; desaceleração da proliferação de células tumorais por meio de vários mecanismos; e efeito anti-inflamatório pela redução de fatores pró-inflamatórios como a IL-6, IL-1 e TNF- $\alpha$  (Velo *et al.*, 2017; Filho; Zanchett, 2020). No entanto, ao buscar sobre esta espécie vegetal em diversas bases de dados disponíveis, nota-se uma escassez de estudos clínicos que comprovem esses mecanismos e seus respectivos efeitos terapêuticos, criando uma lacuna de informações significativas.

Ao se tratar da espécie vegetal, os metabólitos secundários exercem a função de proteção, evitando ataques de espécies agressoras ao mesmo tempo que atraem polinizadores, favorecendo a propagação da espécie. Os compostos flavonoides, taninos e triterpenos são responsáveis pela popularização da espinheira-santa, atribuindo-lhe um potencial comercial significativo.

#### **3.1.4 Produtos disponíveis no mercado**

A espinheira-santa possui 6 produtos com registro válido, mas apenas 4 desses são atualmente comercializados, incluindo duas na forma farmacêutica de cápsula, mas também tintura e extrato fluido. Como produtos magistrais, as cápsulas contendo o pó da planta são as formas mais comumente encontradas, mas também existem as

tinturas, folhas secas, sachês contendo o pó e infusões (Almeida *et al.*, 2012). Essas diferentes apresentações oferecem ao paciente a escolha da forma que se adapta às suas necessidades, com os principais fins terapêuticos sendo o tratamento de úlceras e regulação da função gástrica.

A RENAME, constituída a partir da Portaria GM/MS nº 533, de 28 de março de 2012, inclui essa espécie vegetal e a reconhece como uma alternativa para diversos tratamentos no âmbito do Sistema Único de Saúde (SUS), recomendando a dose de 60 – 90 mg de taninos totais expressos em pirogalol, como apresentado no **Quadro 1** (Brasil, 2022b).

Quadro 1- Formas farmacêuticas reconhecidas pela RENAME para tratamentos com *Monteverdia ilicifolia* (Mart. ex Reissek) Biral

<b>Espécie vegetal</b>	<b>Concentração/Composição</b>	<b>Forma farmacêutica</b>	<b>Componente</b>
Espinheira-santa ( <i>M. ilicifolia</i> )	60 a 90 mg taninos totais expressos em pirogalol (dose diária)	Cápsula	Básico
		Tintura	
		Suspensão oral	
		Emulsão oral	

Fonte: O autor (2023). Adaptado de: RENAME (2022)

### 3.1.5 Espécies vegetais usadas como adulterantes de *Monteverdia ilicifolia* (Mart. ex Reissek) Biral

A espinheira-santa possui caracteres botânicos próprios, mas existem espécies morfologicamente similares, o que a torna suscetível a falsificação. A adulteração também é favorecida pela nomenclatura popular, pois existem várias espécies conhecidas popularmente como espinheira que são utilizadas para esse fim

A falsa espinheira-santa, nome popular para a espécie *M. aquifolium*, é a espécie mais comum encontrada em produtos comerciais, seguida de *S. bonplandii*, *C. gongonha*, *J. rhombifolia* e *Z. ilicifolia* que também são espécies comumente utilizadas para fins de adulteração, intencional ou não (UFSC, 2020).

*Monteverdia ilicifolia* e os adulterantes citados, foram objetos de estudo morfoanatômico de Antunes *et al.* (2023a), sendo que esses autores afirmam que é

possível a identificação de *M. ilicifolia* e a diferenciação desta das demais espécies utilizando caracteres macro e microscópicos em amostras inteiras e rasuradas. Contudo esses mesmos autores, indicam a necessidade de análises químicas quando a amostra estiver pulverizada (Antunes *et al.*, 2023b).

A falsificação de produtos à base das folhas secas só é visualmente identificável em situações que a folha se encontra íntegra, no entanto, a similaridade entre as espécies pode ser suficiente para que nem nesse cenário seja possível distingui-las. Além disso, em apresentações como extratos hidroalcoólicos, cápsulas e sachês, essa discriminação é inviável, exigindo metodologias específicas e precisas para identificar as diversas formas farmacêuticas disponíveis.

Espécies adulterantes trazem consigo problemas relacionados à saúde e ao comércio. No espectro da saúde, o emprego de outras plantas, com diferenças metabólicas relevantes, compromete o tratamento e o torna ineficaz. Os principais grupos fitoquímicos afetados com essa troca são os taninos e catequinas, ambos devendo atender aos valores mínimos preestabelecidos pela FB 6ª edição. Do ponto de vista econômico, a substituição por espécies de menor custo promove uma competição desleal e proporciona vantagens ilegais ao detentor do produto alterado.

As principais espécies utilizadas como adulterantes de produtos à base da *Monteverdia ilicifolia* (Mart. ex Reissek) Biral podem ser encontrados na **Tabela 1**.

Tabela 1 - Espécies adulterantes da *Monteverdia ilicifolia*

Espécie	Família	Morfologia	Distribuição Geográfica	Característica Microscópica	Composição Química	Atividades Biológicas
<b><i>Monteverdia aquifolium</i> (Mart.) Biral</b>	Celastraceae	Arbusto ou árvore pequena (até 15 m de altura). Folhas dísticas, elípticas a oblongo-elípticas, margens espinescentes e revolutas, oito a doze pares de espinhos. Lâminas cartáceas, nervura primária plana na face adaxial e proeminente na abaxial, com 8-12 pares de nervuras secundárias.	Brasil, Paraguai, Argentina.	Estômatos anomocíticos, células epidérmicas retangulares. Estípulas de 0,6-1,6 mm de comprimento, prontamente caducas, com formato largo-triangular e margens fimbriadas. Os pecíolos são canaliculados e glabros.	Taninos, flavonoides, alcaloides celastrina.	Antimicrobiana, antiulcerativa, anti-inflamatória.
<b><i>Sorocea bonplandii</i> (Baill.)</b>	Moraceae	Árvore de pequeno a médio porte, com folhas simples, alternas, e frutos múltiplos pequenos. O caule possui estípulas caducas. As folhas apresentam lâminas com ápice acuminado, base aguda ou obtusa, margens espinulosas dentadas e face abaxial que pode ser glabra ou puberulenta. Os pecíolos podem ser puberulentos ou glabros.	Brasil, Paraguai, Argentina.	Células epidérmicas com cristais de oxalato de cálcio, estômatos anomocíticos. As flores estaminadas apresentam pedicelos com estames que se inflexionam durante a antese e pistilódios às vezes presentes	Compostos fenólicos, flavonoides, alcaloides.	Antioxidante, hepatoprotetora, antidiabética.
<b><i>Citronella gongonha</i> (Mart.)</b>	Cardiopteridaceae	Arbusto com folhas ovais, coriáceas, com nervuras salientes, e frutos drupáceos. As folhas possuem pecíolos com comprimento de 0,5 a 1 cm, base aguda ou rotunda, e ápice rotundo ou cuspidado. As lâminas	Brasil, Bolívia.	Células esclerenquimáticas nas folhas, estômatos paracíticos.	Óleos essenciais, ácidos fenólicos, flavonoides.	Antimicrobiana, anti-inflamatória, analgésica.

<p>apresentam formas elípticas, obovadas, lanceoladas, oblongas ou ovadas. A venação secundária apresenta de 5 a 8 pares.</p>		
<p>Arbusto ou árvore pequena, folhas simples, alternas, coriáceas, com margens dentadas e frutos pequenos. A árvore é pequena, com 2 a 4 m de altura e diâmetro de 5 a 30 cm, possuindo ritidoma fendilhado. As folhas são alternas, com formato romboédrico e um espinho em cada ângulo, glabras e coriáceas. A nervura primária é conspicua e as secundárias pouco visíveis. Pecíolo curto, com 0,5 cm.</p>	<p>Brasil, Uruguai, Argentina.</p>	<p>Células parenquimáticas com inclusões fendílicas, cristais de oxalato.</p> <p>Taninos, flavonoides, alcaloides.</p> <p>Anti-inflamatória, antioxidante, antifúngica.</p>
<p>Árvore de médio porte, folhas compostas pinadas com folíolos elípticos, e frutos leguminosos. O caule apresenta tronco liso a fissurado, de coloração verde clara a escura. As folhas são simples, alternas, obovadas a elípticas, coriáceas, com nervação broquidódroma. A base é cuneada a obtusa, o ápice agudo a retuso e mucronado, e a margem serrilhada. As estípulas são acerosas a falcadas, glabras ou esparsamente tomentosas, rígidas. Os pecíolos são canaliculados, glabros a puberulentos.</p>	<p>Brasil, Paraguai.</p>	<p>Fibras esclerenquimáticas nas folhas, estômatos paracíticos. O cálice é espatáceo, unilobado, cartáceo, glabro na face interna da base e piloso no ápice.</p> <p>Compostos fenólicos, alcaloides, terpenos.</p> <p>Antioxidante, anti-inflamatória, antimicrobiana.</p>

Fonte: O autor (2023).

A dimensão continental do Brasil, aliada a diversidade ambiental, confere-lhe um potencial imensurável no ramo dos fitoterápicos. Esta classe de medicamentos teve seu uso aprovado no âmbito do Sistema Único de Saúde (SUS) a partir da Portaria nº 971, de 3 de maio de 2006, um marco importante, mas controverso. Enquanto diversos especialistas vêm na fitoterapia um alicerce para diversos tratamentos, muitos ainda acreditam que espécies vegetais devem ser apenas utilizadas como métodos alternativos e complementares, na forma de plantas medicinais (Brasil, 2006; Figueredo; Gurgel; Gurgel Junior, 2014).

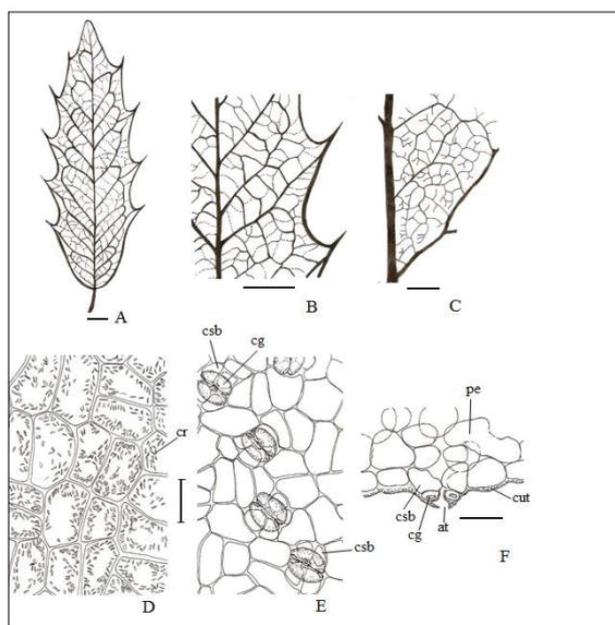
O uso de fitoterápicos, no entanto, cresce ano após ano, e o registro destes produtos é regulado pela RDC nº. 26, de 13 de maio de 2014. Esta resolução estabeleceu que o relatório técnico de segurança e eficácia do produto pode ter suas informações baseadas em quaisquer dos seguintes pontos: ensaios pré-clínicos e clínicos; conhecimento em literatura técnico-científica; presença na "Lista de medicamentos fitoterápicos de registro simplificado" da Instrução Normativa (IN) nº 2, de 13 de maio de 2014, e suas atualizações; e tradicionalidade de uso, comprovada por diversas alternativas, como o uso popular e seguro documentado por mais de 30 anos, e a utilização em episódios e/ou para enfermidades de baixa gravidade. Este último ponto traz consigo problemas quanto ao que se deve considerar como fitoterápico, abrindo espaço para dezenas de produtos sem real eficácia e normalizando o emprego de espécies vegetais para fins a que não são destinados. Ocorre também uma normalização do uso abusivo de fitoterápicos devido à falsa premissa de que, por serem "naturais", não causam efeitos colaterais. Aliado a isso, a presença de produtos sem registro ou notificação na ANVISA e a fiscalização com baixo rigor abrem espaço para falsificações e adulterações. São necessário mudanças em duas frentes distintas: ampliar as evidências científicas sobre a segurança e eficácia, com base em estudos consistentes, e desenvolver metodologias capazes de detectar possíveis desvios de qualidade nestes produtos (Santos *et al.*, 2011; Brasil, 2014a; Brasil, 2014b; Figueredo; Gurgel; Gurgel Junior, 2014).

A problemática apresentada, quanto a identificação e diferenciação da espinheira-santa de espécies morfológicamente similares, aliada ao fato de sua distribuição territorial consistir basicamente na América do Sul, mostra a falta atenção da sociedade científica para tal problema e destaca a importância de se desenvolver métodos que sejam eficazes de garantir a autenticidade destas espécies.

### 3.1.6 Métodos de identificação

A identificação da espinheira-santa, por meio de método estabelecido pela FB 6ª edição, ocorre em nível macro e microscópico, e químico. O método macroscópico envolve a análise das folhas, avaliando a coloração, tamanho, presença de pecíolo, características morfológicas foliares e presença de projeções pontiagudas no bordo foliar. Na microscopia, são verificados os marcadores anatômicos como epiderme contendo cristais de oxalato de cálcio, mesofilo dorsiventral, formato da nervura central e do pecíolo em secção transversal, sendo que os nós também passam por análise microscópica (Beltrame *et al.*, 2012; Brasil, 2019; Antunes *et al.* 2023b).

Figura 7 – Características macro e microscópicas de folhas de *Monteverdia ilicifolia* (Mart. ex Reissek) Biral



Fonte: Farmacopeia Brasileira 6ª Edição (2019)

O método analítico de identificação consiste na cromatografia em camada delgada (CCD), onde a banda apresentada pela amostra é comparada com uma epicatequina de referência (Brasil, 2019). Além da identificação, a espinheira-santa deve ser submetida a um método de quantificação para aprovação quanto ao controle de qualidade da matéria-prima vegetal. O metabólito protagonista para identificação da espécie, a epicatequina, é um alvo do doseamento, realizado por cromatografia líquida de alta eficiência, junto aos taninos totais. A FB 6ª edição reconhece que a

espinheira-santa deve conter, no mínimo, 2,0% de taninos totais e 0,28% de epicatequinas. Ambos os compostos são cruciais para a atividade fitoterápica e identificação da espécie vegetal.

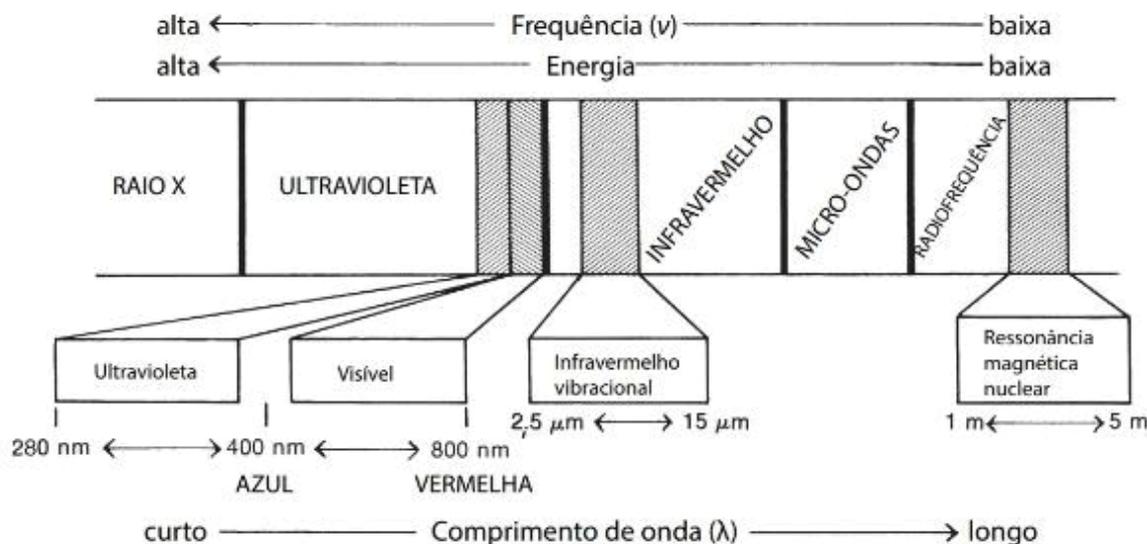
Diferentes técnicas cromatográficas são a base da identificação e doseamento dos metabólitos secundários em *M. ilicifolia*. No entanto, a desvantagem dessas metodologias está no impacto ambiental devido ao excesso de solvente utilizados. Nas diversas bases de dados disponíveis, não existem estudos que combinem eficiência com química verde, evidenciando a necessidade do surgimento de metodologias inovadoras (Lopes *et al.*, 2010).

Uma técnica em potencial para essa finalidade, ainda pouco explorada, é a espectroscopia por infravermelho. Esse método é uma alternativa viável para a análise de diversas espécies vegetais e compostos em geral, pois a radiação no infravermelho interage com moléculas que possuem momento dipolar associado a suas ligações covalentes, as quais estão presentes em praticamente todas as estruturas químicas, permitindo identificar diferentes grupos funcionais e compostos (Pavia *et al.*, 2015; ICNIRP, 2013).

### 3.2 ESPECTROSCOPIA DE INFRAVERMELHO

É uma técnica espectroscópica que utiliza a região do espectro eletromagnético entre 780 nm e 1 mm (**Figura 8**) baseada na medição da radiação infravermelha absorvida por uma amostra. Classificado como um método não-destrutivo e custo-eficiente, a espectroscopia de infravermelho é amplamente empregada nas indústrias alimentícia, farmacêutica, petrolífera, entre outras. Duas das principais funções atribuídas a esta metodologia são a quantificação de concentrações e a detecção de adulterações em diversos produtos (Capuano; Ruth, 2016; Panikuttira; O'Donnell, 2018; Subramanian; Prabhakar; Rodriguez-Saona, 2016; Veerasingam *et al.*, 2020).

Figura 8 - Região do espectro magnético contendo o infravermelho



Fonte: Pavia e colaboradores (2015)

A radiação infravermelha atua identificando a estrutura de diversos compostos, pois diferentes moléculas podem atingir um estado excitado ao absorver essa radiação em frequências específicas, na ordem de 8 a 40 kJ/mol. A medição da vibração dos átomos que compõem a estrutura molecular, nos diferentes comprimentos de onda, gera os chamados espectros, que são representações gráficas e o objeto de estudo desse método. Devido à natureza específica da absorção, duas moléculas distintas não produzem o mesmo espectro, permitindo a identificação individual de cada uma (Pavia *et al.*, 2015).

Como vantagens da espectroscopia de infravermelho pode ser citada a elevada especificidade de identificação associada à química verde, resultando em uma reduzida geração de resíduos e uma diminuta necessidade de reagentes e amostras. A rápida realização e pouca exigência de treinamento também tornam esta técnica uma referência para a identificação de amostras (Capuano; Van Ruth, 2016).

A faixa de comprimentos de onda que a radiação infravermelha (IR) engloba pode ser dividida em três tipos: IR-A (infravermelho próximo, cobrindo os comprimentos de onda de 800 a 2400 nm), IR-B (infravermelho médio, compreendendo 2,4 até 30  $\mu\text{m}$ ) e IR-C (infravermelho longo, o mais distante da luz visível, englobando 30  $\mu\text{m}$  e 1 mm). Em cada comprimento de onda, a energia absorvida pela amostra permite identificar as ligações presentes na estrutura

molecular em questão. A consequência dessa interação é a geração do *fingerprint* da amostra (IFSC, 2013).

No caso do estudo de espécies vegetais, as principais ligações detectadas são CH, OH, NH, SH e CO. A riqueza metabólica das plantas se traduz em bandas que representam simultaneamente essas ligações, porém de forma e concentração diferentes em cada espécie (Oliveira *et al.*, 2018).

Os espectros adquiridos, por si só, podem não fornecer informações suficientes para diferenciar espécies vegetais com constituição semelhante. Uma investigação mais aprofundada se torna possível pelo uso de métodos estatísticos capazes de avaliar minuciosamente as bandas detectadas. As análises estatísticas relevantes para este estudo incluem tanto a análise exploratória, que busca identificar as espécies, quanto a análise discriminatória, que visa diferenciá-las. A quimiometria, um ramo do estudo analítico, é capaz de atender a essas necessidades.

**CAPÍTULO 1: ESPINHEIRA-SANTA, ASPECTOS FITOQUÍMICOS, CLÍNICOS E  
TECNOLÓGICOS: UMA REVISÃO DE ESCOPO**

## 4 CAPÍTULO 1

### 4.1 DESCRIÇÃO GERAL

Com destaque dentro das classes dos fitoterápicos, a espinheira-santa é reconhecida por conta de seus efeitos cicatrizante, anti-inflamatório, antiulcerogênico, analgésico, entre outros. Ocorre, no entanto, uma extensa utilização desta espécie vegetal para fins não comprovados cientificamente (uso *off-label*), colocando em risco a população.

Uso *off-label* é a utilização de produtos farmacêuticos para fins que não seguem as indicações comprovadas e homologadas, sendo empregados para outros propósitos sem a devida comprovação, servindo como uma alternativa sem segurança e/ou eficácia, mas possivelmente útil (FDA, 2018). O impacto desta prática não engloba apenas medicamentos sintéticos, mas também os fitoterápicos, como é o caso dos produtos à base da espinheira-santa.

Com o passar de gerações e culturas, muitos foram os usos de *M. ilicifolia*. A difusão destes conhecimentos ocasionou equívocos do ponto de vista farmacêutico, tornando popular a aplicação desta planta para fins diferentes daqueles com evidências pré-clínicas. Usos como promotor de espermatogênese, anticarcinogênico, prevenção de ototoxicidade, entre outros, são alguns exemplos (Araújo Júnior *et al.*, 2013; Jorge *et al.* 2004; Kasse *et al.* 2008; Montanari; Carvalho; Holder, 1998). Vale ressaltar como essa promoção de conhecimentos não se limita ao efeito, mas também a apresentação do produto, com dezenas de preparações que podem ser ineficientes quanto ao uso previsto e otimizado.

Dentro deste contexto é que se insere a revisão de escopo. O levantamento de evidências, sua síntese e resposta à questão norteadora definida para o estudo servem de base para o avanço científico. A síntese das informações disponibilizadas nas bases de dados permite agregar conhecimentos a partir de pontos de vista diferentes. A execução da revisão de escopo sintetiza as descobertas relacionadas a um determinado assunto, o que possibilita visualizar aspectos que precisam ser aprofundados, e assim orienta estudos futuros (Mak; Thomas, 2022; Silva *et al.*, 2018).

Buscando conhecer as lacunas de informação sobre aspectos tecnológicos, de segurança e eficácia relacionados ao uso da espinheira-santa foi conduzida uma

revisão de escopo com enfoque na parte fitoquímica, farmacológica e tecnológica de formulações inovadoras contendo essa espécie.

Como resultado procura-se conhecer para quais finalidades a espinheira-santa pode ser utilizada, as principais formas farmacêuticas, concentrações de uso, as formas como a planta foi utilizada (folha, pó, extrato), toxicidade, entre outros, e apontar lacunas para futuras pesquisas ou desenvolvimentos tecnológicos.

## 4.2 OBJETIVOS

Realizar uma revisão de escopo a partir de artigos e patentes envolvendo as perspectivas fitoquímicas, farmacológicas e inovações tecnológicas da espinheira-santa, buscando responder à seguinte pergunta: “Quais são as evidências científicas sobre a *Monteverdia ilicifolia* (Mart. ex Reissek) Biral com relação aos seus aspectos fitoquímicos, farmacológicos e tecnológicos?”

### 4.2.1 Objetivos específicos

- Desenvolver uma estratégia de busca que compreenda estudos e inovações envolvendo a espinheira-santa;
- Manejar artigos e patentes a partir do gerenciador de referências *Rayyan*, excluindo estudos duplicados e que não se adequem ao tema.
- Extrair as informações relevantes dos estudos que se adequem aos critérios de inclusão propostos;
- Correlacionar os aspectos fitoquímicos, farmacológicos e tecnológicos encontrados para a espinheira-santa.

### 4.3 METODOLOGIA

Este estudo foi conduzido seguindo as diretrizes do Instituto Joanna Briggs (JBI) e os dados foram reportados conforme a extensão “*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*” para revisão de escopo (PRISMA-ScR) (Tricco *et al.*, 2018; Peters *et al.*, 2015). O protocolo da revisão foi registrado na base de dados OSF (Zein *et al.*, 2023, October 18. Espinheira-Santa phytochemical, clinical and technological aspects: A scoping review. <https://doi.org/10.17605/OSF.IO/ZSPU2>).

As pesquisas se deram a nível de artigos e patentes. Para a busca dos artigos, termos relacionados a estudos fitoquímicos, farmacológicos e conhecimento técnico-científico, nomenclatura científica e popular atual ou anteriormente relacionado a espinheira-santa foram utilizados. Esses termos foram pesquisados por título, resumo e/ou palavras-chave, e adaptados para as plataformas de busca *PubMed*, *Scopus* e *Web of Science* a partir dos operadores Booleanos “AND” e “OR”. (Tabela 2). Uma busca manual também foi conduzida.

Tabela 2 – Estratégia de busca para os artigos

Plataforma de busca	Queries
PubMed	<p>((("Maytenus ilicifolia"[Title/Abstract]) OR ("Monteverdia ilicifolia"[Title/Abstract])) OR ("espinheira-santa"[Title/Abstract])) OR ("espinheira santa"[Title/Abstract])) OR ("M. ilicifolia"[Title/Abstract])</p> <p>((("folk medicine"[Title/Abstract])) OR ("traditional medicine"[Title/Abstract])) OR ("technological approach"[Title/Abstract])) OR ("pharmacological"[Title/Abstract])) OR ("clinical trial"[Title/Abstract])</p> <p>((("Maytenus ilicifolia"[Title/Abstract]) OR ("Monteverdia ilicifolia"[Title/Abstract])) OR ("espinheira-santa"[Title/Abstract])) OR ("espinheira santa"[Title/Abstract])) OR ("M. ilicifolia"[Title/Abstract])) AND ((("folk medicine"[Title/Abstract])) OR ("traditional medicine"[Title/Abstract])) OR ("technological approach"[Title/Abstract])) OR ("pharmacological"[Title/Abstract])) OR ("clinical trial"[Title/Abstract])</p>
Scopus	<p>( TITLE-ABS-KEY ( "maytenus ilicifolia" ) OR TITLE-ABS-KEY ( "monteverdia ilicifolia" ) OR TITLE-ABS-KEY ( "espinheira-santa" ) OR TITLE-ABS-KEY ( "espinheira santa" ) OR TITLE-ABS-KEY ( "m. ilicifolia" ) )</p>

( TITLE-ABS-KEY ( "folk medicine" ) OR TITLE-ABS-KEY ( "traditional medicine" ) OR TITLE-ABS-KEY ( "technological approach" ) OR TITLE-ABS-KEY ( "pharmacological" ) OR TITLE-ABS-KEY ( "clinical trial" ) )

( ( TITLE-ABS-KEY ( "maytenus ilicifolia" ) OR TITLE-ABS-KEY ( "monteverdia ilicifolia" ) OR TITLE-ABS-KEY ( "espinheira-santa" ) OR TITLE-ABS-KEY ( "espinheira santa" ) OR TITLE-ABS-KEY ( "m. ilicifolia" ) ) ) AND ( ( TITLE-ABS-KEY ( "folk medicine" ) OR TITLE-ABS-KEY ( "traditional medicine" ) OR TITLE-ABS-KEY ( "technological approach" ) OR TITLE-ABS-KEY ( "pharmacological" ) OR TITLE-ABS-KEY ( "clinical trial" ) ) )

Web of Science

((((TS=("Maytenus ilicifolia")) OR TS=("Monteverdia ilicifolia" )) OR TS=("espinheira-santa")) OR TS=("espinheira santa")) OR TS=("M. ilicifolia"))

((((TS=("folk medicine")) OR TS=("traditional medicine")) OR TS=("technological approach")) OR TS=("pharmacological")) OR TS=("clinical trial"))

((((TS=("Maytenus ilicifolia")) OR TS=("Monteverdia ilicifolia" )) OR TS=("espinheira-santa")) OR TS=("espinheira santa")) OR TS=("M. ilicifolia") AND (((TS=("folk medicine")) OR TS=("traditional medicine")) OR TS=("technological approach")) OR TS=("pharmacological")) OR TS=("clinical trial"))

Fonte: O autor (2023)

Para pesquisa das patentes buscou-se os termos “*Maytenus ilicifolia*” e “*Monteverdia ilicifolia*” nas plataformas de busca *Google Patents*, *Espacenet*, *Derwent Innovations Index* e *The Lens* (Tabela 3).

Tabela 3 – Estratégia de busca para as patentes

Base de dados	Termos pesquisados
<i>Google Patents</i>	
Espacenet	<i>Maytenus ilicifolia/Monteverdia ilicifolia</i>
Derwent Innovations Index	
The Lens	

Fonte: O autor (2023)

#### **4.3.1 Critérios de elegibilidade**

Os critérios de inclusão e exclusão para ambos artigos e patentes foi similar, e buscou delimitar a pesquisa, direcionando o estudo para que respondesse a seguinte questão norteadora: “Quais são as evidências científicas sobre a *M. ilicifolia* com relação aos seus aspectos fitoquímicos, clínicos e desenvolvimentos tecnológicos?”

#### **4.3.2 Critérios de inclusão**

Os critérios de inclusão foram estabelecidos com base no acrônimo PCC, sendo P, população, estudos que utilizaram a *M. ilicifolia* para fins medicinais, C, conceito, estudos que identificaram e correlacionaram o perfil fitoquímico, atividades farmacológicas e perspectivas tecnológicas envolvendo a *M. ilicifolia*, e C, contexto, artigos e patentes que exploraram aspectos pré-clínicos, clínicos e inovações envolvendo a *M. ilicifolia*.

#### **4.3.3 Critérios de exclusão**

Estudos publicados em caracteres não-romanos; Revisões narrativas, revisões de escopo, revisões sistemáticas, estudos do tipo caso-controle; Artigos científicos baseados em espécies vegetais que não a *M. ilicifolia* ou que não a tenham relação com a atividade biológica, aspecto fitoquímico e formas farmacêuticas foram excluídos.

#### **4.3.4 Gerenciamento dos resultados**

Todos os artigos e patentes encontrados foram adicionados à plataforma de gerenciamento de referências *Rayyan*. Foram identificadas e excluídas as duplicatas, enquanto os trabalhos restantes foram verificados quanto à adequação ao tema proposto.

#### **4.3.5 Processo de seleção**

Após a remoção das duplicatas, cada estudo foi analisado individualmente, a nível de títulos e resumos, por dois revisores (AKEZ e JTE). Com os estudos elegíveis selecionados, a próxima etapa foi revisar na íntegra os artigos e patentes restantes. Em ambos os momentos, todas as incoerências quanto ao julgamento dos revisores passaram por uma análise adicional por um terceiro revisor (RELL). Artigos e patentes que não se adequaram à proposta foram excluídos, com sua identificação e razão de exclusão registrados. Estudos aprovados passaram por uma etapa de extração dos dados, visando extrair e agrupar informações pertinentes ao tema.

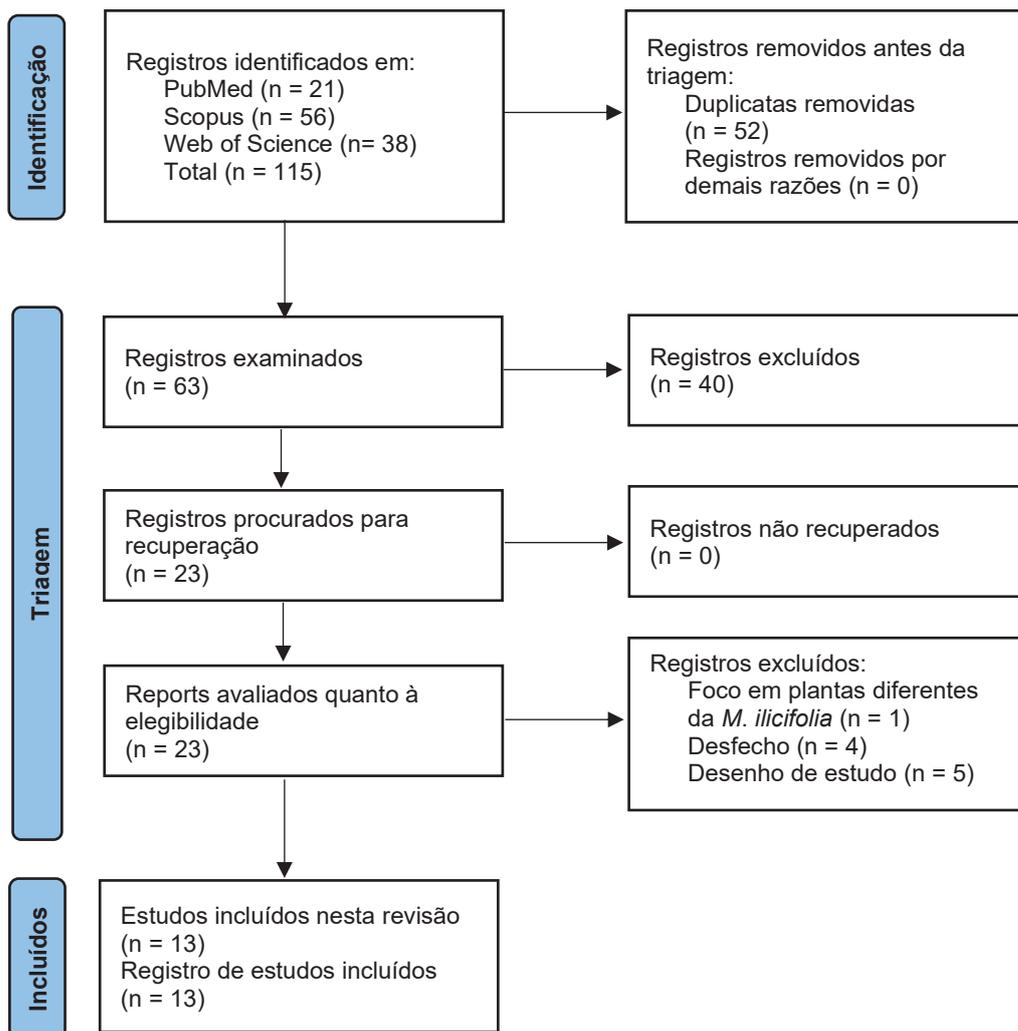
#### **4.3.6 Extração dos dados**

Ambos os revisores responsáveis pela elegibilidade dos estudos também conduziram a extração de dados. A metodologia de extração consistiu em obter as informações de artigos e patentes julgadas relevantes para o presente estudo e as inserir em planilhas do programa *Excel* (versão 2306, 2016). Os dados extraídos nessa etapa foram: **I)** Dados de publicação: Título do artigo/patente, autores, data de publicação (ano, mês e dia), localização (país, estado, cidade), tipo de publicação, e órgão fundador. **II)** Detalhes do estudo: Tipo de estudo, objetivos, critérios de elegibilidade (inclusão e exclusão), forma farmacêutica na qual a espinheira-santa foi empregada, metodologia de produção, concentração da planta na formulação, desfechos, e conclusão quanto a espécie vegetal e produto.

#### 4.4 RESULTADOS E DISCUSSÕES

Um total de 115 artigos foram encontrados nas bases de dados propostas (**Fluxograma 1**). Inicialmente, 52 estudos foram excluídos por se tratarem de duplicatas. Os 63 restantes prosseguiram para a leitura de títulos e resumos, dos quais 40 foram excluídos por não se enquadrarem na proposta de pesquisa, restando 23 artigos para análise integral, ressaltando a baixa quantidade de evidências científicas sobre essa espécie vegetal. Apenas 13 estudos se adequaram aos critérios propostos e tiveram suas informações extraídas (Baggio *et al.*, 2009; Ducat *et al.*, 2011; Ecker *et al.*, 2017; Ferreira *et al.*, 2004; Horn; Vargas, 2003; Jorge *et al.*, 2004; Kohn *et al.*, 2012; Oliveira *et al.*, 1991; Queiroga *et al.*, 2000; Leme *et al.*, 2013; Schindler *et al.*, 2021; Tabach; Duarte-Almeida; Carlini, 2017; Wonfor *et al.*, 2017).

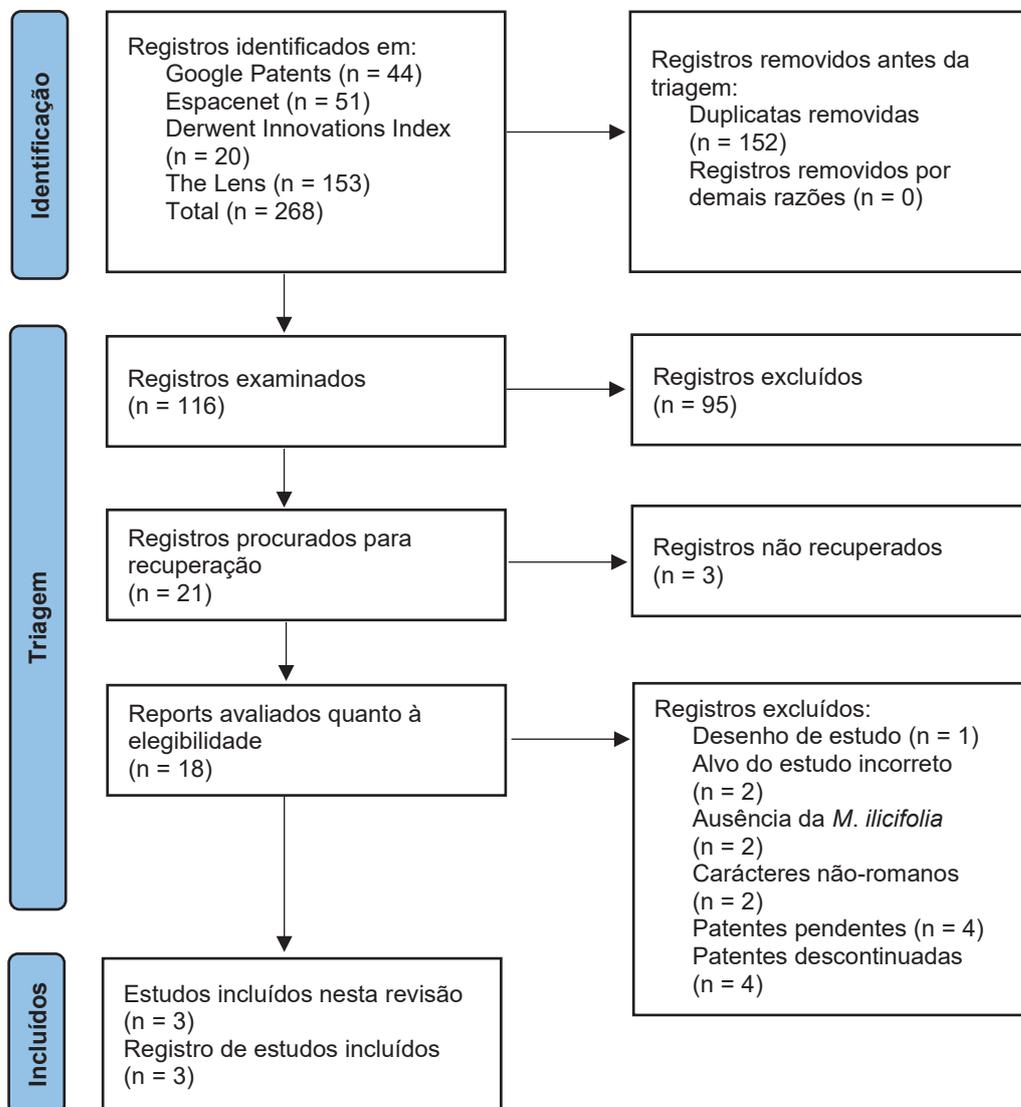
Fluxograma 1 - Fluxograma para revisão de escopo (artigos)



Fonte: O autor (2024)

Para as patentes foram encontrados 268 registros (**Fluxograma 2**). 152 destas foram excluídas na sequência por se tratarem de duplicadas. As 116 patentes restantes seguiram para a etapa de leitura do título e resumo, dos quais foram incluídos 18 registros, e excluídas 95 patentes, constando também três registros não recuperados. Após leitura na íntegra dos registros aprovados, constatou-se que apenas três registros prosseguiriam para extração de informações, com as demais patentes (15) não se adequando ao critério de inclusão (Nakamura *et al.*, 1997; Nakamura *et al.*, 1997; Salvador *et al.*, 2020).

Fluxograma 2 - Fluxograma para revisão de escopo (patentes)

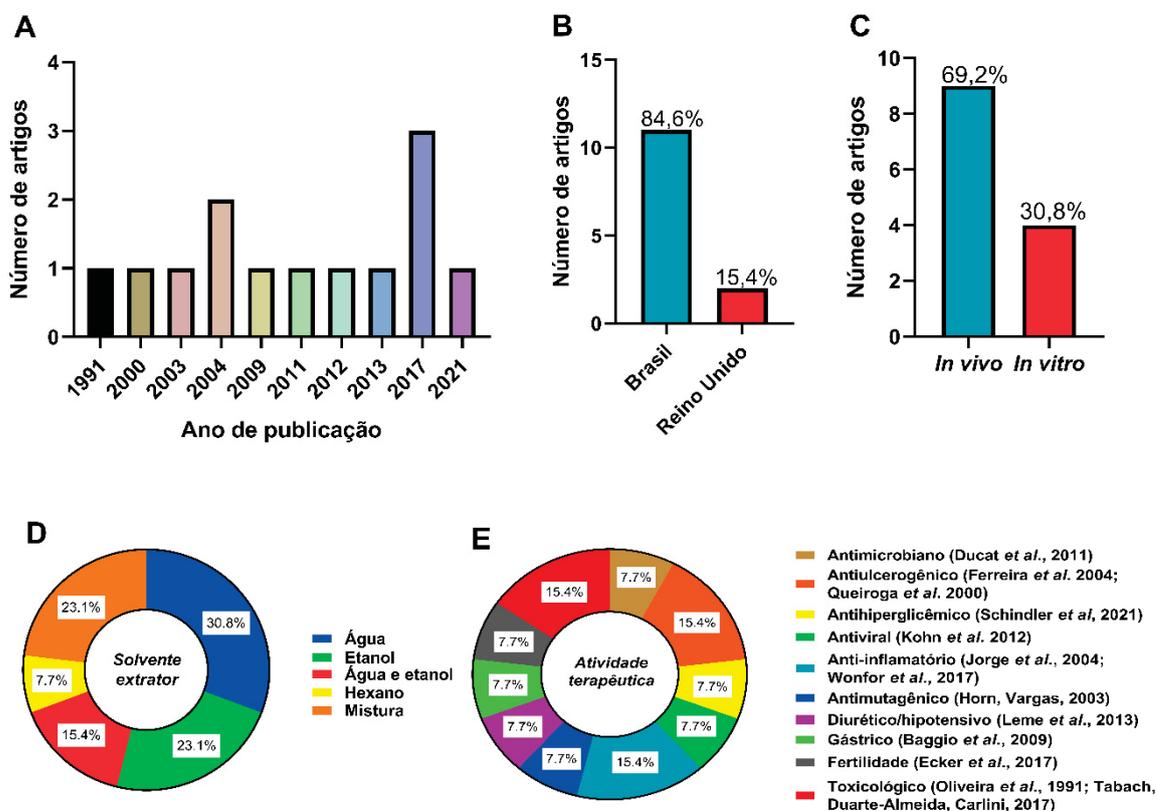


Fonte: O autor (2024)

#### 4.4.1 Inovações científicas

Foi encontrada uma tendência com relação à região de publicação dos 13 estudos aprovados, sendo 11 deles publicados no Brasil (11/13, 84,6%, Baggio *et al.*, 2009; Ducat *et al.*, 2011; Ecker *et al.*, 2017; Ferreira *et al.*, 2004; Jorge *et al.*, 2004; Kohn *et al.*, 2012; Oliveira *et al.*, 1991; Queiroga *et al.*, 2000; Leme *et al.*, 2013; Schindler *et al.*, 2021; Tabach; Duarte-Almeida; Carlini; 2017), e os dois restantes no Reino Unido (2/13, 15,4%, Horn; Vargas, 2003; Wonfor *et al.*, 2017) (**Figura 9**). Os efeitos estudados nos artigos incluíam antimicrobiano (1/13, 7,70%, Ducat *et al.*, 2011), antiulcerogênico (2/13, 15,4%, Ferreira *et al.*, 2004; Queiroga *et al.*, 2000), hipoglicêmico (1/13, 7,70%, Schindler *et al.*, 2021), antiviral (1/13, 7,70%, Kohn *et al.*, 2012), anti-inflamatório (2/13, 15,4%, Jorge *et al.*, 2004; Wonfor *et al.*, 2017), antimutagênico (1/13, 7,70%, Horn; Vargas, 2003), diurético/hipotensivo (1/13, 7,70%, Leme *et al.*, 2013), atividade gástrica (1/13, 7,7%, Baggio *et al.*, 2009), fertilidade (1/13, 7,70%, Ecker *et al.*, 2017) e avaliação de segurança (2/13, 15,4%, Oliveira *et al.*, 1991; Tabach; Duarte-Almeida; Carlini, 2017)) (**Figura 9**). Os anos de publicação compreenderam 1991 a 2021, sendo 2017 o ano com o maior número de estudos selecionados, sendo três registros (Ecker *et al.*, 2017; Tabach; Duarte-Almeida; Carlini; 2017; Wonfor *et al.*, 2017) (**Figura 9**). O **Quadro 2** descreve as principais características dos estudos. A água foi o principal solvente usado para preparar o extrato da folha (4/13; 30,8%), seguida pelo etanol (3/13; 23,1%), uma mistura de diferentes solventes, como hexano, clorofórmio, acetato de etila etc. (3/13; 23,1%), água e etanol individualmente (2/13; 15,4%) e apenas hexano (1/13; 7,7%) (**Figura 9**). Nove dos 13 estudos utilizaram modelos *in vivo* (9/13, 69,2%, Baggio *et al.*, 2009; Ecker *et al.*, 2017; Ferreira *et al.*, 2004; Jorge *et al.*, 2004; Oliveira *et al.*, 1991; Queiroga *et al.*, 2000; Leme *et al.*, 2013; Schindler *et al.*, 2021; Tabach; Duarte-Almeida; Carlini, 2017), sendo um deles um estudo clínico (Tabach; Duarte-Almeida; Carlini, 2017), enquanto os outros quatro foram baseados em modelos *in vitro* (4/13, 31,8%, Ducat *et al.*, 2011; Horn; Vargas, 2003; Kohn *et al.*, 2012; Wonfor *et al.*, 2017) (**Figura 9**). Todos os estudos usaram folhas como seu principal farmacógeno e extratos/infusões em formulações farmacêuticas.

Figura 9 – Características dos artigos incluídos no estudo



Legenda: A = Número de artigos publicados ao longo dos anos; B = Número de artigos por região de publicação; C = Principal modelo experimental utilizado nos estudos; D = Principal solvente de extração utilizado nos estudos; E = Fim terapêutico para cada um dos estudos.

Fonte: O autor (2024)

O crescente número de artigos publicados com o passar dos anos pode ser resultado da conscientização sobre a importância do monitoramento de produtos naturais e da necessidade de colocá-los em evidência. Observa-se que o Brasil é o centro de estudos sobre a *M. ilicifolia*, fato condizente com a tradicionalização da espécie vegetal no país. Os artigos demonstraram como esta espécie vegetal apresenta um amplo espectro de ação, servindo como opção para tratamentos diversos. Os modelos *in vivo* são o foco principal dos estudos com a *M. ilicifolia*, pois esta espécie faz parte de muitos tratamentos e não apresenta efeitos adversos importantes. Todas essas informações, além do principal solvente extrator e conclusões encontradas, foram registradas no **Quadro 2**.

Quadro 2 - Informações extraídas dos registros na forma de artigos

Proposta terapêutica	Referência	País	Extrato	Modelo	Forma farmacêutica	Concentração	Mecanismo de ação	Principais achados
<b>Estudo clínico</b>								
Avaliação de segurança	Tabach, Duarte-Almeida, Carlini, 2017	Brasil	Água	Humanos (12 mulheres e 12 homens) entre 20 e 40 anos de idade	Extrato seco	100, 200, 500, 1000, 1500 e 2000 mg	-	O MIE* se mostrou seguro até uma dose de 2000 mg/dia, com ausência de mudanças em fatores neuropsicomotores. A elevada concentração de taninos se mostrou antidiarréica e reduziu acidez, náusea, pirose, enquanto não apresentou toxicidade
<b>In vivo</b>								
Atividade antiulcerogênica	Queiroga et al. (2000)	Brasil	Hexano, clorofórmio e acetona	Ratos Wistar machos	Extratos	50 mg/kg ou 100 mg/kg	-	O extrato se mostrou rico em friedelan-3-β-ol, friedelina e sua mistura não apresentou atividade antiulcerogênica,

								independentemente da concentração usada. A inibição das úlceras se deu na mesma magnitude que a solução salina empregada no estudo
	Ferreira <i>et al.</i> (2004)	Brasil	Água	Mucosa gástrica de sapos ( <i>Rana catesbeiana</i> )	Extrato aquoso	7-28 mg%	Desconhecido mas parece estar relacionado com a inibição da secreção de ácido gástrico de forma similar a cimetidina, além de antagonista da atividade dos receptores H2	As vias que o LAEMIL* e a cimetidina inibem a secreção ácida foram semelhantes e suas doses subterapêuticas também foram efetivas nesta redução. O principal alvo destes extratos foi a superfície serosa da mucosa
Atividade gastrointestinal	Baggio <i>et al.</i> (2009)	Brasil	Clorofórmio, metanol, água, etanol, acetato de etila, butanol	Camundongos Swiss fêmeas	Extratos	10, 30 e 100 mg/kg	Inibição da secreção ácida gastrointestinal e redução de formação de óxido nítrico	Extratos de espinheira-santa apresentaram altos níveis de flavonoides, os quais são capazes de reduzir a motilidade gastrointestinal através do

Atividades anti-inflamatória, antinociceptiva, e antiulcerogênico	Jorge <i>et al.</i> (2004)	Brasil	Hexano, acetato de etila	Ratos Wistar machos e fêmeas	80, 160, 320 e 640 mg/kg	Estímula a síntese de muco e/ou manunteção dos altos níveis de prostaglandina na mucosa gástrica	antagonismo muscarínico. A fração insolúvel do extrato de acetato de etila apresentou a maior inibição do esvaziamento gástrico e trânsito intestinal, com uma taxa de 56 e 49%, respectivamente, quando em concentração de 9,7 mg/kg
	Extratos na dose de 320 mg/kg	demonstraram um significativo efeito anti-inflamatório e antinociceptivo, além da capacidade de reduzir a severidade de lesões gástricas induzidas pelo método de estresse por restrição em frio					

Atividade anti-hiperglicêmica	Schindler <i>et al.</i> (2021)	Brasil	Etanol	Ratos Wistar machos	Extrato alcoólico	150 mg/kg, 300 mg/kg, 600 mg/kg	Redução do glicogênio hepático. Ampliação do colesterol HDL e tióis, além de reduzir as substâncias reativas ao ácido tiobarbitúrico (TBARS) nesta região	Os extratos foram efetivos em reduzir a glicemia dos ratos, além de apresentar efeitos antioxidantes no fígado e oxidantes no pâncreas
Atividade diurética (efeito anti-hipertensivo)	Leme <i>et al.</i> (2013)	Brasil	Etanol e água	Ratos Wistar machos	Extrato aquoso e alcoólico	30, 100 e 300 mg/kg (SEI*), 15, 50 e 150 mg/kg (Fr.H2O*), 10, 30 e 100 mg/kg (Fr.EtOAc*)	Relação direta com a ativação da via prostaglandina/AMPc e óxido nítrico/guanilato ciclase	A fração Fr.EtOAc* apresentou efeitos diuréticos, natriuréticos e hipotensivos pela via prostaglandina/AMPc, causando diurese em apenas duas horas. O Fr.H2O* não foi capaz de produzir os mesmos efeitos. Não foram encontrados sinais de toxicidade provenientes dos extratos

Influência sobre o desenvolvimento neuronal/fertilidade	Ecker <i>et al.</i> (2017)	Brasil	Etanol e água	Ratos Wistar machos e fêmeas	Extrato aquoso e alcoólico	1360 mg/kg/dia	–	O tratamento com altas doses do extrato não causaram manifestações clínicas nas cobaias gestantes e machos.
Avaliação de segurança	Oliveira <i>et al.</i> (1991)	Brasil	Água	Machos de ratos Wistar e camundongos Swiss	Extrato aquoso	6.8; 170, 340, 680, e 1360 mg/kg	–	As diferentes concentrações testadas do extrato se mostraram praticamente ausentes de toxicidade. Os extratos não tiveram efeito sedativos quando administrados oralmente, mas esse efeito foi significativo pela via intraperitoneal. Efeitos analgésicos e anticonvulsivos não foram identificados

***In vitro***

Atividade antimicrobiana	Ducat <i>et al.</i> (2011)	Brasil	Etanol	Teste de suscetibilidade por difusão em ágar com culturas de microrganismos	Extrato alcoólico	10% (m/v)	Possível queação dos metais com flavonoides	O PCA* indicou que os níveis de compostos fenólicos, principalmente flavonoides, são importantes para a atividade contra <i>K. pneumoniae</i> , mas não combate o <i>S. aureus</i> .
Atividade antiviral	Kohn <i>et al.</i> (2012)	Brasil	Etanol	CRIB*, CER*, L929, CRFK*	Extrato alcoólico	0,25; 2.5, 25, e 250 µg/mL	Inibição dos receptores responsáveis pela adesão durante a replicação viral no estágio de adsorção	A atividade antiviral da espinheira-santa foi observada durante a etapa de adsorção do ciclo de replicação viral, com uma atividade de em potencial de 99% contra HBV-5*. Dos extratos testados, a <i>M. ilicifolia</i> apresentou a menor toxicidade junto ao maior potencial antiviral
Atividade anti-inflamatória	Wonfor <i>et al.</i> (2017)	Reino Unido	Hexano	Células Caco-2	Extrato	12,5; 25, 50, 100 ou 200 mg/mL	Possível redução dos sítios ativos de estímulo dos receptores de reconhecimento	A <i>M. ilicifolia</i> não demonstrou toxicidade às células Caco-2 e seus terpenos puderam reduzir

								de patógenos pela redução da expressão proteica de TLR-2*	lesões gástricas. Os efeitos anti-inflamatórios dos extratos foram ligados a regulação negativa a expressão de TLR2*, inibição da secreção de IL-8* quando aplicado o desafio de LTA*
Atividade antimutagênica	Horn, Vargas, 2003	Reino Unido	Água	Teste <i>Salmonella</i> /microsomo	Extrato aquoso	25, 75, 125, 250, 375 e 500 mg/matéria seca em placa	Oxidação de pró-mutágenos pela ativação do citocromo P450	A infusão a base de espinheira-santa teve a maior capacidade inibitória a substituição de pares de bases e lesões ao DNA. Nos extratos que estavam presentes os flavonoides e taninos, elevadas taxas de modulação desta atividade mutagênica, quando comparados a grupos contendo apenas taninos	

Legenda: \*LAEMIL: Extrato aquoso liofilizado das folhas de *M. ilicifolia*; SEI: Sobrenadante etanólico da infusão; Fr.H2O: Fração aquosa do sobrenadante etanólico da infusão; Fr.EtOAc: Fração orgânica do sobrenadante etanólico da infusão; MIE: Extrato de *M. ilicifolia*; PCA: Análise de Componentes Principais;

CRIB: Células derivadas do modelo Mardin-Darby de fígado bovino; CER: Células relacionadas ao embrião e galinha; L929: Linha celular de fibroblastos de camundongos; CRFK: Linha celular de fígado felino; BHV-5: Herpesvírus bovino tipo 5; TLR2: *Toll-like receptor 2*; IL-8: Interleucina 8; LTA: Ácido lipoteicoico  
Fonte: O autor (2024). Adaptado de: Baggio *et al.* (2009); Ducat *et al.* (2011); Ecker *et al.* (2017); Ferreira *et al.* (2004); Hom, Vargas (2003); Jorge *et al.* (2004); Kohn *et al.* (2012); Oliveira *et al.* (1991); Queiroga *et al.* (2000); Leme *et al.* (2013); Schindler *et al.* (2021); Tabach, Duarte-Almeida, Carlini, (2017); Wonfor *et al.* (2017)

As finalidades terapêuticas trabalhadas nos artigos puderam ser correlacionadas com os principais metabólitos secundários extraídos da planta. Os domínios fitogeográficos de todos os exemplares de *M. ilicifolia* também foram registrados, uma vez que bioma onde a planta teve seu crescimento pode ter um impacto significativo no perfil e nas concentrações de seus metabólitos secundários. A técnica usada para a extração desses compostos foi registrada, permitindo um entendimento de quais metodologias devem ser usadas para a otimização da extração. Todas as classes de metabólitos secundários e seus efeitos correlatos podem ser encontrados na **Quadro 3**.

Quadro 3 – Compostos encontrados em *M. ilicifolia*

Metabólito secundário	Proposta terapêutica	Referência	Local de colheita	Bioma	Compostos	Técnica utilizada	Modelo	Concentração
<b>Triterpenos</b>	Atividade antiulcerogênica	Queiroga <i>et al.</i> (2000)	Campinas, São Paulo, Brasil	Mata atlântica	Friedelan-3- $\beta$ -ol, friedelina	Cromatografia gasosa-espectroscopia de massas	Ratos Wistar machos	50 mg/kg e 100 mg/kg
	Atividade antihiperglicêmica	Schindler <i>et al.</i> (2021)	Pelotas, Rio Grande do Sul, Brasil	Pampas	Esqualeno	-	Ratos Wistar machos	150 mg/kg, 300 mg/kg, 600 mg/kg
	Atividade anti-inflamatória	Wonfor <i>et al.</i> (2017)	Irati, Paraná, Brasil	Mata atlântica	-	-	Células Caco-2	12,5; 25, 50, 100 e 200 mg/mL
<b>Taninos condensados</b>	Atividade antiulcerogênica	Ferreira <i>et al.</i> (2004)	-	-	-	-	Mucosa gástrica de sapos ( <i>Rana catesbeiana</i> )	7-28 mg%



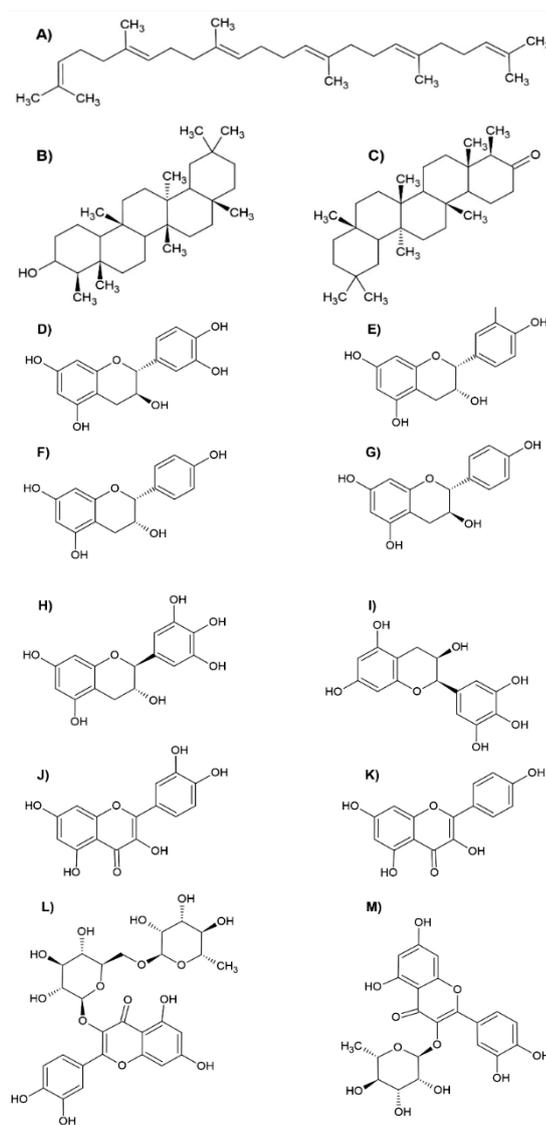


<b>Flavonoides, ácido gálico e ácido cafeico</b>	Influência sobre o desenvolvimento neuronal e fertilidade	Ecker <i>et al.</i> (2017)	Santa Maria, Rio Grande do Sul, Brasil	Mata atlântica /Pampas	Epicatequina, catequina, rutina, quercetina, kaempferol, ácido gálico, ácido cafeico	Cromatografia líquida de alta performance	Ratos Wistar machos e fêmeas	1360 mg/kg/dia
<b>Compostos polifenólicos, taninos condensados, flavonoides e triterpenos</b>	Atividades anti-inflamatória, nociceptiva e antiulcerogênica	Jorge <i>et al.</i> (2004)	Londrina, Paraná, Brasil	Mata atlântica	Compostos polifenólicos, taninos condensados, flavonoides e triterpenos	Testes de coloração, cromatografia em camada delgada, cromatografia fracionada em coluna com sílica gel, espectroscopia de ultravioleta, espectroscopia	Ratos Wistar machos e fêmeas	80, 160, 320 e 640 mg/kg



Os metabólitos secundários mais relevantes, de acordo com os artigos, para os efeitos de *M. ilicifolia* tiveram suas estruturas químicas evidenciadas na **Figura 10**.

Figura 10 - Estrutura molecular dos principais metabólitos secundários presentes na espinheira-santa



Legenda: A) = Esqualeno (Schindler *et al.*, 2021) B) = Friedelan-3- $\beta$ -ol (Queiroga *et al.*, 2000); C) = Friedelina (Queiroga *et al.*, 2000); D) = Catequina (Baggio *et al.*, 2009; Ecker *et al.*, 2017; Leme *et al.*, 2013); E) = Epicatequina (Baggio *et al.*, 2009; Ecker *et al.*, 2017; Leme *et al.*, 2013); F) = Epiafzelequina (Baggio *et al.*, 2009; Leme *et al.*, 2013); G) = Afzelequina (Baggio *et al.*, 2009; Leme *et al.*, 2013); H) = Galocatequina (Leme *et al.*, 2013); I) Epigallocatequina (Leme *et al.*, 2013); J) = Quercetina (Baggio *et al.*, 2009; Ecker *et al.*, 2017; Leme *et al.*, 2013); K) = Kaempferol (Baggio *et al.*, 2009; Ecker *et al.*, 2017; Leme *et al.*, 2013); L) = Rutina (Baggio *et al.*, 2009; Ecker *et al.*, 2017); M) = Quercitrina (Baggio *et al.*, 2009).

Fonte: O autor (2024)

#### 4.4.1.1 Efeitos gástricos

Como uma das protagonistas da medicina popular, a espinheira-santa é eficaz para diversos tratamentos. A razão desta capacidade se encontra nos metabólitos que compõe as folhas desta planta, como os triterpenos. Dentro deste grupo um destaque é dado ao friedelan-3- $\beta$ -ol e friedelina, compostos característicos de *M. ilicifolia* com reconhecida eficiência no tratamento de úlceras e gastrites estomacais, além de marcadores para certificação de veracidade da espécie vegetal (Mariot; Barbieri, 2006; Nossack *et al.*, 2000). Neste contexto, Queiroga e colaboradores desenvolveram uma metodologia para isolar e avaliar a eficácia antiulcerogênica de ambos estes compostos. Extratos de baixa polaridade a base de folhas de espinheira-santa foram obtidos e comparados a outras alternativas de tratamento, como ranitidina a 50 mg/kg, cimetidina a 100 mg/kg, com a concentração de friedelan-3- $\beta$ -ol a 10, 30 e 100 mg/kg e friedelina a 12,5; 25 e 100 mg/kg. Análises por cromatografia gasosa acoplada a um espectrômetro de massas evidenciaram a presença destes triterpenos em extratos a base de hexano. Os modelos *in vivo* demonstraram que ambos triterpenos friedelan-3- $\beta$ -ol e friedelina não foram capazes de inibir as lesões, independente da concentração utilizada (Queiroga *et al.*, 2000) (**Tabela 4**).

Tabela 4 – Efeitos da administração de friedelan-3- $\beta$ -ol e friedelina

Tratamento	Dose (mg/kg)	n	ULI (média $\pm$ S.E.M)	Inibição ULI (%)
Salina	-	5	36.1 $\pm$ 10.1	-
Ranitidina	50.0	5	4.3 $\pm$ 1.6*	88.9
Cimetidina	100.0	5	2.5 $\pm$ 0.4*	92.3
Friedelan-3- $\beta$ -ol	10.0	5	46.2 $\pm$ 11.8	-
Friedelan-3- $\beta$ -ol	30.0	5	44.4 $\pm$ 12.6	-
Friedelan-3- $\beta$ -ol	100.0	5	33.8 $\pm$ 8.0	-
Friedelina	12.5	4	38.8 $\pm$ 11.9	-
Friedelina	25.0	4	35.3 $\pm$ 10.8	-
Friedelina	50.0	4	36.5 $\pm$ 14.8	-
Mistura de 1 + 2*	300.0	5	32.6 $\pm$ 8.4	-

Legenda: \*p<0.05: Teste ANOVA seguido de teste de Duncan. n = Número de cobaias testadas;  
ULI = Índice de lesão ulcerativa; 1 - Friedelan-3- $\beta$ -ol; 2 - Friedelina.  
Adaptado de: Queiroga *et al.*, 2000.

Concluiu-se que os compostos em questão não são os responsáveis pelo efeito antiulcerogênico da espinheira-santa. Uma alternativa para aprimorar o estudo seria o desenvolvimento de um extrato a base de raízes de *M. ilicifolia*, uma vez que os triterpenos em questão são produzidos nas folhas, mas tendem a se translocar até as regiões mais baixas da planta, onde são transformados em maitenina e pristimerina (Mariot; Barbieri, 2006).

Ainda no tema da atividade antiulcerogênica da espinheira-santa, Ferreira e colaboradores investigaram a ação de um extrato aquoso liofilizado (LAEMIL) sobre a secreção gástrica de sapos (Ferreira *et al.*, 2004). A adição do LAEMIL à superfície serosa da mucosa gástrica foi capaz de reduzir a secreção de íons H<sup>+</sup> nesta região, mas também reduziu a secreção basal de ácido de forma rápida quando adicionada na superfície luminal, na ordem de  $5.9 \pm 0.3$  para  $3.9 \pm 0.2 \mu\text{Eq.g}^{-1}$  em 15 minutos. Além do efeito bloqueador, o extrato foi capaz de reduzir a secreção de ácido estimulada por fármacos como a histamina. A secreção ácida basal não foi reduzida de forma significativa com o LAEMIL (7 mg%) aplicado em baixas doses e individualmente, no entanto, quando administrado em conjunto com cimetidina (10 IM), o efeito é acentuado, indicando uma possível similaridade quanto ao mecanismo de inibição de ambos. É atribuída à presença de constituintes fenólicos, como taninos condensados, a atividade antiulcerogênica e redutora de secreção das folhas de espinheira-santa. O extrato liofilizado apresentou significativo efeito de redução gástrica concentração-dependente. Mesmo com o desconhecimento do mecanismo de ação do LAEMIL, o presente estudo indica que a rota de inibição da secreção ácida se assemelha ao de fármacos como a cimetidina, além de uma atividade antagonista aos receptores de histamina H<sub>2</sub> (Ferreira *et al.*, 2004).

*Monteverdia ilicifolia* tem sido empregada na medicina popular como uma alternativa para tratamento de distúrbios gastrointestinais. Investigando a influência que o uso desta planta traz sobre a motilidade desta função, Baggio e colaboradores baseiam seus resultados em modelos *in vivo* compostos por ratos Swiss. O estudo empregou uma purificação bioguiada com um extrato primordial rico em flavonoides a base de folhas da espinheira-santa, a fim de identificar os principais compostos ativos da planta. Formaram-se diversas frações a partir do solvente empregado, sendo eles o clorofórmio (CF), acetato de etila (EAF), butanol (BuF) e água (AqF). A porção EAF, quando tratada com 20 mL de acetato de etila, gerou mais duas frações também

estudadas, uma insolúvel (insF) e outra solúvel (solF). Uma investigação química a partir da espectrometria de massas permitiu estabelecer o perfil metabólico desses extratos, os quais se apresentaram ricos em taninos na forma de catequina, epicatequina, afzelequina e epiafzelequina, além de flavonoides glicosídicos de uma a quatro unidades de monossacarídeos, com quercetina e kaempferol como porção aglicona, flavonóides diglicosídicos (rutina e quercitrina) e triglicosídicos. Os extratos produzidos apresentavam concentração de 10 a 100 mg/kg. O tratamento com estas soluções promoveu o esvaziamento gástrico de forma mais significativa que fármacos anticolinérgicos como a atropina a 3 mg/kg e metoclopramida a 30 mg/kg (**Tabela 5**). O trânsito intestinal não demonstrou alteração pela administração da metoclopramida, porém os extratos apresentaram mudanças a nível de  $45 \pm 4\%$ , valor acima dos 38% encontrados para a atropina. A solução CF não demonstrou impacto para estes parâmetros, porém para os extratos AqF (40 mg/kg), BuF (28,1 mg/kg) e EAF (19,3 mg/kg), encontrou-se, respectivamente, 75 e 67%, 63 e 76%, e 58 e 72%. A fração solF (9,6 mg/kg) não apresentou influência sobre os fatores estudados, enquanto a fração insF promoveu o esvaziamento gástrico e reduziu o trânsito intestinal em 56 e 59%, respectivamente, demonstrando uma relevante eficiência quanto a dose-concentração empregada. Todas as frações se mostraram superiores a atropina e metoclopramida. Os achados do estudo corroboram para uma redução potente da motilidade gastrointestinal por mecanismos de antagonismo muscarínico (Baggio *et al.*, 2009).

Tabela 5 – Efeito comparativo das frações solF e insF quanto a motilidade gastrointestinal das cobaias

<b>Grupo de tratamento</b>	<b>Dose (mg/kg)</b>	<b>Trânsito intestinal (% do efeito)</b>	<b>Esvaziamento gástrico (% do efeito)</b>
Controle	-	$48.1 \pm 5.1$	$57.6 \pm 4.1$
solF	9.6	$35.9 \pm 8.6$	$72.2 \pm 9.7$
insF	9.7	$24.4 \pm 4.4^*$	$26.7 \pm 7.6^*$
Atropina	3	$24.2 \pm 2.1^*$	$32.5 \pm 3.6^*$
Metoclopramida	30	$54.1 \pm 1.7^*$	$79.5 \pm 6.3^*$

Legenda: \*Análise da variância (ANOVA), seguido pelo teste de Bonferroni ( $p < 0,05$ ).  
Adaptado de: Baggio *et al.* (2009)

É possível deduzir que os mecanismos envolvidos em outras funções da planta, como aumento da secreção gástrica e antiulcerativo, são distintos do objeto de estudo deste artigo. O extrato também mostrou efeito vasodilatador dependente do óxido nítrico. Efeitos espasmolíticos e anti-diarreicos também puderam ser percebidos, exigindo maiores estudos para comprovação. A fração insF apresentou cerca de 11 vezes mais quercetina quando comparado ao extrato rico em flavonoides do qual teve origem, enquanto o valor de taninos e demais flavonoides se mostrou reduzido. Esta redução não impactou a capacidade de ação do insF, o que permite deduzir a importância exclusiva da quercetina neste quesito. Estudos demonstram que este composto age a partir do relaxamento da musculatura gástrica lisa por meio de canais K<sup>+</sup>ATP e independe da via do óxido nítrico para tal, tendo um mecanismo ainda a ser elucidado (Modzelewska *et al.*, 2021).

#### 4.4.1.2 Atividade antimicrobiana e antiviral

Existe uma errônea concepção que tratamentos a partir de fitoterápicos se resumem a elucidação de alguma enfermidade sem a possibilidade de danos ao usuário no processo. Neste contexto, Ducat e colaboradores viram na composição metálica das plantas uma alternativa de estudo, buscando a quantificação de íons metálicos a fim de garantir a segurança pela ausência de compostos tóxicos como chumbo e cádmio. A riqueza de metabólitos das folhas também foi estudada a partir dos flavonoides e compostos fenólicos, ambos responsáveis por proporcionar efeitos antibióticos, os quais não se apresentaram totalmente elucidados. Ressalta-se como os íons metálicos também são reconhecidos por agirem em sinergia com esta classe de fármacos, sendo importante verificar o potencial desta combinação. Foi empregada a quimiometria na forma de Análise de Componentes Principais (PCA) para correlacionar os aspectos e similaridades entre as amostras testadas. O estudo se deu pelo uso de extratos etanólicos feitos à base das folhas da planta e apresentaram proporção de 70 e 95% (V/V), denominadas ES70 e ES95, respectivamente. A solução de 70% permitiu uma extração mais eficiente dos íons metálicos por conta de sua polaridade mais acentuada, porém, quantidades superiores de flavonoides e compostos fenólicos foram isoladas com o extrato ES95. Não foram encontrados os íons chumbo, cádmio, ferro, cromo e manganês em nenhum dos extratos. Constatou-se que a ES70 foi a única capaz de inibir o *Staphylococcus aureus*, com sua

concentração bactericida mínima (MBC) sendo de 25 mg/mL. A concentração inibitória mínima (MIC) desta solução foi de 0,25 mg/mL. Ambos os extratos foram eficientes para a *Klebsiella pneumoniae*, sendo o MIC semelhante àquele encontrado para a *S. aureus*. O MBC encontrado para esta bactéria, utilizando o extrato ES95, também foi de 0,25 mg/mL. Não se identificou atividade antimicrobiana contra os microrganismos *Pseudomonas aeruginosa* ATCC 27853, *Listeria monocytogenes* ATCC 19111, *Candida albicans* (cl clinicamente isolada). O modelo quimiométrico de PCA foi capaz de identificar corretamente os extratos ES95 e ES70 com um total de 94,73% da variância capturada dividida entre duas componentes principais (Ducat *et al.*, 2011).

O resultado dos estudos *in vitro* demonstrou que a associação dos íons metálicos aos metabólitos secundários da *M. ilicifolia* dão aos extratos uma atividade antimicrobiana em potencial. Metais tendem a se complexar e formar quelatos na presença de antibióticos, ampliando sua ação. Alguns dos íons mais conhecidos por realizar essa atividade são o magnésio, cálcio, cobre, alumínio e zinco, todos encontrados na metodologia proposta pelo estudo. É provável que os metais presentes na planta serviram como facilitadores para a ação dos compostos fenólicos, enquanto os flavonoides se complexaram aos íons pois são capazes de reduzir sua toxicidade (Khurana; Pulicharla; Brar, 2021; Seeder; Agarwal, 2010; Sutradhar *et al.*, 2023).

Seguindo a linha de enfermidades infecciosas, a espinheira-santa apresenta terpenos, flavonoides, compostos fenólicos, proteínas e outros metabólitos que possuem potencial reconhecido e tem sido utilizada para confecção de antivirais. O desenvolvimento de alternativas como essa vem da necessidade de contornar situações como resistência viral, recidivas e aprimorar tratamentos com elevados efeitos colaterais (Kohn *et al.*, 2012). Kohn e colaboradores realizaram estudos *in vitro* avaliando a atividade da *Monteverdia ilicifolia* contra herpesvírus bovino tipo 5 (HBV-5) e metapneumovírus aviário, infecções virais de consequências extremamente prejudiciais. As folhas da planta foram utilizadas para síntese de um extrato etanólico de 25 µg/mL de concentração. A formulação demonstrou 99% de atividade inibitória para o HBV-5, com um potencial de 90% de inibição no estágio de adsorção do vírus por meio do bloqueio dos receptores responsáveis por sua adesão, e nenhuma atividade no estágio de replicação. A concentração necessária para alcançar 50% de citotoxicidade viral foi próxima a 100 µg/mL. O mecanismo de inibição não foi

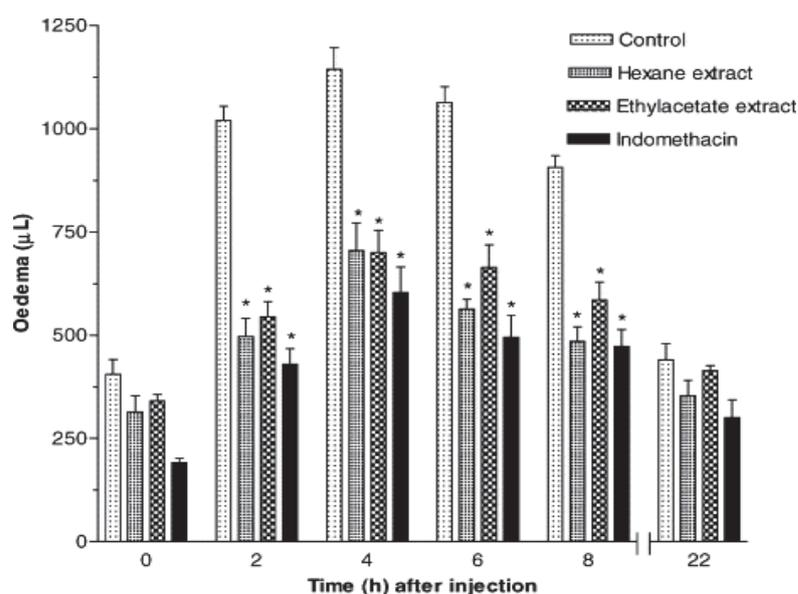
totalmente elucidado. Não foi identificada atividade durante a replicação viral contra parvovírus, betacoronavírus (MHV-3) ou vírus sincicial respiratório bovino (BRSV). Além da eficiência contra organismos virais, o extrato também apresentou baixa citotoxicidade e estabilidade (Kohn *et al.*, 2012). A espinheira-santa tem um potencial antiviral expressivo, porém, estudos adicionais são necessários para identificar seu mecanismo de ação, sendo a sua riqueza de flavonoides glicosídicos uma possibilidade, por conta da capacidade desta classe compostos de inibirem a entrada, replicação e translação de proteínas virais (Badshah *et al.*, 2021; Silva *et al.*, 2020).

#### 4.4.1.3 Atividade anti-inflamatória

Além da influência sobre o sistema digestivo e potencial antiviral, os flavonoides contidos nas folhas da espinheira-santa têm uma relação direta com reações inflamatórias, principalmente no sistema gástrico. Compostos como a quercetina e kaempferol pertencem a esta classe e protegem o organismo por inibir a secreção de enzimas e citocinas pró-inflamatórias como as lisozimas,  $\beta$ -glucuronidase, interleucina-1 $\beta$ , interleucina-6 e TNF-  $\alpha$ , além de regular a expressão de genes que promovem este efeito (Al-Khayari *et al.*, 2022; Ginwala *et al.*, 2019). A luz do exposto, Jorge e colaboradores ampliaram o espectro da busca para os efeitos anti-inflamatório, antinociceptivo e inibitório quanto a lesões gástricas da espinheira-santa (Jorge *et al.*, 2004). A investigação desses três efeitos se deu por conta da composição da planta que compreende os taninos condensados, triterpenos e flavonoides, trio de metabólitos relacionados com efeitos analgésico, anti-inflamatório, e antiulcerogênico, com eficácia equiparável à cimetidina. A partir das folhas da *M. ilicifolia* e com os solventes hexano e acetato de etila foram produzidos extratos de 80, 160, 320 e 640 mg/kg de concentração. Ambas as soluções tiveram seus efeitos comparados àqueles gerados pela indometacina a 20 mg/kg e cimetidina a 100 mg/kg em ratos Wistar e camundongos Swiss. Todas as soluções foram administradas de forma oral. Notou-se uma diferença relevante entre os tipos de extrato desenvolvidos quanto ao perfil fitoquímico: enquanto a formulação a base de hexano se viu rica em triterpenos, o extrato com acetato de etila apresentou principalmente compostos polifenólicos, taninos condensados e flavonoides. O modelo de nocicepção utilizado consistiu de edemas induzidos pela aplicação subcutânea de formaldeído à pata de

camundongos, reação controlada principalmente pela formulação de hexano a 640 mg/kg. Extratos a base deste solvente, na concentração de 320 e 640 mg/kg, foram mais eficientes na redução do edema em relação às outras formulações, com a primeira produzindo uma inibição das lesões de severidade 1+, 2+ e 3+, superior a própria indometacina. Esta eficiência não se manteve para as lesões de mucosas, as quais se apresentaram em quantidade similar ao dos animais tratados com cimetidina. As soluções de hexano e acetato de etila reduziram os traumas em, respectivamente, 71,30% e 57,13%, quando comparados ao controle, enquanto a indometacina ampliou este valor em 25,27%. Ambos os extratos na concentração de 320 mg/kg aumentaram o pH e o volume gástrico em relação ao controle e à indometacina. A cimetidina amplia a basicidade do estômago de forma semelhante, no entanto, não aumenta o volume gástrico de forma significativa como os extratos foram capazes. Os extratos nesta mesma concentração demonstraram potencial anti-inflamatório e antinociceptivo. Soluções à base de hexano com concentração de 640 mg/kg exibiram efeito nociceptivo superiores àquelas de 320 mg/kg, no entanto, a dose mais baixa também se mostrou anti-inflamatória, tornando-a mais indicada para o tratamento de úlceras gástricas (**Figura 11**) (Jorge *et al.*, 2004).

Figura 11 – Efeito dos extratos de *M. ilicifolia* a base de hexano e acetato de etila sobre o edema induzido em cobaias



Legenda: Os asteriscos (\*) indicam diferença significativa com o grupo controle ( $p < 0.05$ ).

Adaptado de: Jorge, et al. (2004)

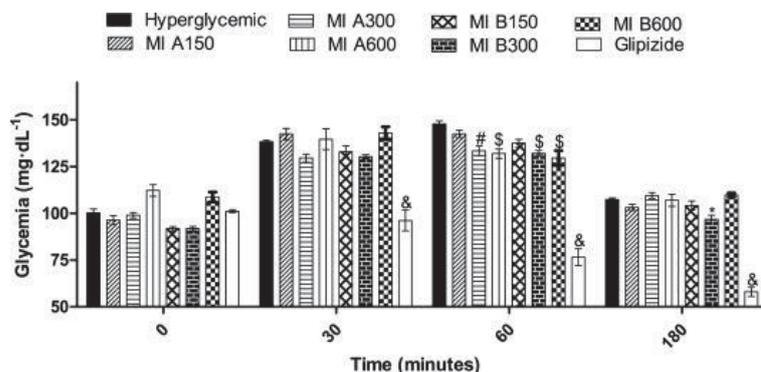
O estudo do potencial anti-inflamatório também foi o objeto de estudo de Wonfor e colaboradores, os quais avaliaram esta atividade para extratos a base de hexano da *M. ilicifolia* (Wonfor *et al.*, 2017). Soluções de concentração 12,5; 25, 50, 100 e 200 µg/mL foram testadas utilizando células Caco-2, verificando suas toxicidades, influência sobre a expressão de genes associados à cascata do receptor toll-like 2 (TLR2) e capacidade anti-inflamatória após desafio com ácido lipotecóico (LTA). A toxicidade foi determinada pela medição da resistência elétrica transepitelial (TEER) a partir de uma monocamada diferenciada. A aplicação da solução produzida manteve a integridade total da monocamada, demonstrando-se atóxica. Verificou-se que a expressão do TLR2 foi dose-dependente. Soluções de concentração 50 µg/mL tenderam a reduzir este gene, enquanto uma significativa diminuição foi identificada com a aplicação de extratos de 100 e 200 µg/mL para células sem LTA. Para células com o ácido o resultado foi semelhante, sendo a formulação de 100 µg/mL suficiente para criar uma tendência a redução da expressão do gene. Quando a 200 µg/mL verificou-se uma diminuição significativa, ao passo que aumenta a expressão do gene CXCL8. A secreção de interleucina-8 (IL-8) foi avaliada para determinação do potencial anti-inflamatório da espinheira-santa. Em células tratadas com LTA, os extratos de concentração 12,5; 25, 50 µg/mL ampliaram a secreção da citocina, enquanto nas soluções de 100 e 200 µg/mL este parâmetro não apresentou diferenças significativas. Considerando todos os resultados, a *M. ilicifolia* apresentou um efeito anti-inflamatório relevante e com baixa toxicidade. As células Caco-2 são naturais do intestino humano e o fato dos extratos atuarem sobre elas vai de acordo com estudos sobre a influência positiva que a espinheira-santa tem sobre o sistema digestivo. No entanto, é necessário avaliar o grau da ação anti-inflamatória desta planta em outros sistemas do organismo (Souza; Freitas; Storpirtis, 2007; Wonfor *et al.*, 2017).

Os triterpenos, presentes nas folhas de espinheira-santa, tem capacidade de inibir a expressão de citocinas pró-inflamatórias como a óxido-nítrico sintase induzível (iNOS), ciclooxigenase-2 (COX-2) e o fator de necrose tumoral-alfa (TNF-  $\alpha$ ), com mecanismos de ação variados de acordo com a cadeia e substituintes que compõe o metabólito. A riqueza deste composto no extrato com hexano o torna um alvo de estudos futuros a fim de verificar sua viabilidade terapêutica e importância clínica (Almeida *et al.*, 2019; Rios *et al.*, 2000; Miranda *et al.*, 2022).

#### 4.4.1.4 Efeito hipoglicemiante

A *M. ilicifolia* tem seu uso tradicional para diversas funções e, junto a função anti-inflamatória, este tem se expandido para o controle glicêmico. Explorando o potencial hipoglicêmico de extratos etanólicos das folhas da planta, Schindler e colaboradores utilizaram de dois acessos introduzidos em ratos Wistar machos (Schondler *et al.*, 2021). O grupo de estudo utilizou uma solução de tween a 1% como controle negativo, glipizida a 10 mg/kg como controle positivo e os extratos da planta (MIA e MIB, sigla para “*M. ilicifolia*” com a última letra representando qual o acesso utilizado) em concentrações de 150, 300 e 600 mg/kg. Os resultados alcançados indicam que os extratos foram efetivos para reduzir a glicemia dos ratos testados. Dentro de 60 minutos após sobrecarga de glicose, os animais tratados com 300 e 600 mg/kg de ambos os extratos tiveram seus níveis de glicose reduzidos em relação ao grupo hiperglicêmico sem tratamento. Após 180 minutos da sobrecarga de glicose, animais tratados com o extrato MIB a 300 mg/kg ainda apresentaram uma redução significativa da glicose sanguínea.

Figura 12 – Níveis de glicose sanguínea (mg.dl<sup>-1</sup>) em ratos hiperglicêmicos após tratamento com diferentes doses de extratos etanólicos de *M. ilicifolia*



Legenda: Níveis de glicose (mg.dL<sup>-1</sup>) em ratos hiperglicêmicos após o tratamento com diferentes doses dos extratos de *M. ilicifolia*. MIA: *M. ilicifolia* no acesso A; MIB: *M. ilicifolia* no acesso B. Os resultados são apresentados como média ± erro padrão e foi analisado por teste ANOVA de duas vias, seguido por teste pós-hoc de Bonferroni. \*p < 0.05; #p < 0.01; \$p < 0.001; &p < 0.0001 comparado ao grupo hiperglicêmico normal do respectivo tempo (n = 6).

Adaptado de: Schindler *et al.* (2021).

Consta-se que ambas as formulações em dose de 150 mg/kg foram capazes de reduzir substancialmente os níveis de glicogênio hepático das cobaias. O

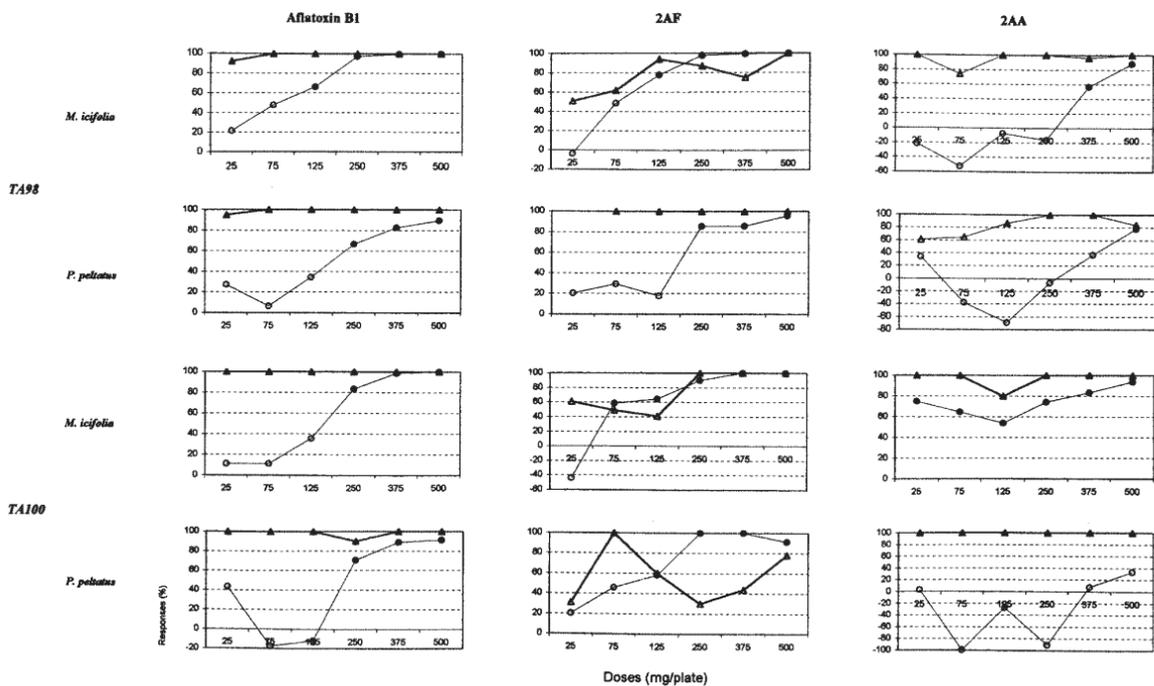
glicogênio muscular e a atividade de dissacarídeos não puderam ser reduzidos por nenhum dos extratos. Não foi possível definir o mecanismo de ação destas formulações. Quanto ao perfil lipídico, o colesterol total das cobaias não apresentou mudanças. Os extratos MIA 300, MIB 150 e 300 ocasionaram aumento nos triglicerídeos e VLDL, ao passo que as soluções MIA e MIB 300 ampliaram o colesterol HDL. Estas soluções também foram capazes de reduzir os níveis de TBARS e ampliar os tióis da região do fígado, agindo como um agente antioxidante. Danos ao fígado e rins não foram verificados. Os extratos em doses de 600 mg/kg podem ter influência indireta para predisposição de diabetes mellitus tipo 2 por conta do desequilíbrio de agentes antioxidantes no pâncreas. Estudos adicionais a fim de explicar o mecanismo dessas atividades são necessários, além de relacionar estes a composição da planta e quais os limites seguros para uma terapia plena (Schindler *et al.*, 2021).

#### 4.4.1.5 Atividade antimutagênica

Um dos principais causadores de mutações no corpo humano é o estresse oxidativo, fenômeno provocado pelo excesso de espécies reativas de oxigênio (ROS) em tecidos e células do corpo. O acúmulo destes compostos é conhecido por provocar disfunções em várias atividades fisiológicas, inclusive no DNA. O dano ao código genético ocasiona as chamadas mutações, eventos representados por alterações na sequência do DNA que provocam mudanças de resultado variado no organismo, sendo desde reações imperceptíveis a carcinomas. Diversos compostos tem sido estudados visando a interrupção deste processo danoso e/ou impedimento da formação dos ROS, sendo alguns deles a vitamina E, metabolitos polifenólicos e, fazendo parte da composição natural da espinheira-santa, os flavonoides (Clancy, 2008; Page, 2022; Pizzino *et al.*, 2017; Poetsch, 2020). Dentro deste contexto, a capacidade antimutagênica de uma infusão de folhas de *M. ilicifolia* foi o objeto de estudo de Horn e Vargas. Para fins de comparação também se avaliou o extrato de *Peltastes peltatus* (Vell.) Woodson, verificando o sinergismo entre ambas. Os resultados se basearam na capacidade dos extratos de minimizar a atividade de agentes promotores de mutações a partir de compostos como flavonoides e taninos. Uma significativa inibição da mutagenicidade pôde ser percebida por ambas as infusões no teste com *Salmonella*/microsomo, utilizando placas com bactérias

totalmente desenvolvidas. No entanto, extratos à base de *M. ilicifolia* apresentaram melhores resultados por conta de uma superior proteção a danos ao DNA e maiores inativações químicas. Este extrato apresentou tanto flavonoides quanto taninos, enquanto o extrato de *P. peltatus* apenas taninos, fator que influenciou na eficiência dos produtos. Quando aplicadas junto a 4-oxide-1-nitroquinoline (4NQO), as infusões apresentaram efeito co-mutagênico sinérgico, gerando lesões no material genético e citotoxicidade, enquanto na presença de outros agentes a inibição dos danos não foi significativa. Quando aplicadas separadamente, notou-se que a infusão de *M. ilicifolia* apresenta a melhor capacidade de inibição para diversos mutagênicos, em diferentes concentrações, como 125 ou 375 mg/placa para 2-nitrofluorene (2NF) e 500 mg/placa para azida de sódio (AZS) (Horn; Vargas, 2003). Uma comparação entre o perfil dose-resposta dos extratos de *M. ilicifolia* e *P. Peltatus* pode ser encontrado na **Figura 13**.

Figura 13 – Resultados dos ensaios de antimutagenicidade em revertantes/placa a partir de *M. ilicifolia* e *P. peltatus* contra mutagênicos conhecidos sem ativação metabólica



Legenda: 4NQO = 4- oxide-1-nitroquinoline; 2NF = 2-nitrofluorene; AZD = Azida de sódio.

Adaptado de: Horn, Vargas (2003).

Identificou-se um potencial significativo do extrato das plantas envolvidas, principalmente quanto à *M. ilicifolia*, para o combate a danos ao código genético. É possível deduzir que, pelo extrato da espinheira-santa ter apresentado um efeito

superior, os responsáveis pela proteção oxidativa foram os flavonoides. Citados anteriormente, a quercetina e o kaempferol são dois integrantes desta classe que tem efeito sobre esta valência, atuando, respectivamente, ampliando a apoptose de células e replicação de antioxidantes, e estimulando a atividade dos antioxidantes naturais do corpo contra os radicais livres. Efeitos gerais também atribuídos aos flavonoides são a inibição enzimas, quelação de células capazes de gerar os ROS e eliminação dessas substâncias (Roy *et al.*, 2022, Ullah *et al.*, 2020).

#### 4.4.1.6 Efeito anti-hipertensivo

A hipertensão é uma enfermidade de caráter crônico caracterizada pelo aumento da pressão arterial. Os medicamentos diuréticos tiazídicos são de primeira linha para este quadro devido a sua elevada eficiência que, no entanto, é acompanhada de efeitos adversos como disfunções de órgãos e células, desequilíbrio eletrolítico, entre outros. Uma estratégia que vem sendo empregada para contornar esse problema é a utilização de diuréticos naturais produzidos a partir de plantas. A espinheira-santa tem sido utilizada na medicina tradicional para este fim, porém não existem estudos que realmente coloquem esta atividade em evidência (Brasil, 2023; Reinhart *et al.*, 2023). Considerando este fato, Leme e colaboradores realizaram testes *in vivo* a fim de avaliar a atividade diurética aguda de infusões etanólicas (Fr.EtOAc) e aquosas (Fr.H<sub>2</sub>O) à base de folhas de *M. ilicifolia*. O estudo investigou o mecanismo de ação e uma possível função hipotensiva dos extratos. Enquanto a infusão Fr.H<sub>2</sub>O não apresentou potencial diurético, o extrato Fr.EtOAc foi capaz de alterar significativamente a função renal dos modelos testados. Ao analisar as frações do sobrenadante etanólico da infusão, notou-se que a porção Fr.H<sub>2</sub>O apresentou principalmente flavonoides glicosilados, enquanto na Fr.EtOAc se encontrou uma grande quantidade de proantocianidinas como a catequina, epicatequina, afzelequina, epiafzelequina, galotequina e epigalotequina. Nesta fração também foram encontrados glicosídeos de quercetina e kaempferol. Uma dose de Fr.EtOAc nas concentrações de 30 e 100 mg/kg foi capaz de aumentar a diurese após oito horas, além de ampliar a excreção de sódio na urina, agindo em grau semelhante a hidroclorotiazida (HCTZ). O Fr.EtOAc foi o único com efeito poupador de potássio e cloro, sem alterar parâmetros como pH, condutividade, densidade, creatinina, ureia,

sódio e potássio no plasma. O mecanismo de ação do extrato Fr.EtOAc demonstrou ter base na via prostaglandina/AMPC, uma vez que a administração intravenosa de anti-diuréticos como a indometacina e a 2',5'-*dideoxyadenosine* (DDA) reduziu a hipotensão produzida em cerca de 80%, enquanto a administração oral a restringiu totalmente. Pode-se concluir que extratos de espinheira-santa possuem compostos responsáveis por efeitos natriuréticos, diuréticos e hipotensores, sem toxicidade e de mecanismo semelhante a outros compostos já presentes na literatura (Leme *et al.*, 2013). Mesmo com a ausência do mecanismo de ação do extrato, é possível deduzir que os efeitos diuréticos são provenientes de taninos condensados e flavonoides. Estudos com catequinas e epicatequinas catalogam ambas como diuréticas a partir da interação com mediadores de vasodilatadores endógenos, inibição da atividade da CYP3A4 e epóxido hidrolase solúvel, e modulação do metabolismo do ácido araquidônico. Estudos com o flavonoide quercetina também chegaram a um efeito hipotensor relevante, porém sem elucidação de seu mecanismo de ação (Elbarry; Jones; Ung, 2022; Larson; Symons; Jalili, 2012; Mariano *et al.*, 2018).

#### 4.4.1.7 Fertilidade

O foco de uso da espinheira-santa se encontra nas regiões onde essa planta tem crescimento natural, como na Argentina e Paraguai. Dentro destes países, a *M. ilicifolia* também teve seu uso popularizado para diversos fins sem reais evidências científicas e clínicas. Um exemplo seria a função emenagoga, abortiva e contraceptiva da planta, a qual é explorada pelas mulheres a fim de evitar gestações (Montanari, Carvalho, Dolder, 1998). Por conta deste fato, Ecker e colaboradores exploraram a influência de extratos aquosos e etanólicos, desenvolvidos a partir das folhas de *M. ilicifolia*, sobre a fertilidade de ratos Wistar machos e fêmeas e desenvolvimento neuronal dos filhotes. A dose administrada dos extratos foi a nível supratrapêutico, com concentração de 1360 mg/kg/dia. Uma análise por cromatografia líquida de alta eficiência evidenciou a presença dos metabólitos epicatequina, catequina, ácido gálico, ácido cafeico, rutina, quercetina e kaempferol nos extratos. Não foram identificados efeitos tóxicos significativos nas cobaias. Uma redução no peso dos ratos machos foi verificada, mas sem alteração na massa dos órgãos. Ocorreu um aparecimento de hepatócitos eosinofílicos com vacúolos citoplasmáticos nas cobaias.

Análises físicas e neurocomportamentais dos ratos demonstraram conduta semelhante ao esperado e ausência de mudanças morfológicas do hipocampo. Não foram identificadas alterações na histoarquitetura do órgão reprodutor, espermatogênese e nem nos espermatozoides em si. Manifestações clínicas em ratas gestantes foram ausentes. A administração dos extratos não influenciou as funções reprodutivas das cobaias (Ecker *et al.*, 2017). A gestação das cobaias não foi afetada, assim como os filhotes provenientes desta, indicando que a administração de espinheira-santa é segura nestes estágios. Este estudo define que a aplicação tradicional de *M. ilicifolia* como um método contraceptivo não tem eficácia, sendo necessária a reeducação popular.

#### 4.4.1.8 Avaliação da segurança

O uso de fitoterápicos e seu controle são aspectos bastante negligenciados pois se criou um mito que a administração de produtos naturais não apresenta riscos à saúde. Neste escopo é importante determinar quais as reações adversas e limites de segurança relacionados com a administração de formulações disponibilizadas no mercado. Sendo assim, Tabach e colaboradores (2017) buscaram estabelecer a dose diária segura máxima de um extrato aquoso à base de *M. ilicifolia*. A formulação foi administrada de forma oral em 24 voluntários saudáveis, com idade entre 20 e 40 anos, avaliando possíveis mudanças em suas funções psicomotoras. A administração foi feita semanalmente, com um aumento gradativo da dose, por um intervalo de seis semanas. Ao longo do estudo, amostras de sangue e urina foram obtidas dos voluntários a fim de se observar possíveis alterações. O extrato seco apresentou metabólitos polifenólicos e taninos, os quais representaram aproximadamente 59,53% de seu conteúdo. O produto testado não provocou mudanças significativas quanto ao estado de alerta, parâmetros hematológicos e bioquímicos (níveis de creatinina, glucose, colesterol total, HDL, VLDL, LDL, triglicerídeos, ureia, enzimas de função renal, potássio e albumina). Os testes de reação visual simples, velocidade e acurácia tiveram a alteração de seus desfechos explicados pelo processo de aprendizado da tarefa a ser cumprida, uma vez que a realização de forma mais veloz e eficiente das atividades não sofreu influência do extrato. Nenhum dos pacientes necessitou interrupção do tratamento por conta de eventos adversos graves, sendo comuns a

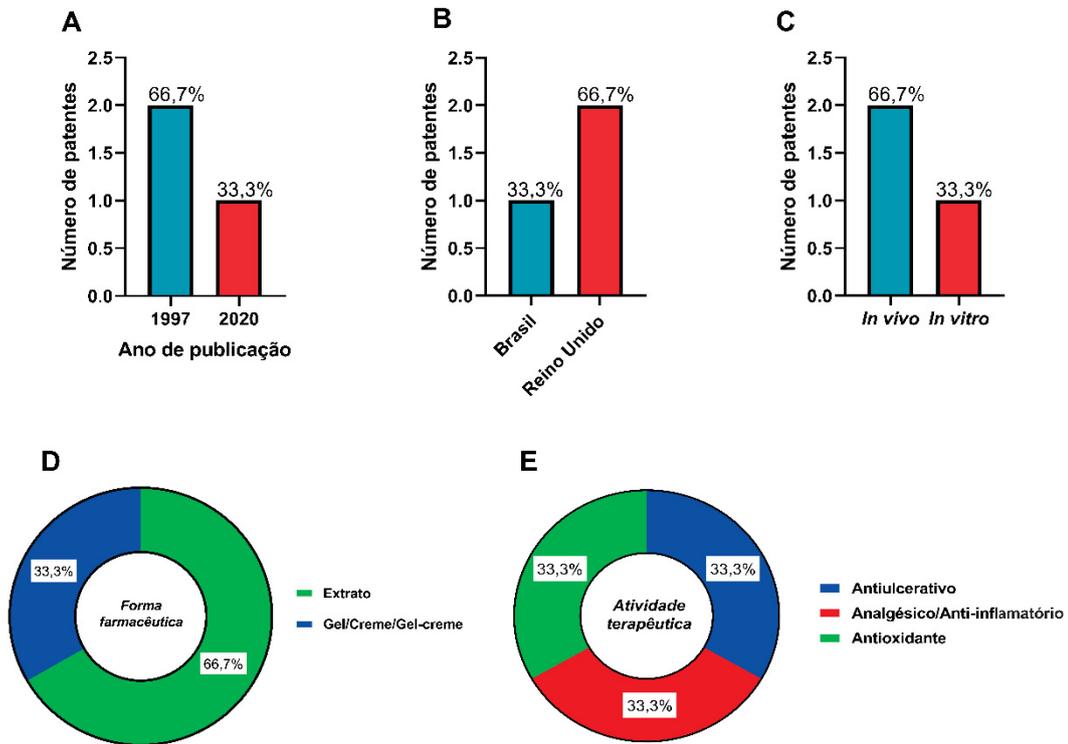
poliúria, observada em cinco pacientes, e xerostomia, observada em quatro pacientes. Ambas estas reações são consideradas de baixa gravidade e resolvidas de forma espontânea. Determinou-se que a utilização do extrato em doses até 2000 mg/dia é segura, com baixa toxicidade e efeitos adversos de severidade mínima (Tabach; Duarte-Almeida; Carlini, 2017).

A avaliação pré-clínica da farmacologia e toxicologia de *M. ilicifolia* também foi o objeto de estudo de Oliveira e colaboradores. Empregando extratos aquosos, com concentração de uma a 50 vezes de 6,8 mg/kg, em ratos Swiss, foram analisados efeitos agudos, crônicos e teratogenicidade da planta. Averiguou-se os efeitos agudos por triagem observacional, atividade locomotora, indução de analgesia, capacidade anticonvulsivante, dose letal mediana (LD50) e teste de desempenho *rota-rod*. Nenhum destes quesitos apresentou mudanças significativas após administração da formulação em diversas concentrações. Não foi identificada toxicidade dos extratos. A única alteração percebida ocorreu no teste de potenciação do tempo de sono com pentobarbital, na qual as três doses mais altas testadas se apresentaram estatisticamente diferentes do grupo controle. As formulações não impactaram no ganho de peso, nível de glucose e nem temperatura corporal das cobaias. Fatores hematológicos e quantificações séricas também não apresentaram mudanças significativas, com exceção da ureia e potássio, os quais se encontraram aumentados. A análise dos órgãos demonstrou ausência de alterações. A ingestão de fluidos das cobaias seguiu inalterada. Efeitos teratogênicos se mostraram ausentes. O número de filhotes por gestação, número e duração das gestações, e malformações congênitas seguiram sem alterações. Os filhotes se apresentaram saudáveis e sem alterações de peso, reflexos, deambulação e abertura de olhos. A espinheira-santa se demonstrou segura nos estudos pré-clínicos, sendo praticamente ausente de efeitos tóxicos em doses equivalentes a até 400 vezes aquelas utilizadas por humanos. Percebe-se que a *M. ilicifolia* demonstra segurança quanto ao seu uso diário aliada a diversos efeitos de eficácia pré-clínica e uso popular documentado, fatores estes que podem ter contribuído para o uso desta planta na medicina tradicional (Oliveira *et al.*, 1991).

#### 4.4.2 Prospecção tecnológica

De forma semelhante aos artigos, as patentes relevantes para este estudo tiveram suas publicações limitadas ao Brasil e Reino Unido, no entanto, enquanto o território brasileiro apresentou apenas um registro (1/3, 33,3%, Salvador *et al.*, 2020), a nação europeia teve dois (2/3, 66,7%, Nakamura, 1997a; 1997b) (**Figura 14**). As invenções exploraram os efeitos antiulcerativo (1/3, 33,3%, Nakamura, 1997a), analgésico, anti-inflamatório (1/3, 33,3%, Nakamura, 1997b) e antioxidante (1/3, 33,3%, Salvador *et al.*, 2020) (**Figura 14**). Os anos de publicação se limitaram a dois registros em 1997 (2/3, 66,7% Nakamura, 1997a; 1997b) e outro em 2020 (1/3, 33,3%, Salvador *et al.*, 2020) (**Figura 14**). Duas patentes estudaram o uso de formulações na forma de extrato (2/3, 66,7%, Nakamura, 1997a; 1997b), enquanto a outra teve sua inovação voltada ao desenvolvimento de géis/cremes/gel-creme (1/3, 33,3%, Salvador *et al.*, 2020) (**Figura 14**). Das três patentes, duas utilizaram de modelos *in vivo* para o desenvolvimento de suas metodologias (2/3, 66,7%, Nakamura, 1997a; 1997b) enquanto o modelo *in vitro* foi utilizado apenas uma vez (1/3, 33,3%, Salvador *et al.*, 2020) (**Figura 14**). Todas as inovações utilizaram as folhas da planta como farmacógeno. As informações principais das patentes foram descritas na **Tabela 6**.

Figura 14 - Características dos artigos incluídos no estudo



Legenda: A = Número de patentes publicadas ao longo dos anos; B = Número de patentes por região de publicação; C = Principal modelo experimental utilizado nas inovações; D = Forma farmacêutica desenvolvida nas patentes; E = Principal atividade biológica explorada na patente.

Fonte: O autor (2024)

Nota-se, no entanto uma grande distância entre o período de publicação de ambos, demonstrando uma possível reemergência de terapias fitoterápicas como alternativas de tratamentos. O desenvolvimento de patentes no Reino Unido pode estar relacionado com o fato de que esta região se apresenta como um polo inventivo de maior impacto que o território brasileiro, sendo constituído também de diversos países. Modelos *in vivo* foram mais comuns, uma vez que o objetivo destas inovações é o desenvolvimento de produtos fitoterápicos para humanos, com a maioria das patentes se encontrando nos últimos estágios de concepção. O extrato da espinheira-santa se mostrou a principal e mais versátil forma farmacêutica desenvolvida pelos estudos, sendo importante até para as formulações em gel/creme/gel-creme, servindo como componente para esta confecção. As quatro funções terapêuticas que alvejam os produtos são usos tradicionais da espinheira-santa, complementando as formas de administração disponíveis.

Tabela 6 - Principais informações extraídas das patentes

Referências	Proposta terapêutica	Forma farmacêutica	Solvente de extração	Concentração	Metabólitos secundários	Modelo	Principais achados
Nakamura et al. (1997a)	Antiulcerativo	Extrato	Metanol, água, clorofórmio e n-hexano	8, 24, 32, 48, 60, 90, 120, 180, e 360 mg/kg	Friedelan-3-β-ol	Ratos Wistar machos	O antiulcerativo produzido a partir dos extratos alcoólicos apresentou friedelan-3-p-ol como componente ativo para o tratamento de lesões na mucosa gástrica sem efeitos adversos relevantes
Nakamura et al. (1997b)	Analgésico e anti-inflamatório	Extrato	Metanol, água, clorofórmio e n-hexano	8, 24, 32, 48, 60, 90, 120, 180, e 360 mg/kg	Quercetin-3-O-glucoside, friedelina, friedelan-3-a-ol ou friedelan-3-p-ol	Ratos Wistar machos	Um analgésico/anti-inflamatório de elevada eficiência foi produzido a partir de um componente de camada aquosa obtido pela distribuição intersolvente do extrato alcoólico de <i>M. ilicifolia</i> , contendo quercetin-3-O-glucoside e friedelina/friedelan-3-a-ol/friedelan-3-p-ol como componentes ativos. O produto tem como vantagem a ausência de efeitos colaterais de medicamentos referência para os fins de tratamento propostos
Salvador et al (2020)	Antioxidante	Gel, creme, e gel-creme	Paraffinum Liquidum, água	2,6% m/m	Quercetin-3-O-glucoside e friedelina/friedelan-3-a-ol/friedelan-3-p-ol	Testes DPPH* e ORACFL*	A inovação consistiu da criação de formulações tópicas e capilares inéditas na forma de gel/creme/gel-creme. Todas apresentaram efeitos antioxidantes e anti-inflamatórios satisfatórios e tiveram a catequina, epicatequina e rutina como principais componentes ativos.

Legenda: \*DPPH: Teste 2,2-diphenyl-1-picrylhydrazyl ; ORAC<sub>FL</sub>: Capacidade de Absorção do Radical de Oxigênio com fluoresceína.  
Fonte: O autor (2024).

No primeiro registro, Nakamura e colaboradores desenvolveram um extrato à base de folhas secas e pulverizadas de *M. ilicifolia* buscando um produto antiulcerativo. As partes aéreas passaram por processos de extração com solventes alcoólicos e água, alcançando uma substância verde escura e oleosa. O principal componente da formulação foi o friedelan-3- $\beta$ -ol, composto responsável pelo efeito terapêutico. Os produtos desenvolvidos podem ser encontrados nas formas de pó, grânulos, drágeas, cápsulas ou preparações líquidas, sendo essa última administrada peroralmente, parenteralmente, entre outras. A forma parenteral pode ser introduzida ao organismo por injeções subcutâneas, intravenosas, intramuscular e até intranasal. Com doses na faixa de 0,1 à 500 mg/kg, dependendo de fatores como idade, sexo, sintomas, etc, a administração do produto deve ser realizada até cinco vezes por dia, alcançando uma dose máxima de 3000 mg/kg, a qual não demonstrou toxicidade. A abundância de friedelan-3- $\beta$ -ol conferiu a formulação um efeito antiulcerativo significativo aliado a um potencial anti-inflamatório, analgésico, intensificador de fatores protetivos, sem quaisquer efeitos adversos, tornando-o apto a atender às expectativas estipuladas (Nakamura, 1997a).

Seguindo a invenção anterior, Nakamura e colaboradores também utilizaram a espinheira-santa para produzir extratos com efeito analgésico-anti-inflamatório. O principal componente deste produto foi a *quercetin-3-O-glucoside*. A composição da formulação também contou com friedelina, friedelan-3-a-ol ou friedelan-3-p-ol, compostos extraídos durante o processamento das folhas e também responsáveis pelo efeito desejado. A administração da formulação se mostrou segura até 3000 mg/kg, recomendando-se a ingestão de 0,1 até 500 mg/kg, de uma até cinco vezes ao dia. A formulação teve enfoque na preparação líquida para aplicação parenteral subcutânea, intravenosa, intramuscular e/ou intranasal. O produto foi capaz de reduzir dores de cabeça, abdominais, neuralgias, dores causadas por cânceres, entre outras. Além da função analgésica-anti-inflamatória, foi identificada uma ação antiulcerosa, sugerindo um potencial para o tratamento simultâneo de diversas enfermidades. Vale ressaltar que o produto teve o foco em associar um efeito potente com a redução de efeitos adversos característicos de medicamentos referência da classe dos anti-inflamatórios não esteroidais e analgésicos (Nakamura, 1997b).

O extrato seco das folhas de *M. ilicifolia* foi utilizado por Salvador e colaboradores (2020) para desenvolver uma formulação farmacêutica tópica e capilar

inédita para essa espécie vegetal, com capacidade antioxidante e anti-inflamatória. Constituído principalmente por terpenos e compostos fenólicos, o produto foi processado a fim de padronizar o teor de catequinas, epicatequinas e rutinas, três compostos de efeitos antioxidantes e anti-inflamatórios. Para desenvolver a formulação empregou-se a técnica de secagem em leito de jorro e o adjuvante dióxido de silício coloidal em proporções de 40, 60 ou 80% (Mi40, Mi60 e Mi80). O extrato fluido teve sua concentração padronizada em 2,6% (m/m). Desenvolveu-se as formas farmacêuticas de creme, gel e gel-creme para aplicação na região da pele e cabelos. A preparação hidroalcoólica demonstrou ausência de toxicidade. A atividade antioxidante foi testada a partir de ensaios *in vitro* como a avaliação de redução do radical DPPH, quantificação de componentes fenólicos totais, ensaio ORAC<sub>FL</sub> e ensaio CCDC com  $\beta$ -caroteno. Dentre todas as concentrações de dióxido de silício coloidal, os resultados para Mi40 aliado a 10% do extrato seco foram os mais promissores e uniram uma maior capacidade antioxidante à maior concentração de compostos fenólicos, principalmente quando incorporado ao gel (Salvador *et al.* 2020).

Em ambos os registros, Nakamura e colaboradores (1997) buscaram desenvolver produtos de administração parenteral com enfoque na ação dos metabólitos *quercetin-3-O-glucoside*, friedelina, friedelan-3-a-ol e friedelan-3-p-ol. Por sua vez, Salvador e colaboradores exploraram a possibilidade de confecção de um gel/creme de propriedades curativas para aplicação capilar e cutânea, forma inédita de apresentação e não registrada como uma alternativa na RENAME atualmente (Brasil, 2022b). Ambos os autores focaram em extração hidroalcoólica, visando completo reaproveitamento dos metabólitos secundários responsáveis pelos efeitos desejados, os flavonoides, taninos e triterpenos. Todas as patentes obtiveram resultados positivos quanto a produção de suas formulações e expandiram as formas de apresentação e fins terapêuticos a qual *M. ilicifolia* é destinada.

## 4.5 CONCLUSÃO

A busca por ensaios clínicos e desenvolvimentos tecnológicos envolvendo a *M. ilicifolia* indicou uma ausência de diversas informações. Apesar de ser uma espécie vegetal presente na medicina tradicional, várias funções desta ainda não foram elucidadas.

Por meio dos artigos foi possível rastrear seus efeitos e correlacioná-los com os principais metabólitos secundários encontrados na planta. Triterpenos como friedelan-3- $\beta$ -ol, friedelina e esqualeno foram associados principalmente às atividades antiulcerogênica, anti-hiperglicêmica e anti-inflamatória. Os flavonoides catequina, epicatequina, galocatequina, epigalocatequina, epiafzelequina, kaempferol, quercetina, rutina e quercitrina foram os principais relacionados a atividades antivirais, antimutagênicas e antimicrobianas, além de não se apresentarem danosos ao desenvolvimento neuronal e fertilidade. Por sua vez, os taninos resultaram em efeitos como diurético, antimutagênico e melhora na atividade gastrointestinal. Nenhum estudo encontrou toxicidade nem teratogenicidade nos extratos da espécie vegetal. As patentes indicaram a versatilidade em termos de preparações à base de espinheira-santa, explicitando o processo de produção de géis, cremes, gel-creme e líquidos para aplicação parenteral, combinados com fins terapêuticos como analgésico, anti-inflamatório, antioxidantes e agentes antiúlcera. Estes registros servem como respaldo científico para futuras inovações contendo *M. ilicifolia*.

Embora na literatura sejam encontradas evidências promissoras quanto as atividades biológicas da espinheira-santa, diversos mecanismos de ação ainda são desconhecidos e existe uma lacuna sobre o desenvolvimento de formulações para o uso clínico da espécie vegetal. É importante observar que a *M. ilicifolia* ainda requer mais estudos e pesquisas para gerar evidências científicas de qualidade, evitando se basear apenas no uso popular tradicional. O reduzido número de patentes também destaca a necessidade de se realizar mais pesquisas a fim de entender completamente o potencial da espinheira-santa.

## **CAPÍTULO 2: ESTUDO QUIMIOMÉTRICO DA ESPINHEIRA-SANTA E ESPÉCIES SIMILARES**

## 5 CAPÍTULO 2

### 5.1 DESCRIÇÃO GERAL

#### 5.1.1 Quimiometria

A classificação e distinção da espinheira-santa de outras espécies morfológicamente similares comumente utilizadas para falsificação tornam-se possíveis pela ferramenta baseada na união interdisciplinar entre matemática, estatística e química: a quimiometria (Ferreira, 2015). Essa técnica consiste da seleção e otimização de informações provenientes de análises diversas, o que permite extrair resultados de forma lógica e garantir sua relevância. Seu principal objetivo é selecionar as informações mais úteis em um conjunto de dados e retratá-las de forma mais pontual, manipulando os resultados até que dezenas de relações possam ser estabelecidas. Essa correlação entre as variáveis analisadas é um dos principais atrativos do método, uma vez que permite associar conhecimentos diversos em grandes conjuntos de dados (Vékey; Telekes; Vertes, 2008; Castro, 2022).

Divide-se a quimiometria em duas categorias principais: reconhecimento de padrões e calibração multivariada. Estudos também são divididos em supervisionados e não supervisionados. O primeiro se classifica como qualitativo e se enquadra dentro do grupo de reconhecimento de padrões, sendo essencial para a construção de modelos focados na identificação e discriminação. Por sua vez, modelos não supervisionados são mais simples, e não exigem conhecimento prévio das classes envolvidas no estudo pelo *software*, baseando-se apenas nos dados contidos na matriz. A Análise de Componentes Principais (PCA) é o método protagonista desse tipo de estudo. Trata-se de uma abordagem exploratória que permite uma aprimorada análise dos dados, identificando padrões na matriz por meio da redução de dimensionalidade e detecção de *outliers*. O PCA é ideal para a classificação de amostras sem conhecimento prévio. No contexto de análises por espectroscopia de infravermelho, os modelos PCA identificam as regiões mais relevantes do espectro e as utilizam para classificar as amostras que compõe uma matriz de dados (Ferreira, 2015; Sarker; Nahar, 2015; Lima; Barbosa, 2019).

Em um conjunto de dados com classes previamente definidas e diversas, destaca-se a Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA). Esta metodologia discriminatória supervisionada baseia-se no popular modelo de regressão Mínimos Quadrados Parciais (PLS), o qual relaciona uma variável alvo (matriz Y) com uma variável mensurável (matriz X). O objetivo do PLS-DA é discriminar as classes dos dados que compõem uma matriz e determinar a qual grupo novas amostras pertencem. A aplicação prévia do método PCA auxilia no reconhecimento das classes e facilita o desenvolvimento do PLS-DA. (Lavine; Rayens, 2009; Singh *et al.*, 2013; Ferreira, 2015; Lima; Barbosa, 2019).

A quimiometria é aplicada de várias formas no ramo farmacêutico, especialmente através da espectroscopia por infravermelho, auxiliando na interpretação de dados. Outros métodos em que esta ferramenta é aplicada incluem a cromatografia líquida de alta eficiência (CLAE/HPLC), espectrometria de massas, estudos de dissolução e avaliação de outras características de medicamentos, como fluidez de pós, entre outros (Singh *et al.*, 2013).

Retratadas como organismos vivos, as plantas possuem um metabolismo complexo e de rica composição química. Dentro do contexto clínico, o protagonista deste arranjo são as folhas, órgãos vegetais especializados e responsáveis por diversas funções da planta por conta de seus metabólitos secundários, estruturas químicas variadas que servem como base para a atividade de plantas medicinais. A garantia da veracidade dessas plantas é dificultada por conta da existência de espécies com aparência foliar semelhante e microscopia similar, em determinadas formas farmacêuticas. Torna-se necessária a criação de métodos analíticos capazes de individualizar estas espécies, estabelecendo um *fingerprint*. A espectroscopia é uma metodologia empregada para alcançar este fim, produzindo espectros dentro de faixas de comprimento de onda previamente definidas. A dificuldade deste recurso reside na impossibilidade de interpretação íntegra da resposta gerada ao olho humano, sendo necessário o uso de ferramentas estatísticas para tal. Mesmo com todos esses empecilhos, a quimiometria é capaz de suprir essa demanda através da criação de diversos modelos.

Diante do exposto, o presente trabalho consistiu na aplicação da quimiometria, associada à inteligência artificial, para a construção de um método inovador capaz de

identificar a espécie *M. ilicifolia* e distingui-la de espécies de morfologia foliar similar, utilizando, respectivamente, modelos do tipo PCA e PLS-DA.

A técnica quimiométrica englobou toda a cadeia de processamento das plantas. Cada uma das espécies contidas no estudo foi analisada por espectroscopia de infravermelho médio com transformada de Fourier (MIR-FTIR) nas formas de folha íntegra, pó e extrato. Para cada matriz de dados foi desenvolvido modelos de PCA e PLS-DA, com o objetivo de construir métodos robustos e de elevada acurácia, capazes de detectar falsificações/adulterações em fitoterápicos à base de espinheira-santa, independente da forma de apresentação. Esses métodos foram desenvolvidos a fim de serem aplicáveis aos diversos produtos disponíveis no mercado, uma vez que, embora seja possível processar as folhas para obter as forma de pó e extrato, o processo inverso não é viável.

### 5.1.2 *Machine Learning*

Há uma crescente tendência mundial sobre o uso de tecnologia, e uma que evoluiu exponencialmente nos últimos anos foi a inteligência artificial, especialmente o *machine learning*, termo em inglês que se refere ao aprendizado de máquina. Essa ferramenta auxilia no gerenciamento de dados, permitindo a extração eficiente de informações de diferentes tipos de matrizes disponíveis. Sua nomenclatura se baseia na evolução constante e progressiva da máquina à medida que mais informações são fornecidas (Sarker, 2021).

Esta metodologia se baseia na criação de modelos analíticos, chamados algoritmos, os quais são inicialmente treinados com uma matriz de dados complexa, porém com relações conhecidas, para estabelecer a lógica pela qual a máquina deve seguir. O resultado é um sistema inteligente capaz de realizar previsões, estabelecer ligações entre os objetos de estudo e interpretar padrões observados (Janiesch; Zschech; Heinrich, 2020).

A versatilidade do *machine learning* permite a criação de dezenas de modelos diferentes para uma mesma matriz de dados. No entanto, é essencial selecionar aquele que melhor se adequa ao objetivo do estudo. Alguns pontos relevantes para essa escolha incluem a natureza dos dados disponíveis, o tamanho da matriz de dados e o desfecho desejado. Esse caráter multifacetado do aprendizado de máquina

o torna apto a atender demandas de diversos ramos, incluindo a identificação e discriminação de produtos variados, como espécies vegetais (Nichols; Chan; Baker, 2018).

O uso desta tecnologia se estende a dezenas de áreas, como a clínica, com o desenvolvimento de métodos computacionais de detecção de metástases, tratamento de enfermidades cardiovasculares e descoberta de novos medicamentos. Essa técnica não tem sua utilidade limitada ao campo farmacêutico, também sendo empregada nas áreas de cibersegurança, agricultura, nanotecnologia e economia (Pugliese; Regondi; Marini, 2021).

No contexto deste estudo, a metodologia de aprendizado de máquina segue a técnica quimiométrica em termos de progressão. À medida que mais testes e previsões são realizados, tanto a análise quanto a eficiência de ambos melhoram. A aplicação conjunta dessas metodologias visa investigar suas similaridades e complementaridades, aprimorando a qualidade do uso de inteligência artificial para a determinação de amostras vegetais.

## 5.2 OBJETIVOS

Desenvolver e validar um método integrado baseado em MIR-FTIR combinado com técnicas de quimiometria e aprendizado de máquina para a identificação da espinheira-santa e prováveis contaminantes.

### 5.2.1 Objetivos específicos

- Coletar amostras representativas de espinheira-santa e seus contaminantes e transformá-las em pó e extratos hidroalcolóicos;
- Analisar as folhas íntegras, os pós e os extratos das espécies por MIR-FTIR;
- Desenvolver metodologia quimiométrica para identificação (PCA) e discriminação (PLS-DA) das espécies vegetais;
- Complementar a técnica quimiométrica a partir de algoritmos de *machine learning*;
- Realizar a validação dos métodos propostos, avaliando métricas como precisão, sensibilidade e especificidade.

## 5.3 MATERIAIS E MÉTODOS

### 5.3.1 Amostras para o desenvolvimento do modelo quimiométrico

O desenvolvimento metodológico pela quimiometria foi realizado pela análise de seis espécies vegetais distintas: *Monteverdia ilicifolia* (Mart. ex Reissek) Biral (E1), *Monteverdia aquifolium* (Mart.) Biral (E2), *Sorocea bonplandii* (Baill.) W.C. Burger et al. (E3), *Citronella gongonha* (Mart.) R.A. Howard (E4), *Jodina rhombifolia* (Hook & Arn.) Reissek (E5) e *Zollernia ilicifolia* (Brongn.) Vogel (E6). Todas essas espécies apresentam similaridades morfológicas. A espécie E1 foi obtida de três localidades distintas, enquanto E2 foi adquirida de dois locais diferentes (**Tabela 7**). A espécie E3, por sua vez, foi obtida da mesma localidade, porém de dois lotes diferentes, sendo necessário dividir ambos para evitar alterações provenientes do tempo de colheita. As análises foram realizadas em três formas: folha íntegra, triturada (pó) e extrato seco. O estudo foi realizado utilizando diferentes apresentações das plantas, a fim de representar todas as formas disponíveis no mercado.

Tabela 7 - Número de amostras para cada espécie analisada

Espécie	Número de amostras	Código herbário	Localização
<i>Monteverdia ilicifolia</i> (Mart. ex Reissek) Biral	3	ECT9000	Ponta Grossa, PR
		BGE107	Esmeralda, RS
		BGE105	Vacaria, RS
<i>Monteverdia aquifolium</i> (Mart.) Biral	2	ECT9001	Ponta Grossa, PR
		BGES106	Esmeralda, RS
<i>Sorocea bonplandii</i> (Baill.) W.C. Burger et al.	2	ECT0008979	Pelotas, RS
<i>Citronella gongonha</i> (Mart.) R.A. Howard	1	UPCB 30838	Curitiba, PR

<i>Jodina rhombifolia</i> (Hook & Arn.) Reissek	1	ECT0008978	Pelotas, RS
<i>Zollernia ilicifolia</i> (Brongn.) Vogel	1	MBM 568533	Antonina, PR

Fonte: O autor (2023)

Todas as espécies utilizadas foram cedidas pela Universidade Estadual de Ponta Grossa a partir de coleções reconhecidas. Os espécimes foram identificados pelo taxonomista Dr. Gustavo Heiden. A autenticação e controle de qualidade das espécies vegetais incluídas no estudo foi também realizada pela equipe da Universidade Estadual de Ponta Grossa, liderada pela Professora Doutora Jane Manfron. A identificação das espécies foi realizada por uma sequência de procedimentos que envolveram a análise morfológica das folhas e caules, microscopia óptica e eletrônica de varredura, análises histoquímicas e espectroscopia de energia dispersiva com raios X (Antunes *et al.* 2023b). O acesso ao material botânico foi aprovado e licenciado pelo Conselho de Gestão do Patrimônio Genético (CGEN)/ Sistema Nacional de Gestão do Patrimônio Genético (SISGEN) e registrado sob o código AE6D3C6.

### 5.3.2 Preparação das folhas

As espécies vegetais passaram inicialmente por um processo de seleção das folhas. As folhas foram acomodadas, de forma independente, em bandejas plásticas, para procedimento de secagem por 72 horas, em estufa a 40 °C (Nova Ética). Após a secagem, as estruturas vegetais foram acondicionadas em tubos do tipo Falcon de 25 mL. Os tubos foram identificados com a espécie presente e vedados com *insulfilm*, evitando a entrada de umidade externa. O armazenamento dos tubos foi realizado em local com ausência de luz e temperatura ambiente, reduzindo a influência de fatores externos.

### 5.3.3 Preparação dos pós

Os pós foram obtidos a partir da trituração das folhas previamente secas. Utilizando um moinho de facas (IKA A11 *basic*), as folhas foram trituradas separadamente até atingirem tamanho de partícula uniforme, verificado por um tamis de 32 mesh. Após a trituração, cada amostra de pó foi armazenada em tubos Falcon de 25 mL, identificados com a respectiva amostra e selados com *insulfilm* para evitar a entrada de umidade externa. Os tubos foram armazenados em ambiente de baixa umidade, sem insolação e mantidos à temperatura ambiente controlada.

### 5.3.4 Preparação dos extratos

Para a extração dos principais metabólitos secundários da espinheira-santa, 1,0 g de cada amostra, na forma de pó, foi pesada e armazenada em tubo Falcon de 25 mL. Essa massa foi solvatada em uma mistura água:etanol 70% (proporção 1:1 v/v) até completar 10 mL. As amostras foram deixadas em maceração por um dia, permitindo a extração dos compostos pela fase aquosa. Após 24 horas, as amostras foram submetidas a banho ultrassom por mais 30 minutos, finalizando a maceração.

Posteriormente, as amostras foram centrifugadas (5810 R, Eppendorf) por 30 minutos, separando o líquido sobrenadante contendo os produtos de extração do pó restante. O sobrenadante de cada planta foi transferido a tubos Falcon de 15 mL, de forma independente. Esses tubos foram então alocados em secador centrífugo CentriVap (*Acid-Resistant CentriVap Concentrator*, LABCONCO) a 30 °C por 24 horas, concentrando as amostras.

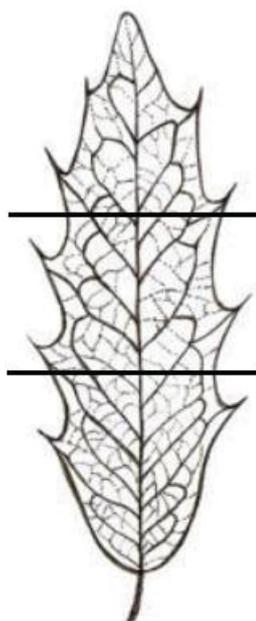
Após o processo de secagem, as amostras resultaram em um gel viscoso, consistente e escurecido. Os tubos contendo as amostras foram selados com *insulfilm* e armazenados em local de baixa umidade, protegidos da luz e mantidos à temperatura ambiente controlada.

### 5.3.5 Espectroscopia por infravermelho

A metodologia utilizada para análise das espécies vegetais foi a espectroscopia por infravermelho com transformada de Fourier, em nível médio (MIR-FTIR). As

análises foram realizadas utilizando espectrômetro Alpha-P Bruker com um cristal de diamante no modo ATR (Bruker OPTIK GmbH, Ettlingen, BW, Alemanha). As folhas de cada uma das espécies vegetais foram submetidas a 500 varreduras na região do infravermelho médio (4000 a 400  $\text{cm}^{-1}$ ), com 24 *scans* compondo cada varredura, em *software* OPUS. As estruturas foliulares foram retalhadas horizontalmente em três partes, e cada fragmento foi analisado separadamente, tanto na face anterior quanto posterior, a fim de identificar distinções (**Figura 15**).

Figura 15 – Método de corte para análise completa da folha de espinheira-santa



Fonte: O autor (2023). Adaptado de: Farmacopeia Brasileira 6ª edição (2019)

As amostras foram acomodadas sobre o cristal de diamante do equipamento, e, em seguida, pressionadas por um pistão, garantindo a passagem de luz pela estrutura vegetal.

Para as espécies vegetais *M. ilicifolia* (E1), *M. aquifolium* (E2) e *S. bonplandii* (E3), as quais apresentavam mais de uma localização ou tempo de colheita, o número de espectros foi dividido igualmente visando um total de 500 espectros (**Quadro 4**).

Quadro 4 - Estratégia de análise das folhas e pós para plantas adquiridas em mais de uma localidade

Identificação da amostra	Divisão das varreduras de acordo com as diferentes amostras		
<b>E1</b>	166	167	167
<b>E2</b>	250		250
<b>E3</b>	250		250
<b>E4</b>	500		
<b>E5</b>	500		
<b>E6</b>	500		

Fonte: O autor (2023)

Seguindo a metodologia aplicada às folhas, foram obtidos 500 espectros de cada espécie a partir dos pós. A divisão das amostras E1, E2 e E3 também foi semelhante ao método utilizado para as folhas. Uma espátula foi utilizada para aplicar os pós no equipamento, analisando-se tanto a porção superior quanto inferior do tubo, garantindo a representatividade íntegra da amostra. Os pós foram acomodados sobre o cristal de diamante e pressionados contra ele por meio de um pistão, de forma semelhante ao procedimento para as folhas íntegras.

Para a análise dos extratos, a metodologia foi semelhante à empregada para o estudo das folhas e pós, com um total de 100 varreduras realizadas para cada espécie. O número reduzido de varreduras foi determinado por conta da natureza mais concentrada do extrato, tornando desnecessária a aplicação de um número extremamente alto de análises para extração de informações. A estratégia para as espécies E1, E2 e E3 também foi ajustada para se adequar ao novo número de análises (**Quadro 5**).

Quadro 5 - Estratégia de análise dos extratos para plantas adquiridas em mais de uma localidade

Identificação da amostra	Divisão das varreduras de acordo com as diferentes amostras		
	E1	33	33
E2	50		50
E3	50		50
E4	100		
E5	100		
E6	100		

Fonte: O autor (2023).

Os extratos foram analisados de forma semelhante às folhas e pós, colocando-os sobre o cristal de diamante e pressionando com o pistão. A matriz de dados para as amostras na forma de folhas íntegras e pós foi composta por 3000 dados, enquanto a matriz dos extratos contou com 600 resultados.

### 5.3.6 Análise dos espectros obtidos

Os resultados obtidos foram organizados utilizando o *software Origin* (versão 9.6.5.169, 2019b) para torná-los compatíveis com o *software MatLab* (versão 7.5.0.342; R2007b, 2007), o qual foi empregado para o desenvolvimento dos modelos quimiométricos. O *software MatLab*, capaz de suportar uma variedade de estudos quimiométricos, permitiu a interpretação e multiprocessamento dos resultados pelos métodos exploratório PCA e discriminatório PLS-DA.

### 5.3.7 Desenvolvimento do método quimiométrico

Empregando o *software MatLab*, foram desenvolvidos os modelos quimiométricos exploratórios (PCA) e discriminatórios (PLS-DA), em sequência. O

*software* foi alimentado com os dados dos espectros, de forma independente, após tratamento e processamento destes. Os dados adicionados ao *software* foram submetidos a diferentes tipos de tratamentos para torná-los aptos para estudos futuros. O conjunto de processos aplicados aos dados buscou corrigir deslocamentos da linha de base, aumentar a razão sinal/ruído, precisão e exatidão, aprimorar a capacidade de interpretação dos resultados, entre outros.

Para desenvolver o método PCA que melhor agrega cada uma das espécies, diversos tipos de pré-tratamentos foram testados. Esses foram aplicados de forma individual e/ou combinada e puderam ser divididos em: **transformações**, sendo aplicadas às amostras (linhas da matriz X), como o **Smoothing** e **derivadas de primeira e segunda ordem**; e **pré-processamento**, aplicáveis às variáveis (colunas da matriz X) como **Mean center**, **Autoscale**, e **GLS Weighting** (com diferentes valores de alfa). Em situações de combinação, os pré-tratamentos foram experimentados em diferentes sequências, visando a otimização dos dados.

Embora apresentem diversas similaridades com o PCA, os modelos PLS-DA não são necessariamente interpretados corretamente com os mesmos tratamentos. Para o estudo discriminatório, aplicou-se a mesma lista de pré-tratamentos, porém com outras combinações. Aliado disso, foi estabelecida a aplicação da *cross-validation* a partir das seguintes metodologias:

1. *Leave one out*
2. *Venetian blinds*
3. *Contiguous block*
4. *Random subsets*
5. *Custom*

Os métodos de *cross-validation* foram aplicados a fim de se testar o modelo, dividindo parte dos resultados para calibração (70%, equivalente a 2100 dados da matriz) e o restante para validação (30%, equivalente a 900 dados da matriz) para as folhas e pós. No caso dos extratos, a calibração e a validação foram feitas com um número reduzido de dados (respectivamente, 70%, equivalente a 420 dados, e 30%, equivalente a 180 dados). A escolha da metodologia levou em consideração fatores como número de amostras, divisão dos dados, número de iterações, entre outros.

### 5.3.8 Critérios de avaliação para os modelos

A certificação da qualidade dos modelos foi realizada por meio da avaliação das amostras que compõem a matriz e resultados dos estudos estatísticos. Os dados foram analisados utilizando o gráfico *Scores Plot*, baseando-se em duas componentes principais. Estabeleceu-se um intervalo de confiança de 95% para as amostras serem válidas para inclusão no estudo. Espectros fora deste limite foram considerados como anômalos, sendo removidos a fim de se garantir maior precisão dos modelos.

Para o PCA, os métodos desenvolvidos foram avaliados com base no número de componentes principais a serem selecionadas, na capacidade do modelo de explicar a matriz de dados (*%variance cumulative*) e na separação das espécies a nível tridimensional. Além disso, o gráfico *Variables/Loadings* também fez parte dos critérios de aceitação do modelo.

A avaliação do PLS-DA considerou a seleção do número de variáveis latentes, a disposição tridimensional dos dados no gráfico *Sample/Scores Plot*, o gráfico *Variables/Loadings* e capacidade do modelo de explicar a matriz de dados (*%variance cumulative*). Todas essas variáveis culminaram em diferenças nos resultados estatísticos para cada coluna do bloco Y, representadas pelos valores de especificidade e sensibilidade a nível de calibração, *cross-validation* e predição, além dos valores de *Root Mean Square Error of Calibration* (RMSEC), *Root Mean Square Error of Cross-Validation* (RMSECV) e *Root Mean Square Error of Prediction* (RMSEP).

### 5.3.9 *Machine learning*

Complementando o estudo quimiométrico, foram desenvolvidos métodos de classificação utilizando aprendizado de máquina. Os dados de MIT-FTIR adquiridos anteriormente foram organizados em formato tabular, garantindo uma estrutura adequada para análise.

Para desenvolver os modelos de *machine learning*, os dados foram inicialmente divididos em conjuntos para treinamento (70% da matriz de dados) e validação (30% da matriz de dados). Em seguida, foram criados 40 algoritmos diferentes para

classificar as diversas espécies incluídas no estudo, utilizando para isso a biblioteca *LazyPredict* da linguagem *Python*.

Após essa etapa, selecionou-se o melhor modelo entre aqueles testados, com base nos valores de acurácia, *F1 Score*, índice *Kappa* e tempo de treinamento. O algoritmo que apresentou o melhor resultado passou por uma otimização de seus hiperparâmetros e validação cruzada utilizando 10 conjuntos de validação (metodologia *10 fold*).

O ambiente *PyCaret* em linguagem *Python* foi configurado para auxiliar na construção dos modelos, automatizando o processo, e especificando a variável alvo, *M. ilicifolia* (espinheira-santa verdadeira) e opções do pré-tratamento. Esta etapa é crucial para garantir a qualidade dos resultados durante o treinamento dos modelos de *machine learning*.

Durante a configuração do ambiente, foram aplicadas técnicas de pré-tratamento, incluindo a normalização dos dados para reduzir variações na escala das características, a seleção de características para identificar as mais relevantes para a classificação, e balanceamento de classes para lidar com desequilíbrios no número de amostras entre as espécies vegetais envolvidas no estudo.

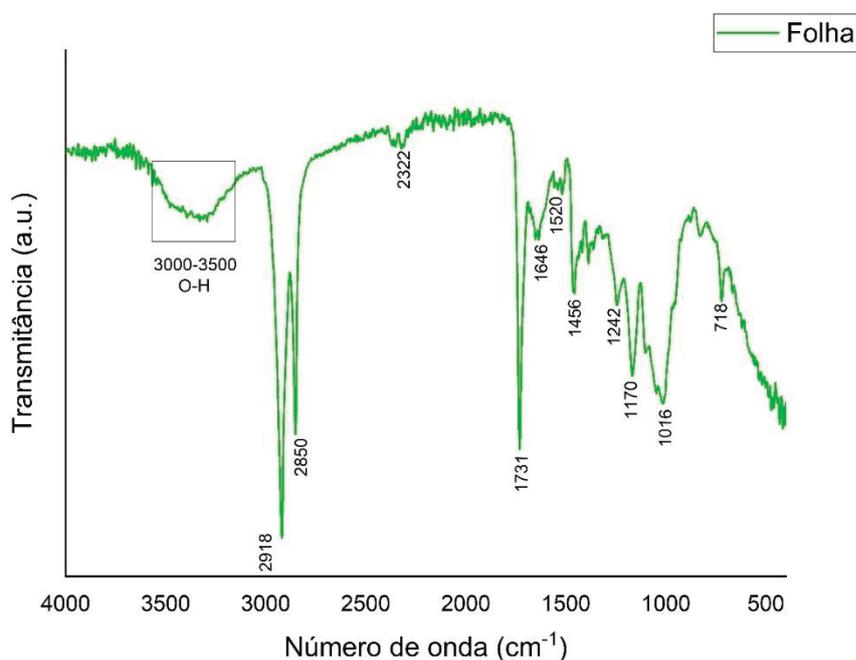
Após a configuração do ambiente, foram visualizados e comparados 40 tipos diferentes de modelos, buscando determinar o mais adequado para a matriz de dados disponível. Estes testes foram realizados utilizando a biblioteca *LazyPredict* da linguagem *Python*. A escolha do melhor modelo foi baseada em métricas como acurácia, *F1 Score*, índice *Kappa* e tempo de treinamento.

Com o modelo escolhido (*Random Forest*), a avaliação da capacidade preditiva do algoritmo foi realizada utilizando validação cruzada com metodologia de 10 conjuntos de validação (*10 Fold*), fornecendo os resultados nas métricas de acurácia, *Recall*, precisão, *F1 Score*, índice *Kappa* e Coeficiente de Correlação de Matthews (MCC). Gráficos de Área sob Curva ROC (AUC ROC), *classification report*, *class prediction error*, *validation curve*, *learning curve* e a matriz de confusão foram gerados para o compreender o desempenho preditivo do melhor modelo. Por fim, gráficos das variáveis (números de ondas) mais importantes foram construídos a partir da otimização do melhor modelo, visando entender quais grupos funcionais contribuem na diferenciação das espécies vegetais.

## 5.4 RESULTADOS E DISCUSSÃO

A primeira etapa do estudo quimiométrico consistiu na análise do padrão espectroscópico gerado para cada uma das formas de apresentação da planta, comparando-as. O estudo se iniciou com as folhas íntegras, estruturas de elevada complexidade. (**Figura 16**).

Figura 16 – Padrão espectroscópico de folhas íntegras de *M. ilicifolia*



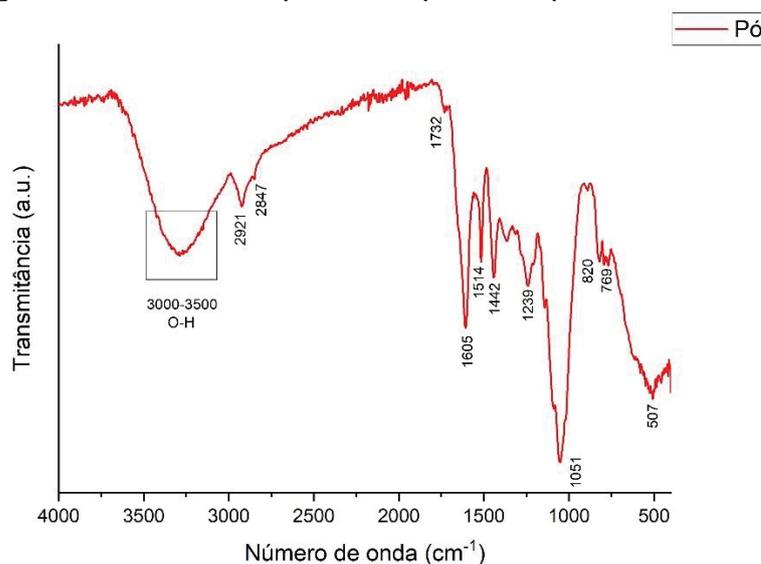
Fonte: O autor (2023)

As folhas de *M. ilicifolia* apresentam um perfil com diversas regiões de relevância mesmo com a padronização de seus cortes e áreas de visualização. A distribuição do espectro demonstrou-se deveras heterogênea na região anterior aos 1800 cm<sup>-1</sup>. Diversos picos importantes foram encontrados e suas presenças são essenciais para caracterização da espécie. Na região inicial do espectro e próximo a 500 cm<sup>-1</sup> se encontram os compostos aromáticos e halogênios que compõe a amostra. A região de 1100 cm<sup>-1</sup> teve bastante destaque e condiz com a deformação C-H de anéis aromáticos. Por sua vez, a região de 1700 cm<sup>-1</sup> está relacionada com o estiramento C=O de grupamentos éster que compõe os taninos hidrolisados, e a 2300 cm<sup>-1</sup> está relacionada com a presença de alquinos (C≡C). A 2900 cm<sup>-1</sup> - 2800 cm<sup>-1</sup> têm-se os picos indicativos do grupamento CHO, constituindo um aldeído, e o

estiramento assimétrico C-H dos grupamentos  $-CH_2$ , estiramentos simétricos C-H dos grupamentos  $-CH_3$  presente em moléculas de açúcares, e os grupamentos alifáticos, como  $-C-H$ ,  $-C-H_2$  e  $-C-H_3$ . Após este número de onda, a tendência do espectro foi permanecer homogêneo, com picos de menor intensidade quando comparados àqueles no início (Alves *et al.*, 2024; Martins, 2020; Shurvell, 2006).

Após a trituração das folhas, foi realizada a análise dos pós (**Figura 17**).

Figura 17 – Padrão espectroscópico dos pós de *M. ilicifolia*



Fonte: O autor (2023)

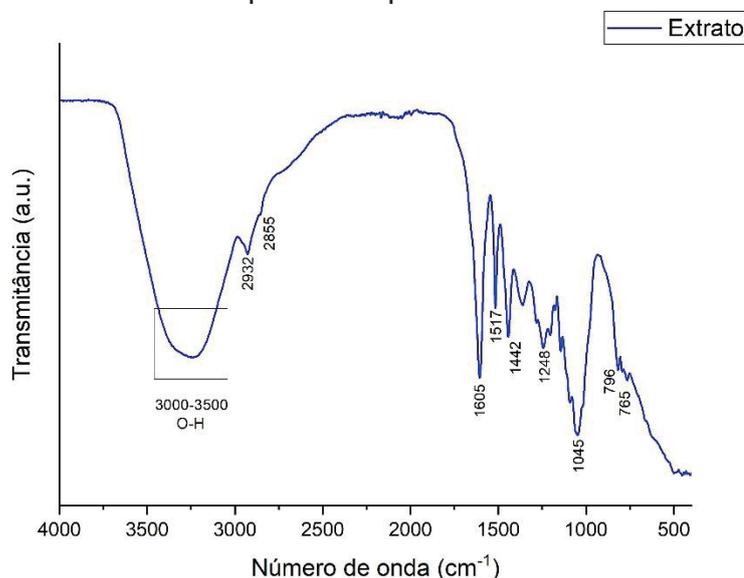
Para os pós, as regiões de maior destaque foram similares às das folhas, porém no número de onda de  $1700\text{ cm}^{-1}$  notou-se um pico de proporções muito superiores, novamente remetendo ao estiramento  $C=O$  encontrado nas folhas. A área inicial do espectro (entre  $400$  a  $1700\text{ cm}^{-1}$ ) se mostrou rica em picos, de forma semelhante às estruturas foliares anteriores. Dentro desta região, alguns dos pontos mais importantes incluíram a faixa de aproximadamente  $0$  a  $750\text{ cm}^{-1}$ , região voltada aos compostos aromáticos e halogênios identificados, e um novo destaque ao pico em  $1250\text{ cm}^{-1}$ , relacionado a grupamentos éter, álcoois e açúcares por conta de grupamentos como  $C-O-C$ ,  $C-OH$  e estiramentos  $C-O$  proveniente de anéis pirano, característicos de moléculas da classe dos flavonoides. Novamente destaca-se a região de  $1100\text{ cm}^{-1}$  com a deformação  $C-H$  de anéis aromáticos, e a  $2900\text{ cm}^{-1}$  -  $2800\text{ cm}^{-1}$  os picos dos grupamentos  $CHO$ , estiramentos assimétricos  $C-H$  dos

grupamentos  $-CH_2$ , simétricos C-H dos grupamentos  $-CH_3$ , além os grupamentos alifáticos  $-C-H$ ,  $-C-H_2$  e  $-C-H_3$ . (Alves *et al.*, 2024; Martins, 2020; Shurvell, 2006).

Houve um aumento coletivo na intensidade dos picos para as amostras em pó em relação às folhas. Esse fato se deve à diferente interação com a luz infravermelha de cada tipo de amostra, uma vez que o pó é mais homogêneo, possui maior superfície de contato e oferece menor impedimento à passagem de luz em comparação à estrutura foliar.

O último grupo de espectros a ser analisado foi o dos extratos (**Figura 18**), originados após a passagem dos pós pelos processos de maceração e extração.

Figura 18 – Padrão espectroscópico os extratos de *M. ilicifolia*



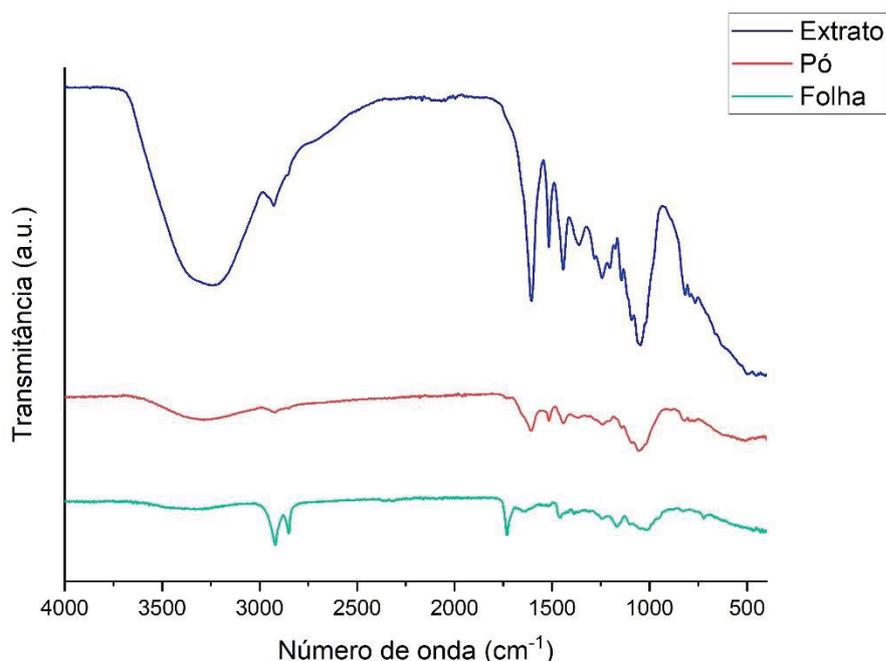
Fonte: O autor (2023)

O perfil espectral dos extratos demonstrou picos de elevada intensidade e mais bem definidos que aqueles apresentados para folhas e pós. Diferentemente dos demais, os extratos apresentaram uma distribuição homogênea ao longo de toda a região do infravermelho médio. O processo de extração concentra a amostra, reduzindo a influência de possíveis interferências por contaminantes e estruturas sem valor para o estudo, fato que contribui para um perfil mais uniformizado. Destacam-se os picos nas regiões de, aproximadamente,  $500\text{ cm}^{-1}$  a  $750\text{ cm}^{-1}$ , relacionados a grupos aromáticos e halogênios, de  $1000$  a  $1750\text{ cm}^{-1}$ , os quais incluem diversos estiramentos diferentes, como as ligações C-O, C=C-H, C=O, englobando ésteres, aldeídos, álcoois, açúcares, alcanos, alcenos, entre outros, tornando essa região

crucial rica nos grupamentos que compõe as principais classes de metabólitos da espinheira-santa, os flavonoides e taninos. Em  $2900\text{ cm}^{-1}$  a  $2800\text{ cm}^{-1}$  repete-se os estiramentos assimétricos C–H dos grupamentos  $-\text{CH}_2$ , simétricos C–H dos grupamentos  $-\text{CH}_3$ , grupamentos CHO e alifáticos ( $-\text{C}-\text{H}$ ,  $-\text{C}-\text{H}_2$  e  $-\text{C}-\text{H}_3$ ). Já a região de  $3500\text{ cm}^{-1}$  está bastante relacionada com a presença de ligações O–H provenientes da presença de água absorvida na amostra, álcool e grupamentos fenólicos (Alves *et al.*, 2024; Martins, 2020; Shurvell, 2006).

O perfil espectroscópico de cada uma das formas de apresentação pode ser comparado a partir de sua sobreposição (**Figura 19**).

Figura 19 – Sobreposição do perfil espectral das formas de apresentação estudadas para *M. ilicifolia*



Fonte: O autor (2023)

Comparando os espectros originados de cada uma das apresentações estudadas, nota-se uma ampliação na transmitância ao passo que mais processada se encontra a amostra. O perfil espectral dos extratos permitiu uma análise mais precisa de quais são os estiramentos relevantes encontrados em *M. ilicifolia*, destacando-se a porção média e final do espectro, além de contar com o aparecimento de uma elevada banda na região de  $3000$  a  $3500\text{ cm}^{-1}$ , oriunda do processo de maceração.

A análise dos espectros de cada uma das apresentações da espinheira-santa é crucial, pois esses são os objetos de estudo para a criação dos modelos. A quimiometria partiu do PCA, visando a identificação da espinheira-santa dentre todas as espécies vegetais contaminantes testadas, para então seguir ao PLS-DA, um modelo discriminatório e facilitado pela construção prévia do PCA.

#### 5.4.1 Modelo PCA

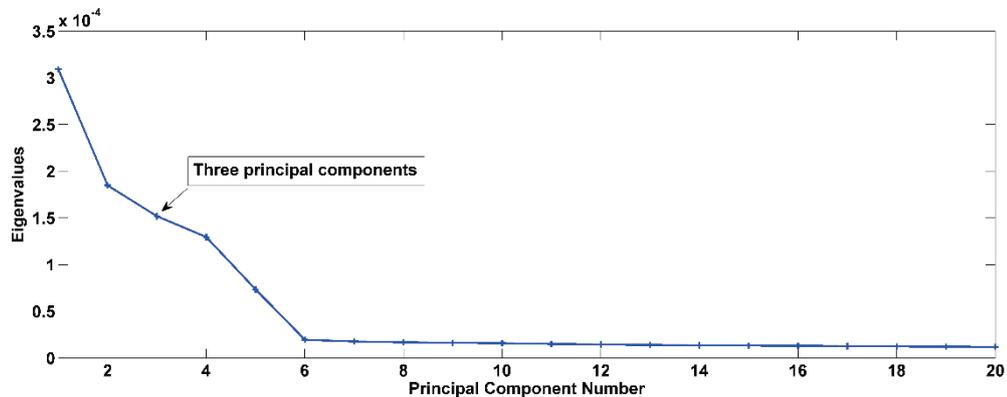
##### 5.4.1.1 Folhas

A etapa de criação dos modelos de PCA teve início com as amostras na forma de folha. Primeiramente, os modelos foram tratados centrando os dados na média (*mean center*) e autoescalamento (*autoscale*), mas os resultados não foram satisfatórios devido ao afunilamento das amostras, impedindo a separação das espécies.

Em seguida, aplicou-se um pré-processamento mais específico e que apresentou real capacidade corretiva. A aplicação do GLS W visa melhorar a precisão das análises estatísticas, corrigindo variações sistemáticas e estabelecendo correlações entre os dados que compõem a matriz (SILVA, *et al.* 2018). Para a construção do atual modelo, essa metodologia se mostrou adequada quando aplicados valores de alfa abaixo de 0,01. O melhor modelo de classificação foi alcançado utilizando GLS W com alfa de 0,0075.

O número de componentes principais para este pré-processamento foi escolhido baseado na % de variância cumulativa. O gráfico expresso na **Figura 20** demonstra que 3 componentes principais são responsáveis pela maior variância dos dados, totalizando **27,88%** da variância cumulativa, corroborando com a complexidade das folhas.

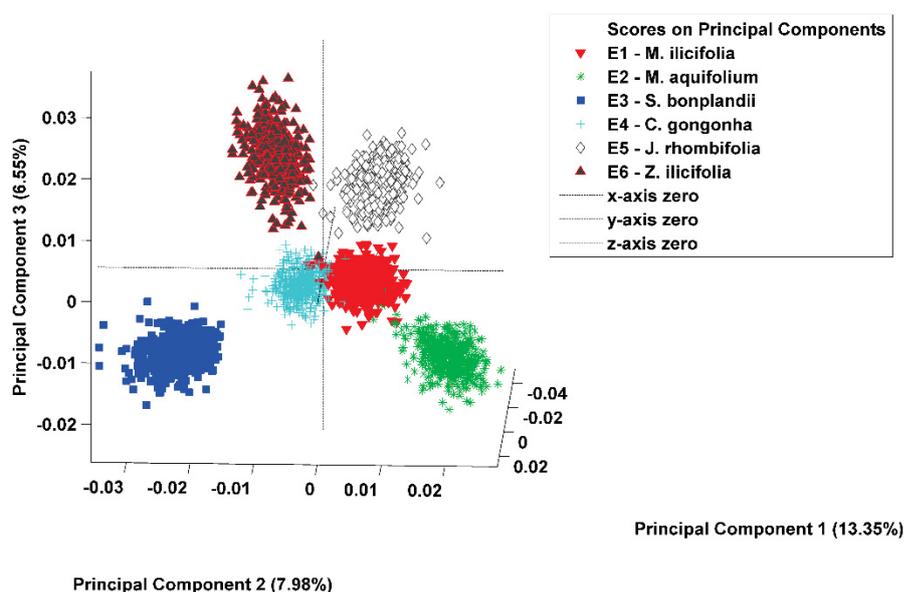
Figura 20 - Análise gráfica da escolha do número de componentes principais



Fonte: O autor (2023)

A escolha do número de componentes principais se dá pela relação entre o total de variância que pode ser explicada por cada componente principal (*Eigenvalues*, eixo Y) e o número de componentes principais propriamente dito (eixo X). Para gerar um modelo de qualidade, é importante otimizar esses parâmetros, buscando uma boa explicação dos dados da matriz sem introduzir viés. Quando o número de componentes principais é excessivo, o modelo perde credibilidade devido à utilização de informações redundantes, como o ruído do equipamento em que se realizou as análises. Por outro lado, um número reduzido de componentes principais oferece uma janela restrita para explicar as informações, limitando a divisão da variância do modelo entre as componentes principais e, por consequência, reduz sua capacidade de identificação (Ferreira, 2015). Essas componentes principais carregam consigo informações cruciais para a interpretação do modelo e classificação das amostras. Após aplicação das três componentes principais, foi possível estabelecer uma representação tridimensional deste modelo, a qual pode ser encontrada na **Figura 21**.

Figura 21 - Representação tridimensional do modelo PCA desenvolvido para folhas íntegras



Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

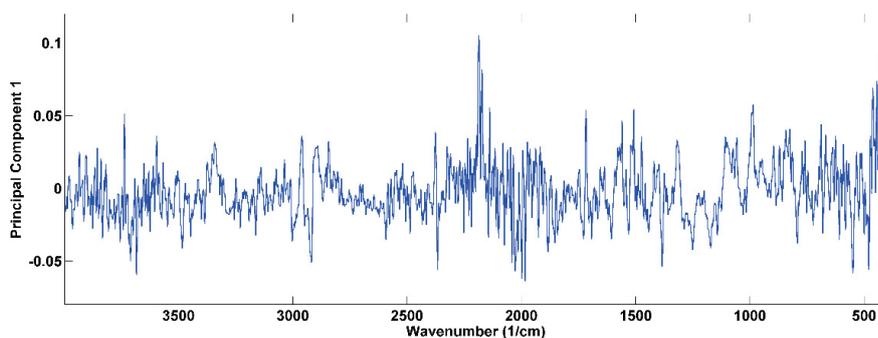
A representação tridimensional aponta que todas as espécies foram classificadas de forma significativamente distinta. Além disso, notou-se que alguns espectros da espécie *M. aquifolium* foram identificados na mesma região que a espinheira-santa oficial, com alguns pontos sobrepostos. Vale ressaltar que este contaminante se trata do adulterante mais comum da espinheira-santa, sendo frequentemente encontrado em amostras comerciais. Por sua vez, a amostra *Citronella gongonha* foi a que mais se aproximou da espécie oficial; com uma quantidade mínima de pontos sobrepostos com a espinheira-santa, o que também indica uma diferenciação efetiva de ambas as espécies,

As amostras das espécies *Citronella gongonha*, *Sorocea bonplandii*, *Jodina rhombifolia* e *Zollernia ilicifolia* apresentaram-se identificadas de forma totalmente distinta das demais, demonstrando diferença significativa em suas estruturas quando analisadas pelo método de infravermelho. A epicatequina, flavonoide base para métodos de referência de identificação da espinheira-santa, é um metabólito encontrado em quantidades reduzidas nessa espécie, fator que torna esse

componente bastante relevante para análises discriminatórias. Vale ressaltar que a diferença entre os componentes das amostras não se limita apenas a este metabólito, mas a um complexo conjunto químico, sendo a epicatequina a protagonista por atual atribuição como referência de *M. ilicifolia*. Por sua vez, a espécie *M. aquifolium* se apresenta similar à *M. ilicifolia* quanto a epicatequina, mas também foi discriminada pelo modelo, reforçando a importância do perfil fitoquímico completo para distinção das espécies. Análises fitoquímicas geram dados complexos que englobam diversas substâncias e suas interações, o que torna vantajosa a aplicação de um modelo multivariado, uma vez que esse torna possível a análise simultânea de múltiplas variáveis, possibilitando a extração de uma grande quantidade de informações da matriz de dados, como neste caso, que permitiu a identificação das espécies vegetais (Ferreira, 2015).

O gráfico *Loading Plots* tem elevada importância para o estudo pois permite identificar quais as regiões dos espectros foram mais relevantes para correta classificação de cada espécie. A **Figura 22** demonstra que existe um grau de importância em praticamente toda a faixa do infravermelho médio, sendo necessária sua avaliação como um todo para a identificação das espécies.

Figura 22 – Gráfico variáveis/*Loading Plots* para as folhas íntegras



Fonte: O autor (2023)

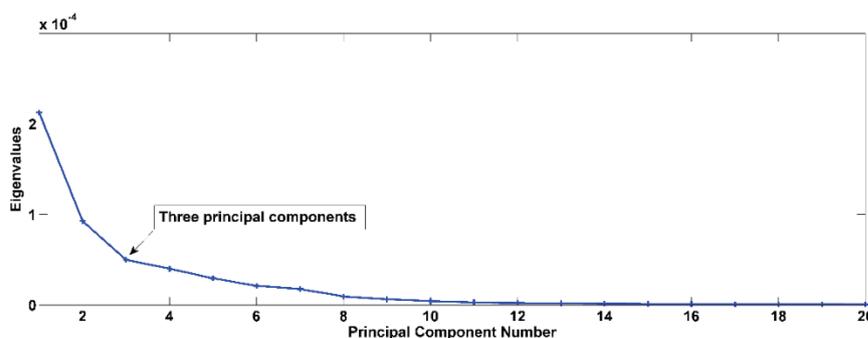
Amostras vegetais são complexas e apresentam diferenças macro e microscópicas entre si, fato evidenciado pelo panorama obtido no infravermelho das folhas íntegras. A importância de cada número de onda pode ser explicada por essa complexidade, sendo necessária a análise de todas as regiões espectrais para identificação de cada espécie.

#### 5.4.1.2 Pós

Em relação aos pós, os pré-tratamentos aplicados foram, simultaneamente, o *mean center*, a aplicação de alisamento com janela de 15 dados associado à primeira derivada e uma polinomial, e GLS W com alfa de 0,02. O *mean center* é um dos tratamentos mais convencionais aplicados a variáveis e tem o objetivo de centralizar os dados em um valor médio calculado para cada um dos eixos que compõem uma matriz de dados, sem alterar o tamanho e estrutura desses dados. A utilização dessa técnica permite melhorias a nível de colinearidade, interpretação e estabilização dos resultados (Ferreira, 2015).

Por sua vez, o alisamento dos dados e aplicação da primeira derivada são métodos frequentemente combinados. O *smoothing* (alisamento) melhora a razão sinal/ruído ao suavizar variações randômicas nos espectros, baseando-se em uma faixa ímpar de números de onda. É crucial aplicar essa técnica com cuidado a fim de não superalisar os dados, tornando-os inadequados para análise. Como parte das técnicas de correção da linha de base, as derivadas ajustam os espectros deslocando-os para o zero de absorbância, usando o método polinomial de primeira ordem (Ferreira, 2015). É importante ressaltar que todos os pré-tratamentos foram aplicados na sequência mencionada, pois a aplicação em ordens diferentes acaba alterando os sucessivos cálculos e modificações da matriz de dados, impactando no resultado final do modelo. A seleção de componentes principais se deu da mesma forma que nas folhas (**Figura 23**), onde 3 componentes principais foram escolhidas por representar **69,35%** de variância cumulativa dos dados.

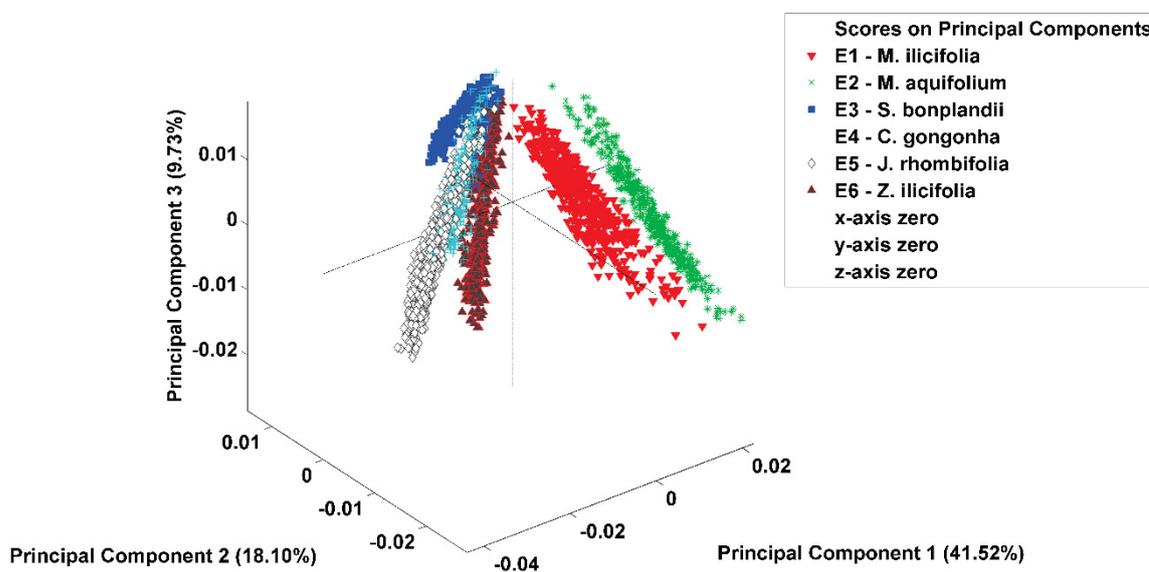
Figura 23 - Análise gráfica para a escolha do número de componentes principais para o modelo PCA dos pós



Fonte: O autor (2023)

A representação tridimensional promovida pelas três componentes principais selecionadas pode ser encontrada na **Figura 24** a seguir.

Figura 24 - Representação tridimensional do modelo PCA desenvolvido para os pós



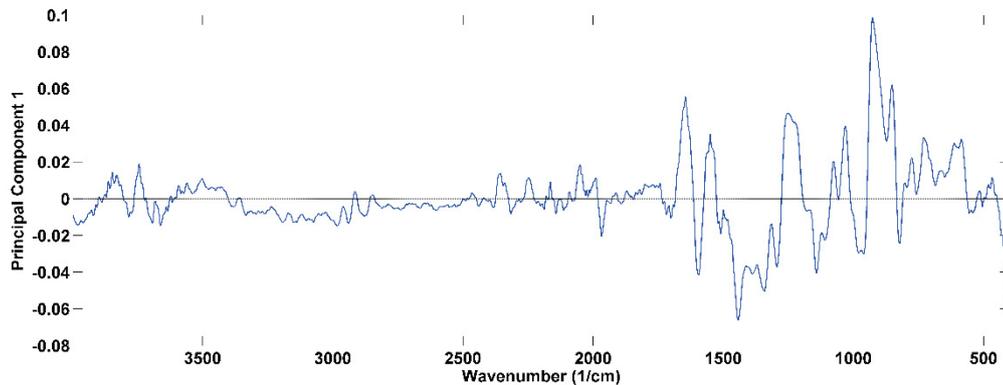
Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

A representação gráfica dos pós, dentre todas as formas de apresentação das plantas, foi a que demonstrou menor uniformidade nos seus espectros, com resultados distribuídos em uma ampla faixa. A espécie E1 foi efetivamente separada das demais, com apenas alguns dados apresentando contato com a espécie E2, considerada o principal contaminante.

O gráfico *Loading Plots* para os modelos dos pós mostrou-se mais seletivo do que o apresentado para o modelo das folhas. Notou-se que os números de onda de maior relevância estão na faixa de  $1750\text{ cm}^{-1}$  a  $500\text{ cm}^{-1}$ , com um impacto menor também presente no intervalo aproximado de  $3000\text{ cm}^{-1}$  a  $2800\text{ cm}^{-1}$ . As regiões entre  $1800$  e  $2800\text{ cm}^{-1}$ , além da faixa de  $3000\text{ cm}^{-1}$ , mostraram menor influência no modelo (**Figura 25**).

Figura 25 - Gráfico variáveis/*Loading Plots* para os pós

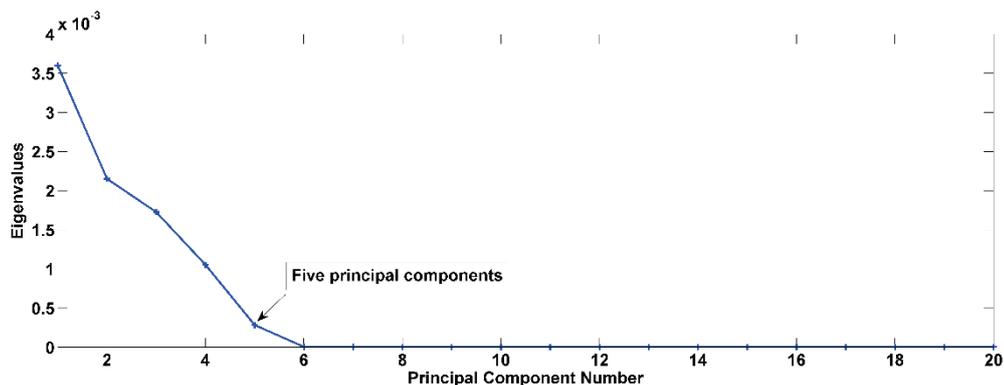


Fonte: O autor (2023)

#### 5.4.1.3 Extratos

Os extratos foram submetidos à combinação dos pré-tratamentos *mean center*, alisamento com janela de 15 pontos e GLS W com alfa de 0,005. Diferentemente dos modelos anteriores, o número de componentes principais empregado foi de cinco, o que permitiu ao modelo cobrir com precisão **96,36%** da matriz dos dados (**Figura 26**).

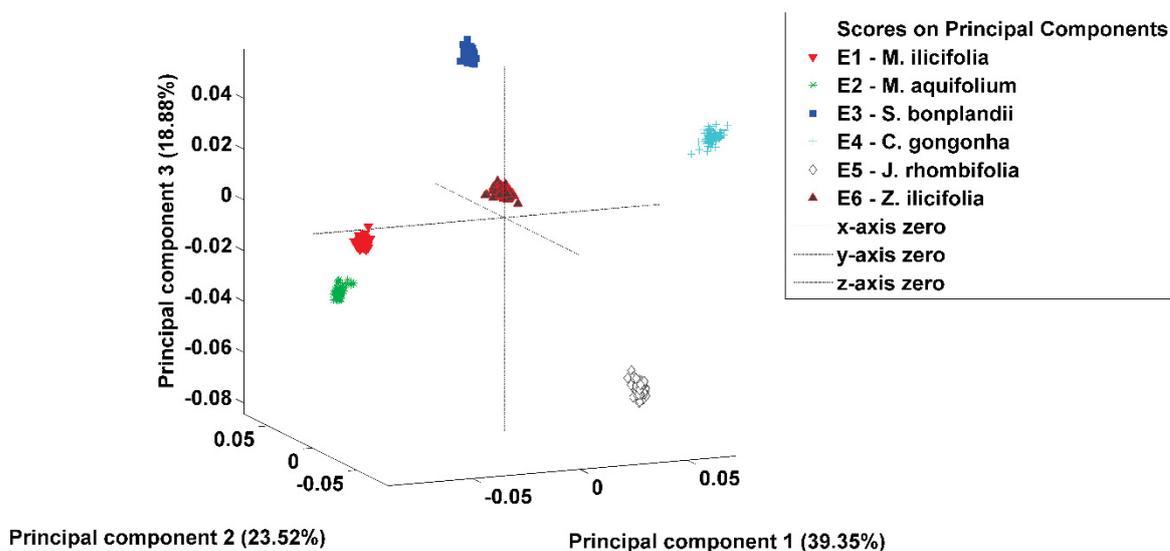
Figura 26 - Análise gráfica para escolha do número de componentes principais para o modelo PCA dos extratos



Fonte: O autor (2023)

A análise tridimensional dos dados permitiu verificar diferenças na classificação de cada uma das espécies (**Figura 27**).

Figura 27 - Representação tridimensional do modelo PCA para os extratos



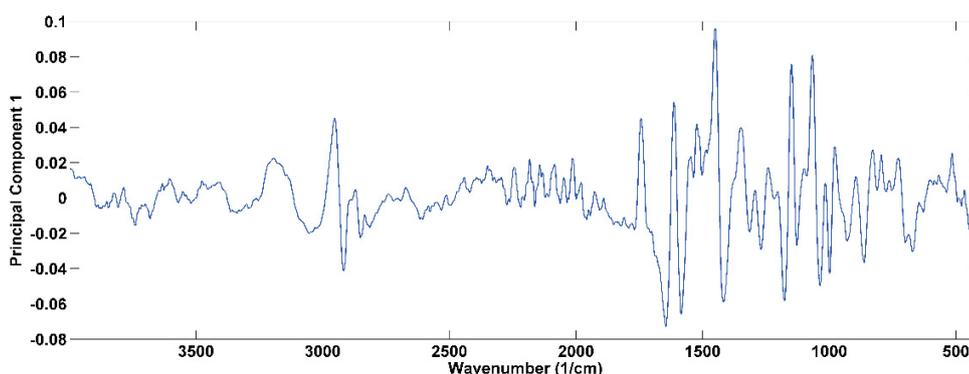
Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

Neste modelo, todas as espécies apresentaram variações intra-amostrais baixas, evidenciando uma maior uniformidade dos dados. Cada espécie foi classificada de forma isolada, com uma maior proximidade entre a espinheira-santa oficial e a espécie E2, mantendo o mesmo padrão observado para os pós e para as folhas íntegras.

O gráfico *Loading Plots* para os extratos sugere a presença de diversas regiões importantes para sua definição ao longo dos números de onda que compõem o infravermelho médio. Destaca-se a faixa dos  $1750\text{ cm}^{-1}$  até  $750\text{ cm}^{-1}$  e a região próxima aos  $3000\text{ cm}^{-1}$ , com outras faixas do espectro também demonstrando importância, porém em menor magnitude (**Figura 28**).

Figura 28 - Gráfico variáveis/*Loading plots* para os extratos



Fonte: O autor (2023)

Cada uma das formas de apresentação das folhas demonstrou um perfil diferente do gráfico *Loading plots*, indicando como o modelo previu a identificação das espécies a partir de comprimentos de onda distintos. Essa variação pode ter diversas origens, como a perda e surgimento de compostos relevantes durante os procedimentos de trituração, extração e secagem por força centrífuga, além de interferências provenientes do próprio equipamento para cada tipo de amostra, mesmo que a metodologia de análise tenha sido semelhante.

## 5.4.2 Modelo PLS-DA

### 5.4.2.1 Folhas

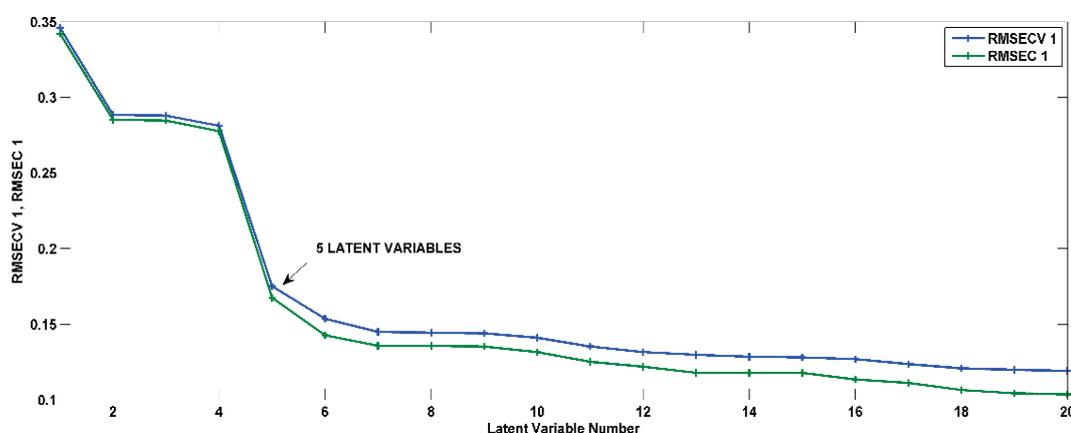
O modelo de discriminação da espinheira-santa de seus contaminantes também foi desenvolvido inicialmente com amostras das folhas. Para criar o modelo PLS-DA, é necessário dividir a matriz de dados em duas porções: uma para calibração, contendo 70% dos dados, e outra para validação, com os 30% restantes. Como o PLS-DA é complementar ao PCA, a matriz utilizada para criar ambos os modelos foi a mesma.

Dados de calibração têm a função subsidiar o desenvolvimento e treinar o modelo a partir de uma fração representativa da matriz de dados que se deseja usar para as predições/decisões. Por se tratar da criação do modelo, a parcela de dados para esta etapa deve ser superior àquela reservada para a avaliação da capacidade

preditiva do modelo gerado, a porção de validação. Este julgamento se baseia no conhecimento prévio de cada uma das classes e sua apresentação ao *software*, o qual não possui essa experiência, tentando designar esses dados às classes estipuladas no momento de criação do modelo (aplicação dos dados de calibração) (Ferreira, 2015; Peris-Díaz; Krezel, 2021; Westad; Marini, 2015).

Assim como no modelo PCA, é necessário escolher o número de componentes principais para o PLS-DA. Contudo, nessa categoria de experimentos, essas componentes são referidas como variáveis latentes. Essas variáveis têm a capacidade de relacionar a explicação da variância (bloco de dados X) com a propriedade de interesse (bloco de dados Y), maximizando a covariância entre ambos. Diferentemente das componentes principais no PCA, a seleção do número otimizado de variáveis latentes está relacionada com as medidas de erro de previsão nos modelos, representadas pelo *Root Mean Square Error of Cross-Validation* (RMSECV) e *Root Mean Square Error of Calibration* (RMSEC) (**Figura 29**). Ambas estas medidas de erro estão relacionadas com a variância explicada pelo bloco de dados X, a qual dita a construção do modelo. Essas métricas se encontram elevadas quando o número de variáveis latentes é reduzido, indicando um subajuste do modelo. Por outro lado, um número excessivo de variáveis latentes pode levar ao sobreajuste, ambos os fenômenos sendo prejudiciais para a predição do modelo. Assim, é necessário selecionar variáveis latentes que contenham informações suficientes para a criação de um modelo efetivo, evitando tanto o subajuste quanto sobreajuste (Ferreira, 2015).

Figura 29 – Gráfico de avaliação das variáveis latentes para amostras de folhas



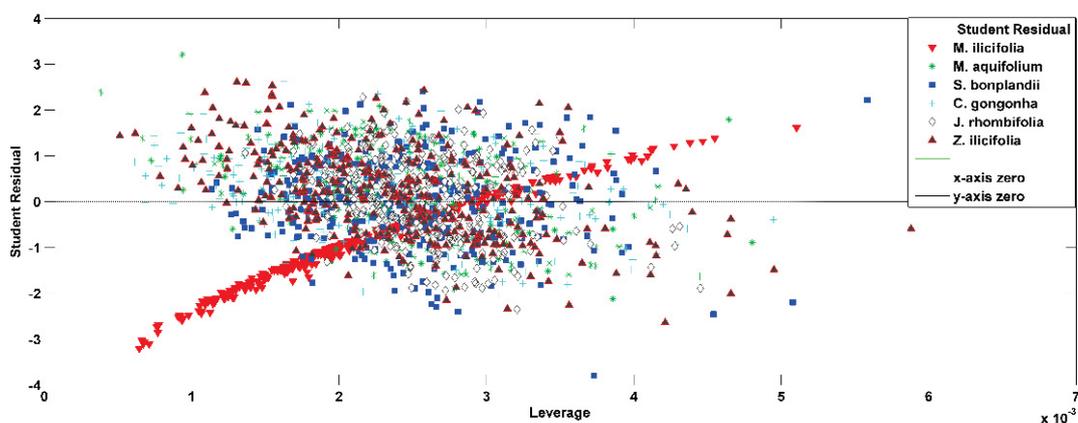
Fonte: O autor (2023)

Selecionou-se cinco variáveis latentes (**Figura 29**), as quais foram suficientes para maximizar a covariância sem sub ou sobreajustar o modelo. Essa escolha equilibra a complexidade do modelo e a precisão das previsões, garantindo um desempenho robusto. Os pré-tratamentos para criação deste modelo foram autoescalamento, GLS W com alfa de 0,07 e alisamento com janela de 9 pontos. Cerca de **82,5%** da variância pode ser capturada pelo modelo a partir desse número de variáveis latentes.

A partir da seleção de variáveis latentes é possível avaliar a presença de dados anômalos na matriz, também chamados de *outliers*. Resultados deste tipo devem ser excluídos, uma vez que não fornecem informações representativas da amostra estudada, podendo ter sido alterados por erros sistemáticos ou aleatórios. A inclusão desses dados na criação dos modelos pode distorcer as previsões devido a seus valores extremos (Santos, 2020).

A detecção e remoção de *outliers* para o modelo PLS-DA foi realizada utilizando o gráfico *Student Residual*, encontrado na **Figura 30**.

Figura 30 – Gráfico *Student Residual* para amostras na forma de folhas



Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

Este gráfico ilustra a relação entre duas medidas de predição: o *Student Residual* e o *Leverage*. O *Student Residual* representa a diferença padronizada entre os valores reais e os preditos para uma mesma observação, podendo ser quantificado

a partir da razão entre o resíduo comum e seu desvio padrão estimado. Um valor de *Student Residual* superior a 2,5, positivo ou negativo, é geralmente considerado como indicador de *outlier*. Por outro lado, o *Leverage* mede o potencial de influência que uma observação específica exerce sobre o modelo de regressão. Para esse parâmetro calcula-se o *Leverage* crítico para o modelo de calibração a partir da seguinte equação: (ASTM, 2017; PENN STATE, 2018; Greenwood, 2023).

$$\text{Leverage crítico} = VLs \times 3/n$$

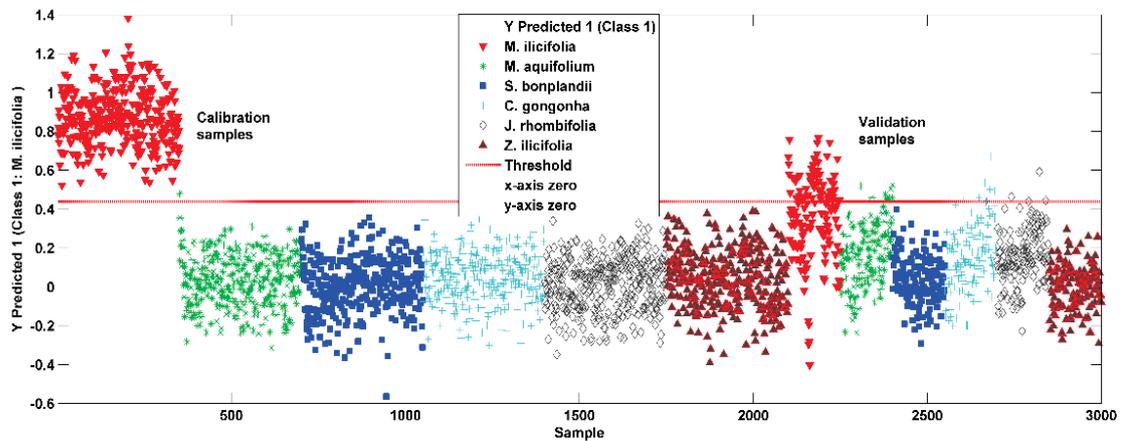
Legenda: VLs – número de variáveis latentes definidas para o modelo; n – número de amostras do conjunto de calibração.

Amostras que apresentarem *Leverage* superior ao valor crítico devem ser investigadas com maior cautela, pois representam observações de grande influência na matriz de dados, podendo se tratar de *outliers*. Os dados apresentaram *Student Residual* entre -2 e 2, e *Leverage* crítico de 0,007, valor não alcançado por nenhuma observação. Optou-se por não remover nenhum dado da matriz, uma vez que não houveram resultados com grande impacto sobre a média do modelo de regressão (ASTM, 2017).

A capacidade discriminatória do modelo PLS-DA pôde ser avaliada de forma individual para cada uma das classes de amostras por meio do gráfico Y predito, o qual ilustra a precisão das predições para cada classe de amostras. Na **Figura 31** pode-se observar como o modelo distingue as diferentes classes com base nos dados preditos.

O gráfico Y predito permite identificar se o modelo está adequadamente capturando as variações entre as classes e se faz predições precisas para cada uma delas. Esta etapa é crucial para garantir que o modelo seja útil em aplicações práticas, como a discriminação entre a espinheira-santa e seus contaminantes.

Figura 31 – Gráfico Y predito para avaliação da capacidade discriminatória do modelo para as folhas, com foco em *M. ilicifolia*



Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

Ao longo do eixo das abcissas, estão alocadas, sequencialmente, todas as varreduras realizadas, com cada espécie vegetal sendo representada por uma cor distinta. A parcela de varreduras destinada a calibração do modelo está à esquerda do gráfico, enquanto aquelas destinadas à validação se encontram na direita. Os dados de *M. ilicifolia* foram o foco deste gráfico, uma vez que esta é a espécie a ser discriminada das demais, representada pela cor vermelha. A linha horizontal que cruza o centro do gráfico, chamada de *threshold*, tem a função de delimitar as decisões do modelo. Amostras acima dessa linha são reconhecidas como o alvo do estudo, *M. ilicifolia*, enquanto aquelas que se encontram abaixo foram identificadas como espécies vegetais distintas.

Observou-se que, para os dados do modelo de calibração, todas as varreduras de *M. ilicifolia* foram corretamente identificadas. No entanto, para os dados de validação, diversas amostras dessa classe se apresentaram abaixo do *threshold*. Por outro lado, o número de amostras contaminantes identificadas como espinheira-santa foi zero. A partir do gráfico Y predito foi possível estabelecer a relação de espécies correta e incorretamente identificadas pelo modelo (**Quadro 6**).

Quadro 6 – Avaliação das amostras do gráfico Y predito do modelo PLS-DA para as folhas

	<b>Espinheira-santa</b>	<b>Demais espécies</b>
<b>Identificado como espinheira-santa</b>	Verdadeiro positivo (VP) = 415	Falso positivo (FP) = 20
<b>Identificado como contaminante</b>	Falso negativo (FN) = 85	Verdadeiro negativo (VN) = 2480

Fonte: O autor (2023)

A precisão da identificação do modelo permite o cálculo de métricas fundamentais para avaliar o desempenho do modelo, sendo elas a sensibilidade, especificidade e acurácia (**Quadro 7**).

Quadro 7 - Cálculo das métricas de qualidade do modelo para as folhas

<b>Sensibilidade</b>	$\frac{VP}{VP+FN}$	83%
<b>Especificidade</b>	$\frac{VN}{VN+FP}$	99,2%
<b>Acurácia</b>	$\frac{VP+VN}{TOTAL}$	96,5%

Fonte: O autor (2023)

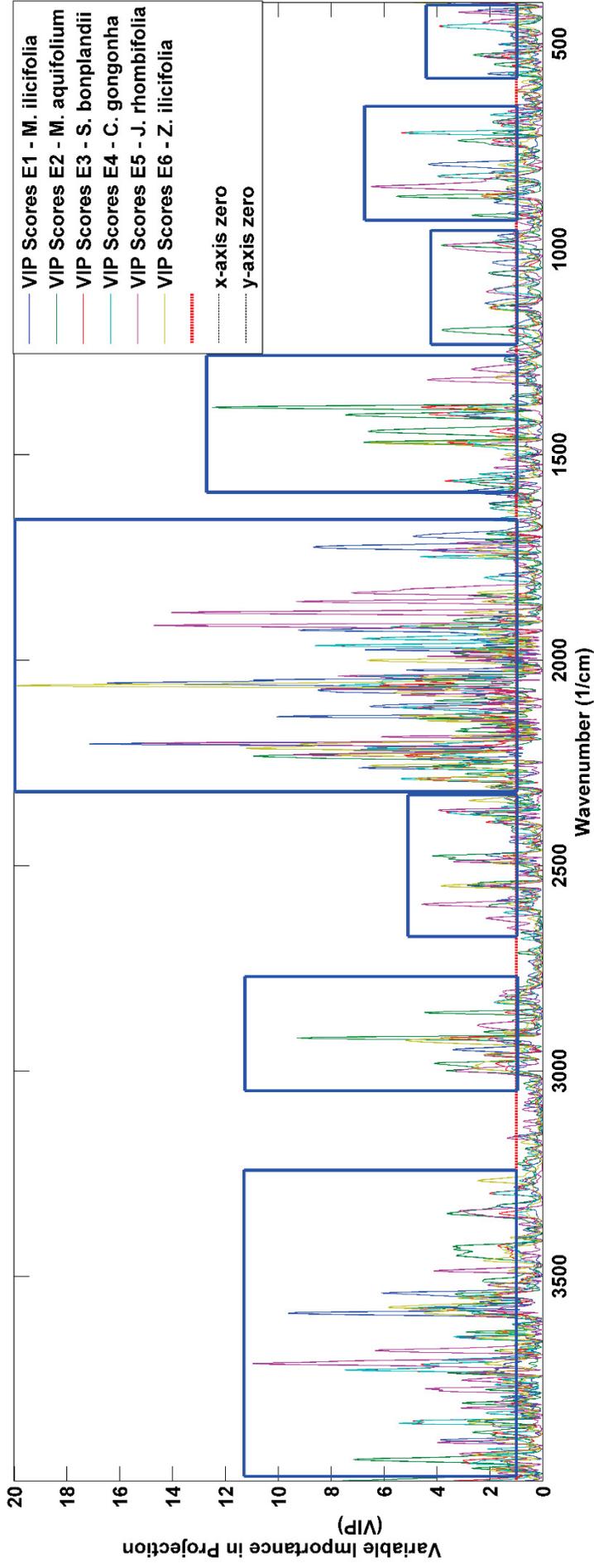
A sensibilidade indica a capacidade do modelo de identificar corretamente a *M. ilicifolia* (verdadeiros positivos), enquanto a especificidade se resume na capacidade de uma identificação acertada de amostras não pertencentes ao grupo alvo (verdadeiros negativos). Por sua vez, a acurácia consiste na medida geral de desempenho do modelo, considerando tanto os verdadeiros positivos quanto os negativos (Chu, 1999).

O modelo PLS-DA desenvolvido demonstrou uma excelente capacidade de detecção de amostras contaminantes dentro da matriz dados, apresentando um valor de especificidade bastante elevado. A acurácia do modelo também se apresentou bastante positiva, evidenciando um desempenho preditivo notável. A sensibilidade, que se refere à capacidade do modelo de predizer a classe de amostras de *M. ilicifolia*,

apresentou resultados positivos, embora reduzido em comparação às demais métricas. O modelo desenvolvido possui resultado comparável aos modelos de discriminação de Vieira et al. e Zhao et al., por exemplo, os quais também utilizaram a espectroscopia de infravermelho e quimiometria para distinguir, respectivamente, amostras de *Ilex paraguariensis* St. Hil e *Gentiana rigescens* Franch. ex Hemsl. Ambos os estudos culminaram em técnicas aptas a utilização em amostras comerciais, alcançando sensibilidade, especificidade e acurácias entre 80 e 100% para modelos de predição, resultado também encontrado para *M. ilicifolia*. A aplicação de uma matriz de dados de validação mais ampla pode ser uma alternativa para aumentar os valores de todas as métricas, aprimorando ainda mais o modelo para a forma de folhas (Vieira et al., 2019; Zhao et al., 2015).

Além desses resultados, foi possível realizar uma análise detalhada das variáveis de maior impacto para a discriminação das espécies vegetais a partir do gráfico VIP (*Variable in Importance Prediction*). Em uma metodologia baseada na espectroscopia de infravermelho, as principais variáveis correspondem aos picos característicos de cada espécie vegetal ao longo do espectro de 4000 a 400  $\text{cm}^{-1}$ . Neste gráfico, a linha de *threshold*, em vermelho e paralela ao eixo das abcissas, delimita as regiões importantes para discriminação (picos acima dessa faixa) daquelas de menor relevância (picos abaixo deste limite) (**Figura 32**).

Figura 32 – Gráfico VIP para a identificação das variáveis de maior relevância para discriminação das espécies na forma de folhas

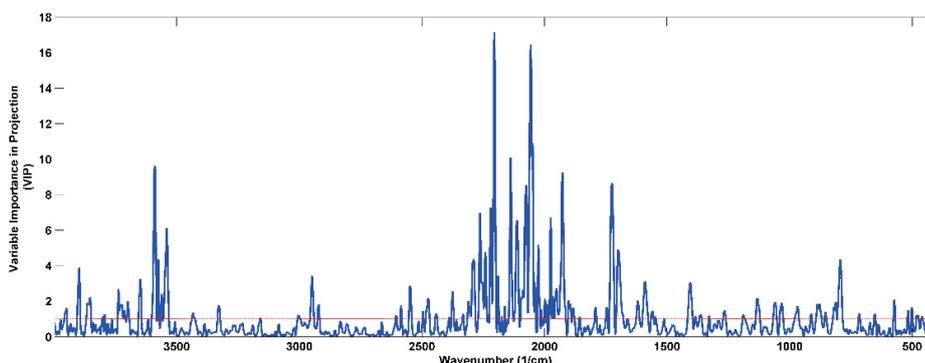


Legenda: Azul escuro – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Verde – *Monteverdia aquifolium* (Mart.) Biral; Vermelho – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Lilás – *Jodina rhombifolia* (Hook & Arn.) Reissek; Amarelo – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

As regiões demarcadas por retângulos azuis foram interpretadas como as de maior importância para discriminar as amostras. As diferentes cores dos picos representam as regiões relevantes para cada classe avaliada. Observa-se que, para as amostras na forma de folhas, a discriminação das espécies vegetais utilizou de diversos pontos do espectro, com destaque para a região entre 2300 a 1500  $\text{cm}^{-1}$ . No entanto, é necessária uma avaliação do perfil espectroscópico completo para a distinguir as espécies vegetais de forma eficiente. O perfil individual encontrado para a *M. ilicifolia* pode ser encontrado na **Figura 33**.

Figura 33 – Perfil VIP individual para folhas de *M. ilicifolia*



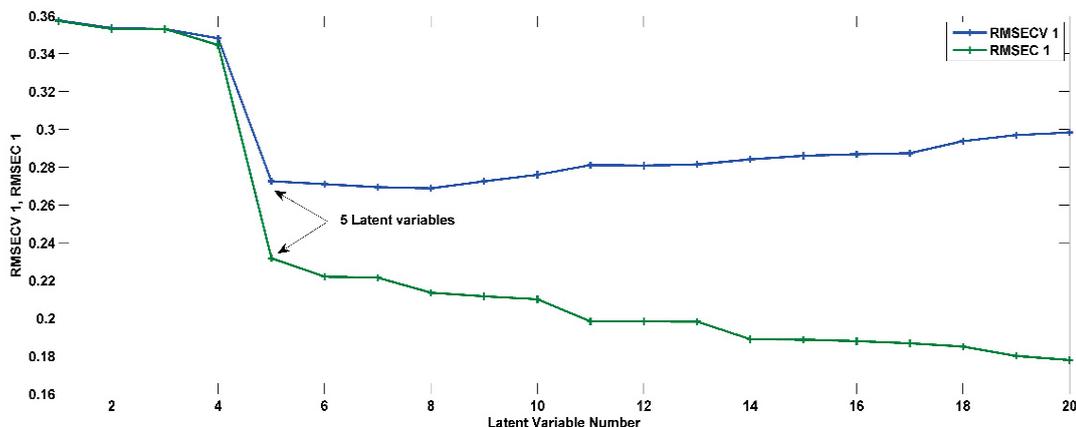
Fonte: O autor (2023)

O perfil individual encontrado para *M. ilicifolia* apresenta algumas diferenças em comparação ao gráfico VIP para todas as espécies vegetais combinadas. A região de maior impacto está situada, aproximadamente, no número de onda 3600  $\text{cm}^{-1}$  e a faixa de 2300 a 1700  $\text{cm}^{-1}$ .

#### 5.4.2.2 Pós

A criação do modelo PLS-DA dos pós seguiu uma metodologia similar ao das folhas, iniciando com a divisão dos dados em porções para calibração e validação. Os pré-tratamentos utilizados foram GLS W com alfa de 0,06 e alisamento com janela de 9 pontos. O número otimizado de variáveis latentes, baseado nas medidas de erro, para o modelo dos pós foi de cinco. (**Figura 34**).

Figura 34 – Gráfico de avaliação das variáveis latentes para amostras em pó



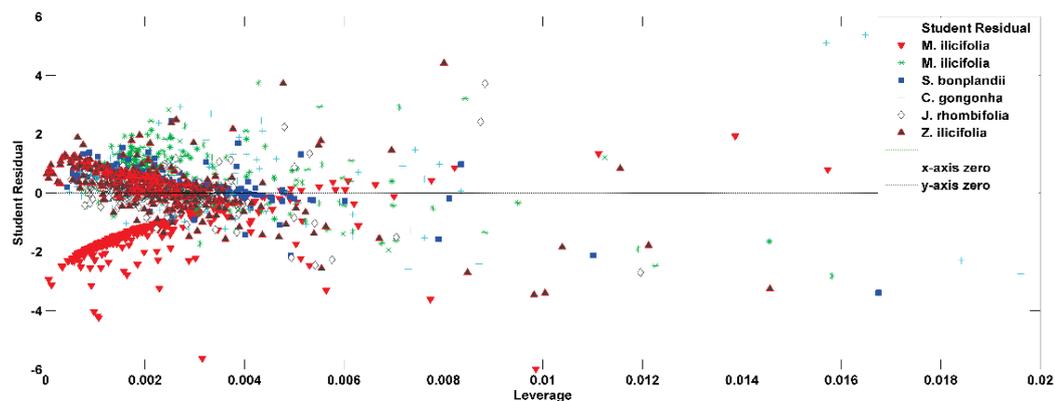
Fonte: O autor (2023)

Foi possível observar, no entanto, uma distinção entre os valores de RMSEC e RMSECV, com o primeiro sendo inferior ao RMSECV no nível de variáveis latentes selecionadas. A variância capturada pelo modelo a partir da aplicação desse número de variáveis latentes foi de **71,4%**.

O fenômeno de sobreajuste do modelo é evidenciado pelo aumento da medida de erro da predição devido à adição excessiva de informações, o que reduz sua efetividade. No presente caso, utilizar um número maior de variáveis latentes deixaria o modelo muito próximo do sobreajuste durante a etapa de validação cruzada. Portanto, cinco variáveis latentes foram consideradas o número otimizado para evitar esse fenômeno, ao mesmo tempo em que reduz o RMSEC.

Os dados destinados à criação e treinamento do modelo passaram por uma triagem usando gráfico *Student Residual* para identificar e excluir *outliers* da matriz (Figura 35).

Figura 35 – Gráfico *Student Residual* para identificação de outliers para as amostras na forma de pó



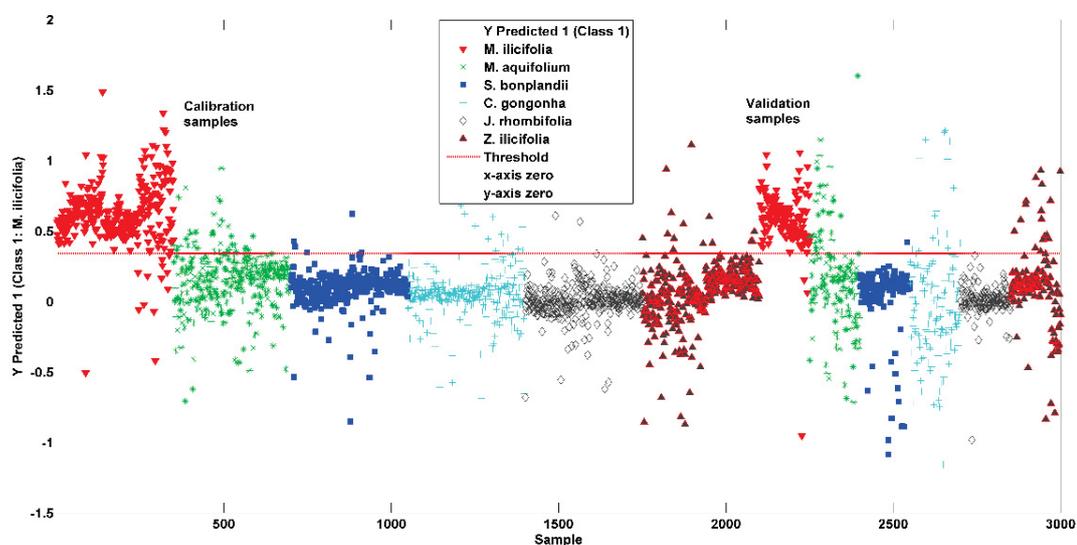
Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

Considerando os valores de *Student Residual* inferiores a -2,5 e superiores a 2,5, além de um *Leverage* crítico de 0,007, foi possível identificar resultados anômalos dentro da matriz de dados. Esses *outliers* foram devidamente retirados para evitar alterações indesejadas no modelo proposto. Os *outliers* não se limitaram a apenas uma classe, estando distribuídos pelas seis classes que compuseram a matriz de dados.

A capacidade preditiva do modelo PLS-DA para os pós foi analisada a partir do gráfico Y predito, focando na discriminação da *M. ilicifolia* (representada cor vermelha) das demais espécies contaminantes (representadas por outros grupos coloridos) (Figura 36).

Figura 36 – Gráfico Y predito com foco em *M. ilicifolia*, avaliando capacidade preditiva do modelo para amostras em pó



Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

As amostras foram alocadas ao longo do eixo X, com cada classe sendo representada por uma cor diferente. Assim como no modelo das folhas, as amostras acima da linha de *threshold* foram reconhecidas como *M. ilicifolia*, enquanto as abaixo da linha foram identificadas como contaminantes. Os resultados obtidos permitem o cálculo das métricas de sensibilidade, especificidade e acurácia (**Quadro 8**).

Quadro 8 – Avaliação das amostras do gráfico Y predito do modelo PLS-DA para os pós

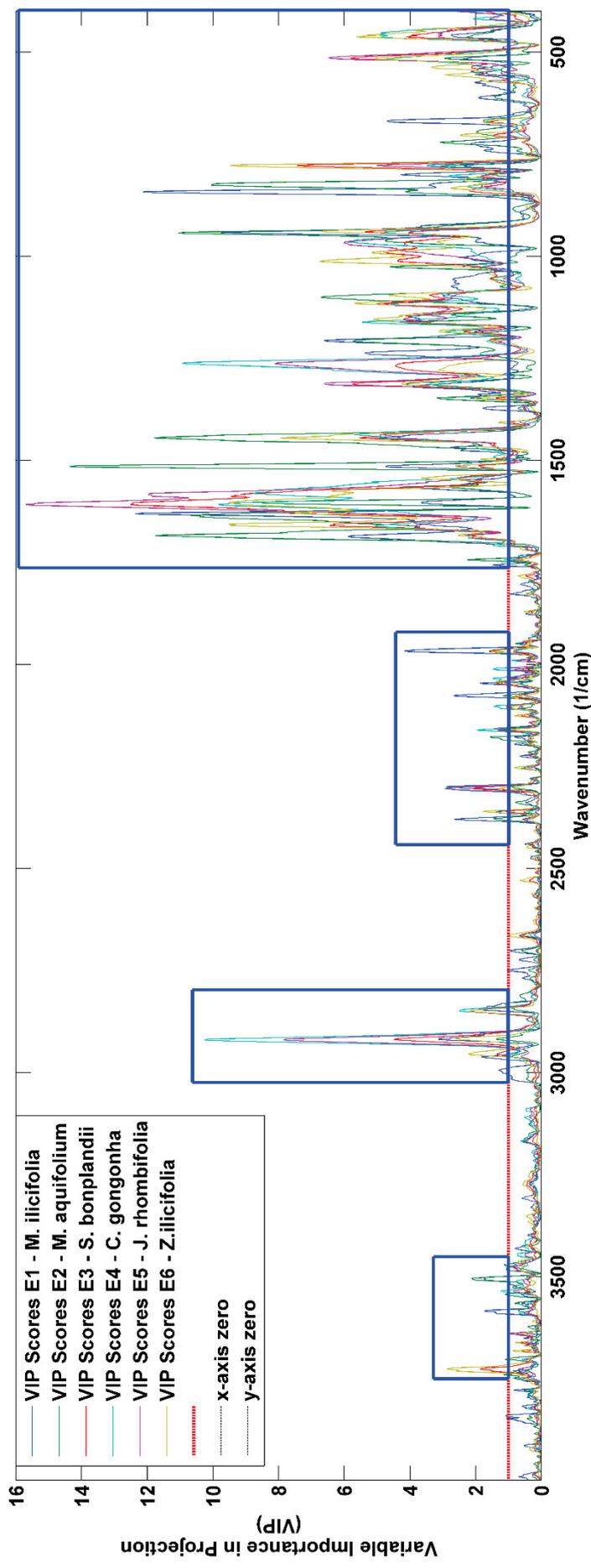
	<b>Espinha-santa</b>	<b>Demais espécies</b>
<b>Identificado como espinheira-santa</b>	Verdadeiro positivo (VP) = 484	Falso positivo (FP) = 127
<b>Identificado como contaminante</b>	Falso negativo (FN) = 16	Verdadeiro negativo (VN) = 2373

Fonte: O autor (2023)

Para o modelo dos pós, os valores de sensibilidade, especificidade e acurácia foram de, respectivamente, 96,8%, 94,9% e 95,2%. A partir da matriz de dados, o *software* construiu um modelo de elevada sensibilidade, especificidade e acurácia. Todas as amostras puderam ser corretamente discriminadas da *M. ilicifolia*. Em comparação ao modelo das folhas, o algoritmo dos pós apresentou uma sensibilidade superior (83% para folhas a 96,8% para os pós), enquanto teve uma especificidade e acurácia semelhantes (99,2% e 96,5% para as folhas, e 94,9% e 95,2% para os pós, respectivamente) demonstrando uma elevada capacidade de detectar a espinheira-santa e interpretar a amostra como tal.

O gráfico de VIP também demonstrou um perfil distinto para as variáveis de maior impacto nos pós com relação às folhas, como pode ser visto na **Figura 37**.

Figura 37 – Gráfico VIP para a identificação das variáveis de maior relevância para discriminação das espécies na forma de pós



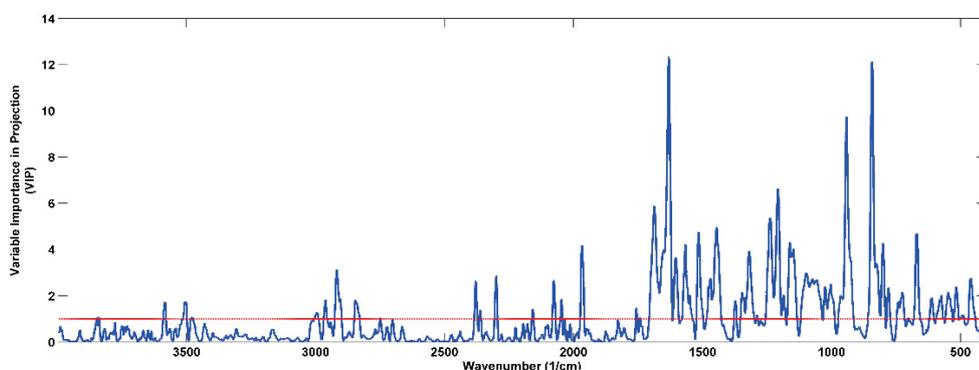
Legenda: Azul escuro – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Verde – *Monteverdia aquifolium* (Mart.) Biral.; Vermelho – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Lilás – *Jodina rhombifolia* (Hook & Arn.) Reissek; Amarelo – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

As regiões de maior impacto para a discriminação das amostras incluem, aproximadamente, 3750 a 3500  $\text{cm}^{-1}$ , 2900  $\text{cm}^{-1}$ , 2400 a 2000  $\text{cm}^{-1}$  e, como área de maior importância, 1800 a 400  $\text{cm}^{-1}$ . Este gráfico apresentou perfil distinto daquele das folhas, demonstrando como o *software* utiliza de diferentes regiões do espectro para distinguir as amostras de acordo com a forma apresentada.

Por sua vez, o gráfico VIP individual para *M. ilicifolia* se mostrou similar ao encontrado para todas as amostras, explicitando sua influência na representação geral das amostras simultaneamente (**Figura 38**).

Figura 38 – Perfil VIP individual para *M. ilicifolia* em pó



Fonte: O autor (2023)

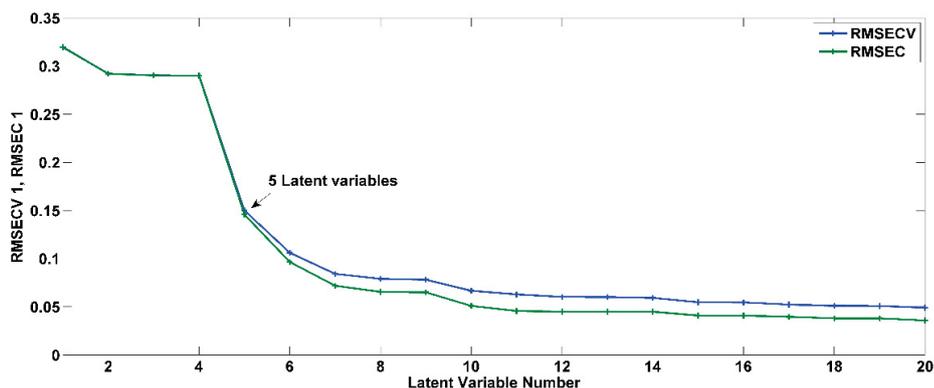
Para as amostras em pó, 500 varreduras por amostra também se mostraram suficientes para a criação de um modelo com alta sensibilidade, especificidade e acurácia. Foram identificados *outliers* em todas as classes de amostras, e o número de variáveis latentes, conciliando a otimização dos valores de RMSEC e RMSECV, foi determinado como cinco. O modelo gerado demonstrou potencial para aplicação comercial e implementação como método de referência para discriminação de amostras.

#### 5.4.2.3 Extratos

Complementando o estudo pela técnica quimiométrica, a metodologia de desenvolvimento do modelo PLS-DA para os extratos seguiu o mesmo procedimento aplicado para folhas e pós. Os pré-tratamentos aplicados foram o *mean center*, alisamento com janela de 9 pontos e GLS W com alfa de 0,04.

A seleção do número de variáveis latentes foi similar aos modelos PLS-DA anteriores, com um valor otimizado para ambas as métricas de erro (RMSEC e RMSECV) de cinco variáveis. O modelo foi capaz de capturar um total de **98%** da variância a partir desse número de variáveis latentes (**Figura 39**).

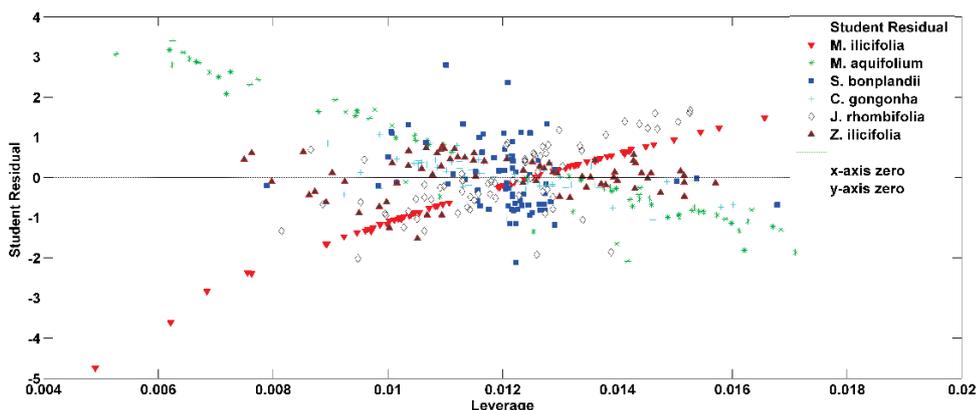
Figura 39 – Gráfico de avaliação das variáveis latentes para amostras em extratos



Fonte: O autor (2023)

A primeira etapa na criação do algoritmo envolveu a verificação de *outliers*, buscando sua presença e subsequente exclusão. (**Figura 40**).

Figura 40 - *Student Residual* para identificação de outliers para as amostras na forma de extratos



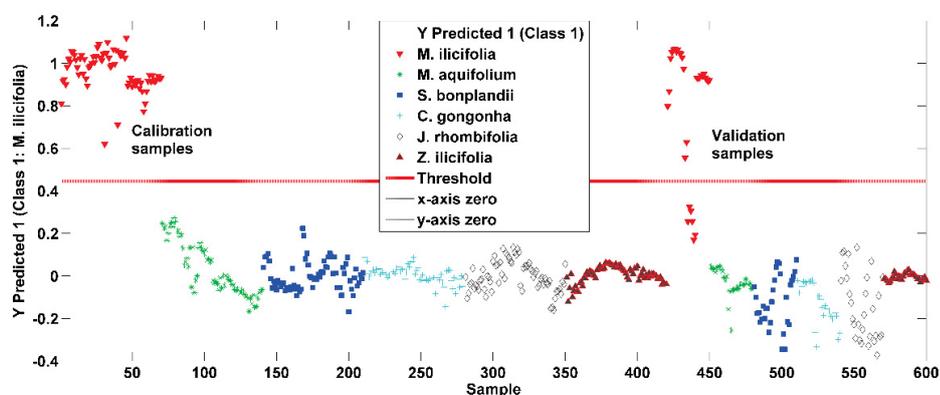
Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

A partir das referências adotadas anteriormente (*Student Residual* > 2,5 e *Leverage* crítico de 0,035), pode-se inferir que não existiram valores anômalos dentro da matriz de calibração. Portanto, para a fração destinada à calibração, nenhum resultado foi removido.

Avaliando-se a capacidade de discriminar a espécie vegetal alvo (*M. ilicifolia*, representada pela cor vermelha) das demais, o gráfico Y predito foi gerado (**Figura 41**).

Figura 41 – Gráfico Y predito com foco em *M. ilicifolia*, avaliando capacidade preditiva do modelo para amostras em extrato



Legenda: Triângulo vermelho – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Asterisco verde – *Monteverdia aquifolium* (Mart.) Biral.; Quadrado azul escuro – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Cruz azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Losango branco – *Jodina rhombifolia* (Hook & Arn.) Reissek; Triângulo preto – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

Todas as amostras contaminantes foram representadas por diferentes cores ao longo do eixo das abcissas. Observa-se a ausência dessas amostras contaminantes para regiões acima do *threshold* estabelecido, ilustrando a precisão do *software* em prever corretamente quais amostras não pertenciam à espécie-alvo do estudo. Por sua vez, as amostras da classe da espinheira-santa ficaram sobre a faixa limite, com exceção de seis dados da porção de validação, erroneamente identificados como contaminantes. Com base nessas características, foram realizados os seguintes cálculos (**Quadro 9**).

Quadro 9 – Avaliação das amostras do gráfico Y predito e o cálculo das métricas de qualidade do modelo PLS-DA de extratos

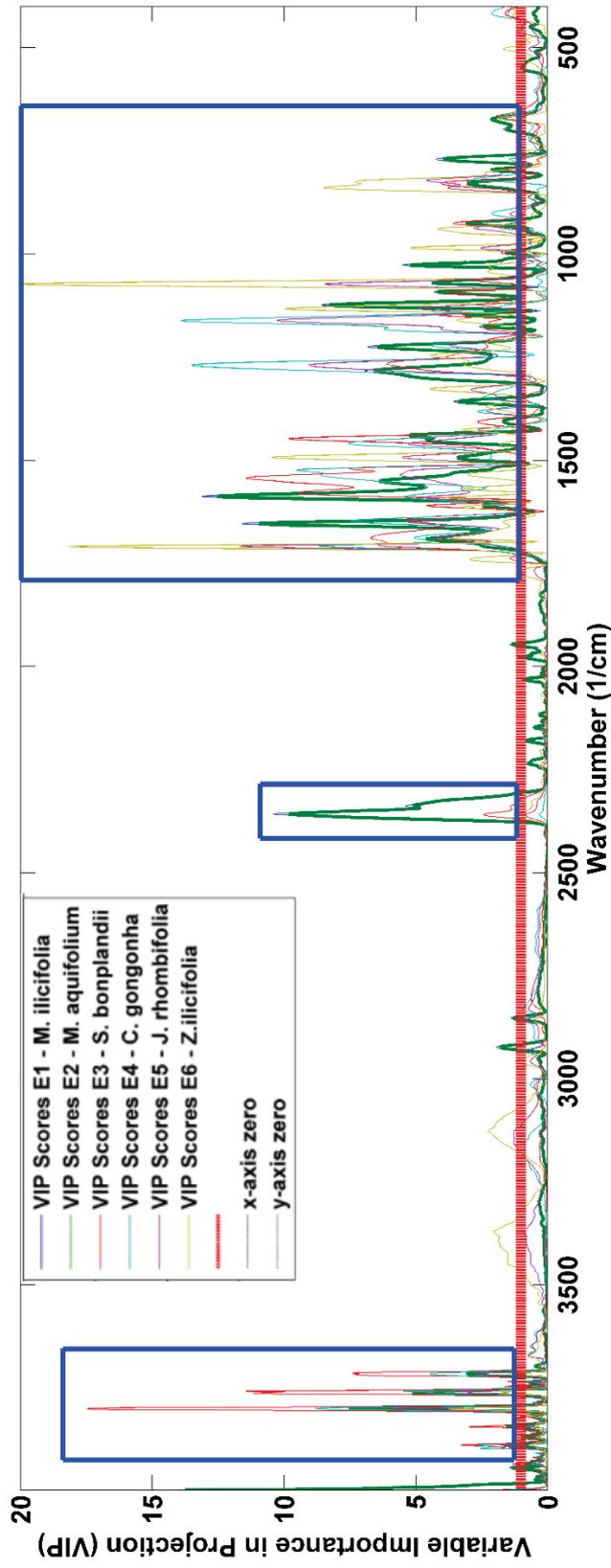
	<b>Espinheira-santa</b>	<b>Demais espécies</b>
<b>Identificado como espinheira-santa</b>	Verdadeiro positivo (VP) = 94	Falso positivo (FP) = 0
<b>Identificado como contaminante</b>	Falso negativo (FN) = 6	Verdadeiro negativo (VN) = 500

Fonte: O autor (2023)

Os resultados de sensibilidade, especificidade e acurácia encontrados para o modelo de extratos foi de 94%, 100% e 99%, respectivamente, demonstrando os melhores resultados entre todas as apresentações testadas. Devido à sua capacidade de discriminar corretamente todas as amostras contaminantes, encontrou-se uma especificidade ideal (100%) para o modelo proposto. As métricas de sensibilidade e acurácia também foram altamente satisfatórias, não atingindo um valor ideal devido apenas às seis amostras da espinheira-santa identificadas como contaminantes. A natureza das amostras nesta forma de apresentação, devido ao processo de extração que concentra os compostos presentes na espécie vegetal, permitiu o desenvolvimento de um modelo de qualidade elevada, sendo a forma mais adequada para o estudo de amostras comerciais pelo método quimiométrico.

Além das métricas de especificidade, sensibilidade e acurácia, o modelo para os extratos também revelou um perfil conciso no gráfico VIP. Para as amostras nesta forma de apresentação, as principais regiões de impacto mostraram-se mais focalizadas, como representado na **Figura 42**.

Figura 42 – Gráfico VIP para a identificação das variáveis de maior relevância para discriminação das espécies na forma de extratos

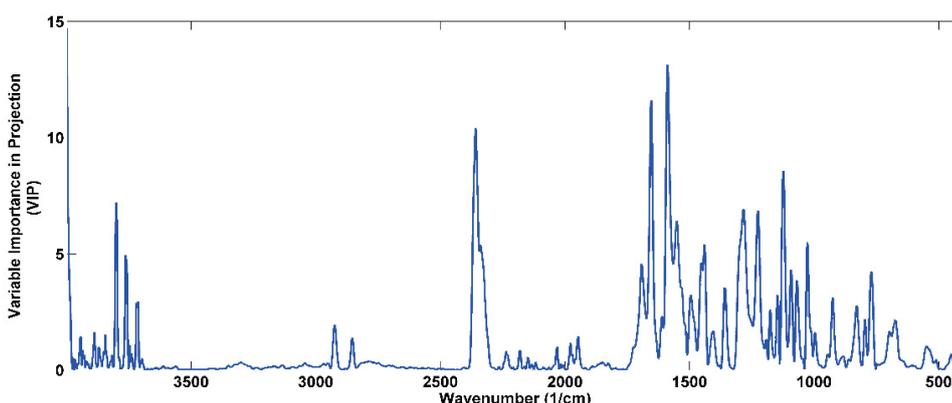


Legenda: Azul escuro – *Monteverdia ilicifolia* (Mart. Ex. Reissek) Biral; Verde – *Monteverdia aquifolium* (Mart.) Biral.; Vermelho – *Sorocea bonplandii* (Baill.) W.C. Burger et al.; Azul clara – *Citronella gongonha* (Mart.) R.A. Howard; Lilás – *Jodina rhombifolia* (Hook & Arn.) Reissek; Amarelo – *Zollernia ilicifolia* (Brongn.) Vogel.

Fonte: O autor (2023)

À luz do exposto, as regiões do espectro de maior impacto para discriminação das espécies vegetais estão localizadas nos números de onda aproximados de 3900 a 3700  $\text{cm}^{-1}$ , 2350  $\text{cm}^{-1}$  e a faixa de 1750 a 600  $\text{cm}^{-1}$ . As regiões mais importantes do espectro isolado de *M. ilicifolia* mostraram-se semelhantes àquelas encontradas para o conjunto das espécies, e pode ser encontrado na **Figura 43**.

Figura 43 – Perfil VIP individual para *M. ilicifolia* em extratos



Fonte: O autor (2023)

Ao comparar os gráficos VIP para todas as formas de apresentação, é possível identificar diferenças significativas nas regiões utilizadas pelo modelo para discriminar as amostras em cada uma delas, principalmente entre as folhas e demais apresentações. A quantidade de números de onda relevantes para definição dos algoritmos foi proporcional à complexidade das amostras. As estruturas foliares em sua forma íntegra apresentaram uma ampla variedade de regiões relevantes. No entanto, ao longo de seus processamentos e concentração da amostra, o número de regiões foi reduzido. Essa característica destaca a importância de desenvolver modelos para cada uma das formas de apresentação e a necessidade de planejar para extrair o máximo de informação possível de cada uma delas.

O modelo PLS-DA para os extratos apresentou elevados valores de sensibilidade e acurácia, juntamente com uma medida ideal de especificidade. O número de variáveis latentes foi o mesmo para as três formas de apresentação. Nenhum *outlier* foi encontrado nos dados de calibração. A quantidade de varreduras suficiente para a criação do modelo foi de 100, sendo que análises acima deste número não contribuíram para a qualidade do modelo.

### 5.4.3 *Machine Learning*

A quimiometria é uma ferramenta difundida no ramo de vigilância em saúde por conta de sua capacidade exploratória e discriminativa. No entanto, essa metodologia apresenta limitações para alguns tipos de amostra, como espécies vegetais, as quais demonstram uma elevada complexidade quanto à sua composição, sendo necessário complementar suas análises com outros métodos mais flexíveis e robustos, como o *machine learning*.

O *machine learning* é o ramo da ciência que busca fazer previsões e tomar decisões com base em informações fornecidas por matrizes de dados. Esse objetivo é alcançado através da construção de algoritmos e modelos estatísticos que aprendem e melhoram sua performance conforme mais dados são fornecidos ao modelo. A característica marcante dessa metodologia é sua capacidade de evoluir continuamente à medida que novas informações sobre a amostra são incorporadas ao algoritmo durante o treinamento. Ao aplicar diversas instâncias de treino, na forma de diferentes matrizes de dados, o modelo se aprimora, dando ao *machine learning* uma capacidade de utilização a longo prazo. À medida que novas amostras são analisadas e adicionadas ao modelo, a precisão preditiva também se eleva. Essa adaptabilidade não apenas aprimora o desempenho do modelo, mas também assegura que ele se mantenha relevante e eficaz em um ambiente de constante mudança (Janiesch; Zschech; Heinrich, 2020).

Embora os resultados pelo método quimiométrico tenham sido bastante positivos para as três formas de apresentação estudadas (folhas, pós e extratos), o *machine learning* pode oferecer uma capacidade preditiva ainda maior e/ou complementar os modelos previamente construídos, elevando ainda mais sua qualidade (Sarker, 2021; Shinde; Shah, 2018).

O modelo que se destacou foi o *Random Forest*, uma metodologia de classificação que utiliza a decisão conjunta de múltiplos classificadores na forma de árvores de decisão. Essas estruturas são construídas de forma aleatória a partir dos dados da matriz, o que amplia a capacidade preditiva do modelo ao reduzir a correlação entre as árvores. O *Random Forest* é uma das mais convencionais metodologias para o desenvolvimento de modelos por aprendizado de máquina por conta de sua eficiência, robustez, aplicabilidade em diversas matrizes de dados,

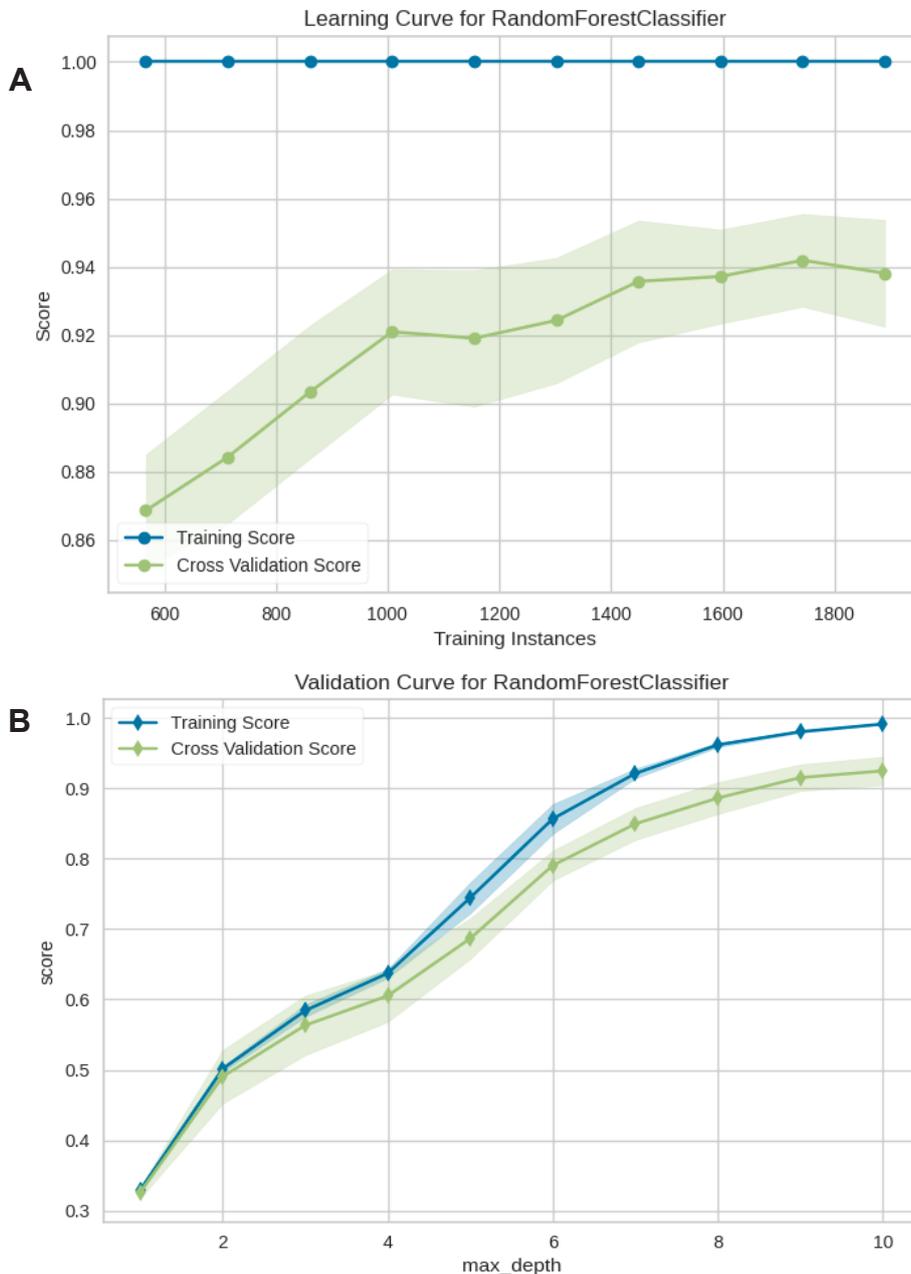
capacidade de lidar com sobreajuste e habilidade para identificar e tratar *outliers* (Ali; Khan; Ahmad, 2012; Breiman, 2001).

Após a seleção e validação do modelo, foram extraídos os resultados do desempenho preditivo, incluindo métricas como acurácia, precisão, coeficiente de correlação de Matthews (MCC), *recall*, índice *Kappa* e escore F1. Essas métricas permitem avaliar o quão bem o modelo faz previsões precisas sobre as diferentes espécies vegetais com base nos dados disponíveis. Além destas, são gerados os gráficos: área sob a curva ROC, *classification report*, *class prediction error*, *validation curve*, *learning curve*, *confusion matrix*, todos destinados a entender o desempenho preditivo do melhor modelo. Por fim, os gráficos das variáveis, neste caso os números de onda, mais importantes foram gerados a partir dos modelos otimizados, demonstrando quais grupos funcionais mais contribuíram para a diferenciação das espécies vegetais

#### 5.4.3.1 Folhas

O desenvolvimento dos modelos de *machine learning* começou com as amostras na forma de folhas. Primeiramente, foi avaliada a acurácia do modelo para calibração e validação (**Figura 44**).

Figura 44 – Progressão da acurácia da calibração e validação do modelo *machine learning* para amostras na forma de folha



Legenda: A - Progressão da acurácia da calibração do modelo; B - Progressão da acurácia da validação do modelo

Fonte: O autor (2024)

Ao longo das 10 instâncias de treino do modelo (eixo X), percebe-se que a acurácia (eixo Y) aumenta de forma gradual, estabilizando ao final da aplicação de todos os dados. Observa-se uma área de coloração verde clara a qual engloba todos

os resultados de acurácia encontrados até aquele momento pelas árvores de decisão formadas. A acurácia geral do modelo, representada pelos círculos verdes ao longo do gráfico, é a média de todos esses valores. Inicialmente, a acurácia do modelo foi de 87%, evoluindo para 94% após a aplicação de todas as instâncias de treino.

Para os dados de validação cruzada, a mesma análise foi realizada, também empregando 10 instâncias de cálculo da metodologia *10 fold*. Inicialmente, a acurácia para esse conjunto de dados foi de 33%, progredindo significativamente até alcançar uma média de 93%. Acompanhando a faixa de crescimento em ambos os gráficos, encontra-se uma linha chamada de *training score*, a qual apresenta a progressão dos resultados em situações ideais, onde a acurácia final seria de 100%.

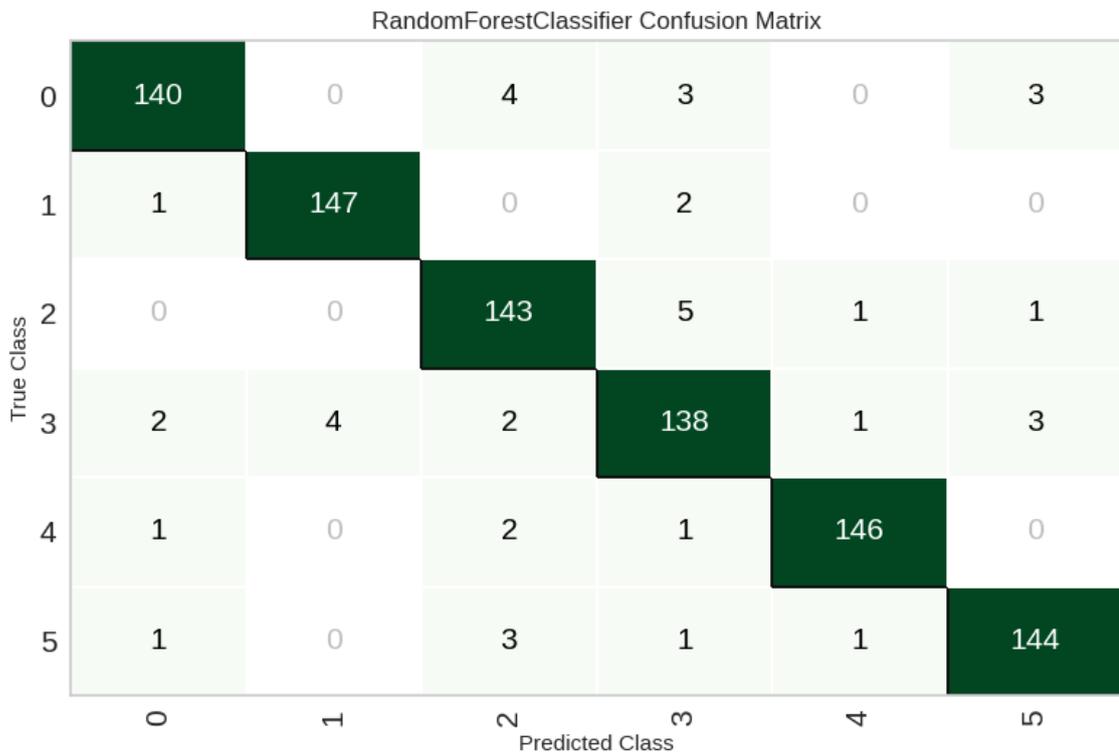
Tanto para os dados de calibração quanto para a validação cruzada, a acurácia encontrada foi elevada e similar. A combinação desses resultados indica uma acurácia geral do modelo extremamente positiva para os dados apresentados, demonstrando uma capacidade preditiva de alto desempenho para este tipo de amostra.

O estudo com aprendizado de máquinas prosseguiu com a construção da matriz de confusão dos dados, visando identificar a relação entre resultados verdadeiramente positivos e falsos positivos. Para as folhas, as classes envolvidas no estudo foram:

- *Classe 0 – Citronella gongonha;*
- *Classe 1 – Jodina rhombifolia;*
- *Classe 2 – Monteverdia aquifolium.;*
- *Classe 3 – Monteverdia ilicifolia (espinheira-santa);*
- *Classe 4 – Sorocea bonplandii;*
- *Classe 5 – Zollernia ilicifolia.*

A matriz de confusão em si pode ser encontrada na **Figura 45**.

Figura 45 – Matriz de confusão do modelo *machine learning* para amostras na forma de folha



Fonte: O autor (2024)

A interpretação da matriz é feita observando como as amostras destinadas à validação foram classificadas, levando em consideração sua classe real (*True Class* – eixo Y) e para qual grupo essas amostras foram alocadas (*Predicted Class* – eixo X). As amostras corretamente identificadas são representadas em células de coloração verde escura e escrita branca, correspondendo ao grupo das amostras verdadeiramente positivas. Os valores restantes apresentados em cada coluna ao longo da linha de cada classe verdadeira formam o grupo dos falsos positivos, representadas nas células de cor verde clara e escrita preta. Células de coloração branca e mais opacas são aquelas que se mostraram ausentes de resultados falsos positivos.

Como 30% das 500 varreduras representam 150 amostras para cada classe, o valor verdadeiramente positivo pode ser encontrado subtraindo-se este total de 150 pelo número de falsos positivos encontrados para cada classe.

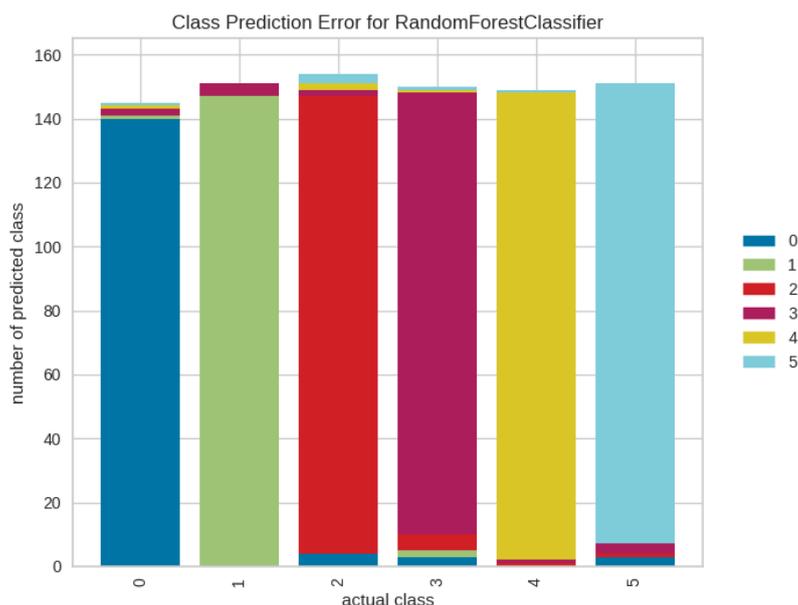
O número de falsos positivos encontrados para cada classe foi:

- Classe 0 – *Citronella gongonha* = 10 resultados;

- Classe 1 – *Jodina rhombifolia* = 3 resultados;
- Classe 2 – *Monteverdia aquifolium* = 7 resultados;
- Classe 3 – *Monteverdia ilicifolia* (espinheira-santa) = 12 resultados;
- Classe 4 – *Sorocea bonplandii* = 4 resultados;
- Classe 5 – *Zollernia ilicifolia* = 10 resultados.

O perfil de identificação para cada uma das classes também é apresentado de forma visual por meio do gráfico de erro de classificação, com cada classe sendo representada por uma cor. Em situações de falso positivo, as extremidades da coluna de uma determinada amostra apresentam colorações diferentes da sua original, sendo a cor distinta aquela que representa a espécie vegetal contaminante (**Figura 46**).

Figura 46 – Gráfico de erro de classificação para a amostra de folhas

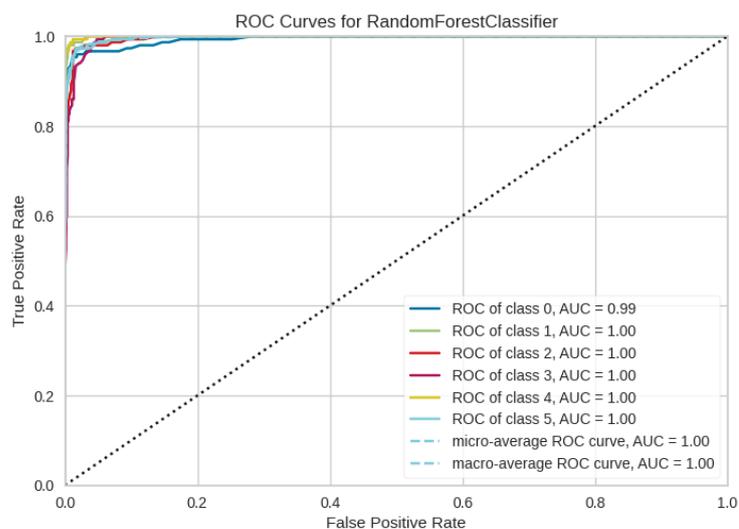


Legenda: Classe 0 – *Citronella gongonha*; Classe 1 – *Jodina rhombifolia*; Classe 2 – *Monteverdia aquifolium*.; Classe 3 – Espinheira-santa (*Monteverdia ilicifolia*); Classe 4 – *Sorocea bonplandii*; Classe 5 – *Zollernia ilicifolia*.

Fonte: O autor (2024)

A taxa de verdadeiros e falsos positivos também é expressa por meio do gráfico da curva ROC, cuja área foi calculada para avaliar da capacidade preditiva do modelo. Esse gráfico se encontra na **Figura 47**.

Figura 47 – Área sob a curva ROC para a amostra de folhas



Fonte: O autor (2024)

Todas as classes são dispostas em linhas e representadas no plano pelas mesmas cores que no gráfico do erro de classificação. No centro do gráfico têm-se a linha de *threshold*, a qual divide o gráfico em duas metades. A localização de cada classe no gráfico baseia-se na relação entre verdadeiros e falsos positivos em relação ao total de dados dessa classe. Abaixo do *threshold* ficam alocadas todas as amostras que apresentam predominantemente resultados falsos positivos, enquanto aquelas com a maioria de verdadeiros positivos ficando acima desse limite. Quanto mais próximo do vértice superior esquerdo, mais próxima de 1 é a área sob a curva ROC, resultado que indica uma predição ideal do modelo. Observa-se que todas as classes estudadas estão bem próximas ao ponto de valor 1, indicando uma excelente capacidade preditiva do modelo e corroborando com resultados encontrados pela matriz de confusão, ou seja, poucos resultados falsos positivos e uma quase totalidade de resultados verdadeiros positivos.

Para a classificação das amostras na forma de folha, também foram calculadas as métricas de Acurácia, *Recall*, *Precisão*, *F1 Score*, Índice *Kappa* e Coeficiente de Correlação de Matthews (MCC). Cada uma dessas métricas foi calculada 10 vezes, para cada instância da metodologia 10 *fold* (**Quadro 10**).

Quadro 10 – Métricas calculadas para o método *machine learning* das folhas

Instância	Acurácia	<i>Recall</i>	Precisão	F1	<i>Kappa</i>	MCC
1	0.9381	0.9381	0.9387	0.9379	0.9257	0.9259
2	0.9429	0.9429	0.9433	0.9425	0.9314	0.9317
3	0.9333	0.9333	0.9334	0.9331	0.9200	0.9201
4	0.9381	0.9381	0.941	0.9378	0.9257	0.9264
5	0.9381	0.9381	0.9405	0.9374	0.9257	0.9264
6	0.9000	0.9000	0.9022	0.9000	0.8800	0.8805
7	0.9333	0.9333	0.9347	0.9333	0.9200	0.9203
8	0.9429	0.9429	0.9462	0.9426	0.9314	0.9322
9	0.9667	0.9667	0.9674	0.9668	0.9600	0.9601
10	0.9476	0.9476	0.9488	0.9477	0.9371	0.9373
<b>Mean</b>	<b>0.9381</b>	<b>0.9381</b>	<b>0.9396</b>	<b>0.9379</b>	<b>0.9257</b>	<b>0.9261</b>
Std	0.0156	0.0156	0.0154	0.0157	0.0188	0.0187
CV (%)	1.6692	1.6692	1.6397	1.6708	2.0281	2.0192

Fonte: O autor (2024)

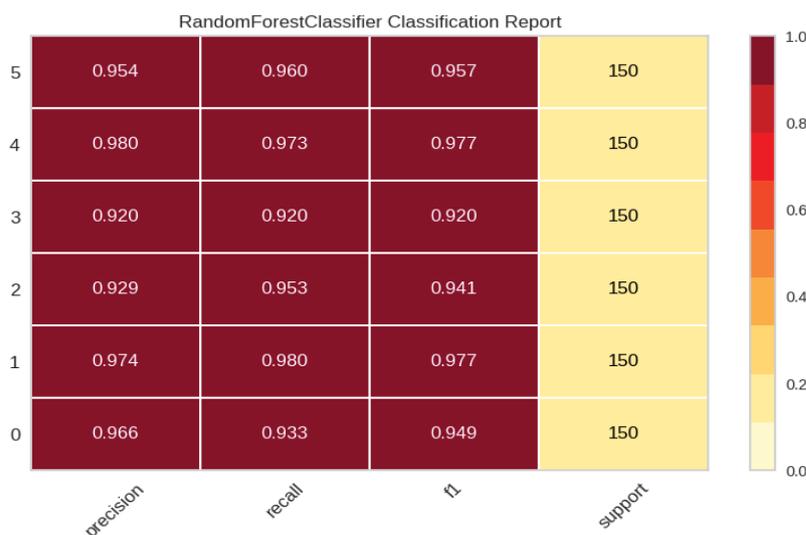
Destacados em amarelo estão os valores médios para cada uma das métricas estudadas. Enquanto a acurácia já foi comentada, as demais grandezas não foram elucidadas no estudo. A medida de *Recall*, também conhecida como sensibilidade, refere-se à capacidade do modelo de corretamente prever amostras verdadeiramente positivas, considerando amostras verdadeiramente negativas para seu cálculo. A precisão do modelo também avalia a capacidade preditiva do modelo para amostras positivas, no entanto, considera apenas amostras de caráter positivo, sejam elas verdadeiras ou falsas positivas. O F1 Score é entendido como a média harmônica entre as medidas de precisão e *Recall*. Todas essas métricas, para os modelos de *machine learning*, tem valores que vão de 0 a 1, sendo desejado que todas cheguem o mais próximo possível de 1, a qual representa uma medição perfeita do modelo. Por sua vez, o Índice *Kappa* relaciona-se com o grau de concordância entre um valor real e aquele alcançado pelo modelo de classificação. O Coeficiente de Correlação de Matthews (MCC) serve como uma medida de qualidade da classificação binária e multiclasse a partir dos valores verdadeiros, e falsos positivos e negativos, sendo bastante utilizado para matrizes de dados desbalanceadas. Ambas

as últimas métricas apresentam valores que vão de -1 a 1 e, assim como as grandezas anteriores, quanto mais próximo do máximo valor positivo, mais eficiente foi a capacidade preditiva do modelo (Ben-David, 2008; Bartholomew; Knott; Moustaki, 2011; Cormanich; Freitas; Rittner, 2011; Powers, 2011; Ferreira *et al.*, 2019; Chicco; Jurman, 2020; Zhu, 2020; Chicco; Warrens; Jurman, 2021; Humphrey *et al.*, 2022).

O valor médio encontrado para todas as métricas ficou em torno de 0,93, valor bem próximo ao ideal (1,0). Além disso, o coeficiente de variação (CV%) pode ser encontrado na porção final do quadro e refere-se à proximidade dos resultados de cada uma das métricas ao longo das 10 instâncias de cálculos. Resultados de CV% abaixo de 2,0 são ideais pois indicam a proximidade dos valores encontrados ao longo da validação. Para o modelo das folhas, o CV% ficou em torno de 1,7% para todas as métricas, com exceção do Índice *Kappa* e Coeficiente de Correlação de Matthews, os quais se apresentaram em torno de 2,0%.

Estes resultados demonstram como o modelo apresentou elevada capacidade preditiva, ao passo que não apresentou alterações bruscas das métricas utilizadas para avaliação de sua qualidade ao longo do procedimento de validação cruzada. Os resultados de precisão, *Recall*, e *F1 Score* para cada uma das seis classes de plantas usadas para a validação do modelo podem ser ilustrados a partir da construção do gráfico do relatório de classificação (**Figura 48**).

Figura 48 – Gráfico do relatório de classificação para a validação do modelo para amostras na forma de folha

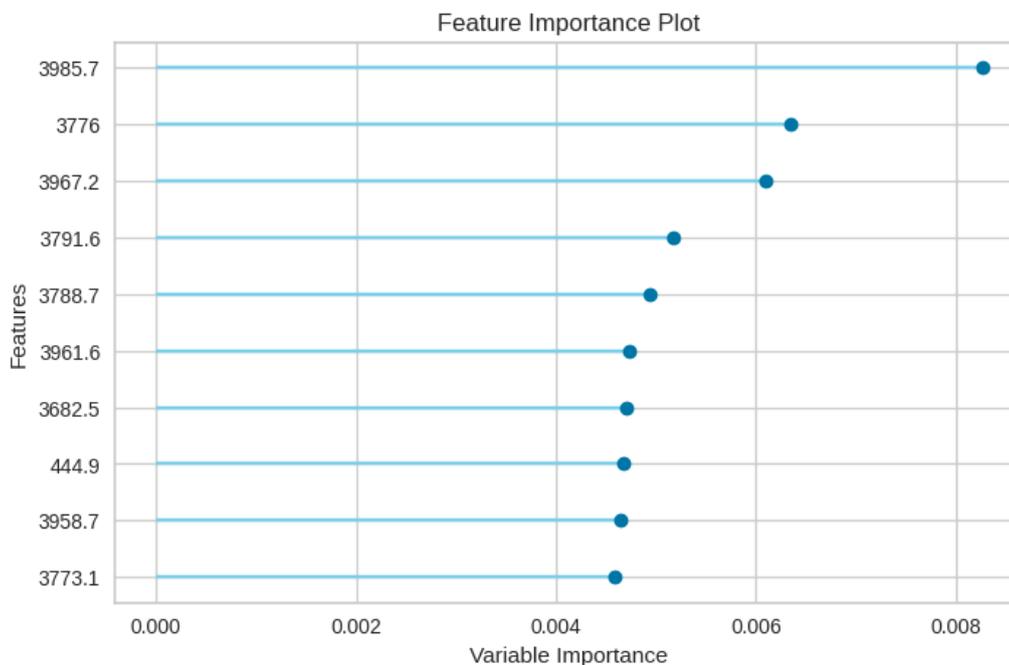


Fonte: O autor (2024)

Abrangendo todas as varreduras destinadas à validação de cada uma das seis classes, este gráfico demonstra a qualidade dos resultados em relação a uma escala de 0 a 1. A proximidade com o valor mais alto dessa faixa é o desejado para todas as métricas, e essa proximidade é representada a partir da cor da célula em que o resultado está inserido no gráfico. Quanto mais próximo do valor ideal, mais avermelhada fica a célula, com resultados extremamente próximos a 1 apresentando esta coloração de forma mais destacada. Percebe-se que todas as métricas foram representadas por esta cor avermelhada, demonstrando uma elevada habilidade preditiva do modelo.

Assim como nos modelos PCA e PLS-DA, a metodologia *machine learning* também permite identificar as principais regiões dos espectros responsáveis pela diferenciação das espécies. O *software* fornece os 10 números de ondas do MIR-FTIR que foram mais importantes nesta discriminação, atribuindo também um grau de importância (eixo X) para cada um desses valores encontrados (**Figura 49**).

Figura 49 – Números de onda mais importantes na classificação das diferentes espécies vegetais na forma de folha



Fonte: O autor (2024)

Para o modelo das folhas, a máquina atribuiu às regiões iniciais do espectro a maior importância, com a região de 3987,7  $\text{cm}^{-1}$  a mais relevante para discriminar as

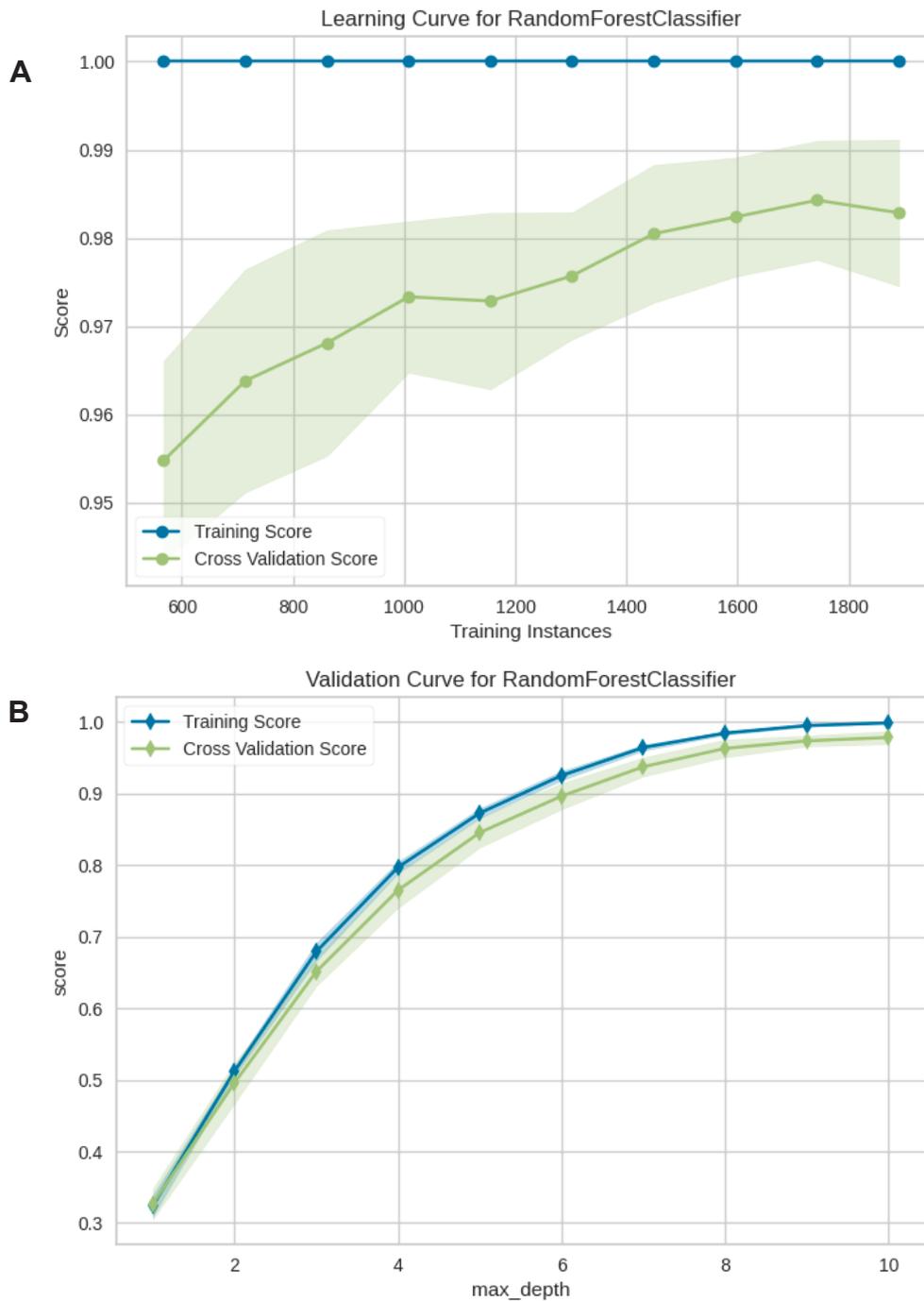
espécies, apresentando uma importância cerca de 34% superior ao próximo número de onda. Dentre os 10 números de onda destacados, o único que diverge dos demais é  $444,9 \text{ cm}^{-1}$ , demonstrando como a porção final do espectro também teve sua influência quanto à discriminação.

O modelo construído para folhas apresentou uma boa progressão da acurácia de calibração e validação ao longo das 10 instâncias de treino do modelo, subindo, respectivamente, de 87% a 94% e 33% a 93%. Pela matriz de confusão, foi identificado um número reduzido de falsos positivos para o modelo, resultado reforçado pelo perfil da curva ROC, a qual teve sua área calculada e também demonstrou bom desempenho do modelo. As métricas de acurácia, *Recall*, precisão, *F1 Score*, índice *Kappa* e MCC foram utilizadas para avaliar a capacidade preditiva do modelo, e todas apresentaram resultados bastante positivos, com valores em torno de 0,93, o que as coloca próximo ao considerado ideal (1,0). O valor das métricas de precisão, *Recall* e *F1 Score* foi avaliado para cada uma das classes de espécie vegetal a partir do relatório de classificação, atribuindo uma coloração à célula conforme sua proximidade do valor encontrado com 1,0. Essas métricas apresentaram resultados imediatos ao ideal, demonstrando bom desempenho preditivo para todas as classes. As regiões iniciais do espectro foram as mais relevantes para a discriminação da espécie pela máquina, principalmente na região de  $4000$  a  $3700 \text{ cm}^{-1}$ . No entanto, não se pode descartar as porções finais do espectro, uma vez que o número de onda de  $444,9 \text{ cm}^{-1}$  também apresentou elevada importância.

#### 5.4.3.2 Pós

Seguiu-se para a aplicação da estratégia de aprendizado de máquina para as amostras em pós, a qual teve sua construção realizada de forma semelhante ao modelo das folhas. Inicialmente avaliou-se a evolução acurácia da calibração e validação do modelo ao longo das 10 instâncias de treino (**Figura 50**).

Figura 50 – Progressão da acurácia da calibração e validação do modelo *machine learning* para amostras na forma de pó



Legenda: A - Progressão da acurácia da calibração do modelo; B - Progressão da acurácia da validação do modelo

Fonte: O autor (2024)

Da mesma forma que no modelo para folhas, as árvores de decisão formaram uma área que agrega todos os resultados encontrados, e sua média representa o valor

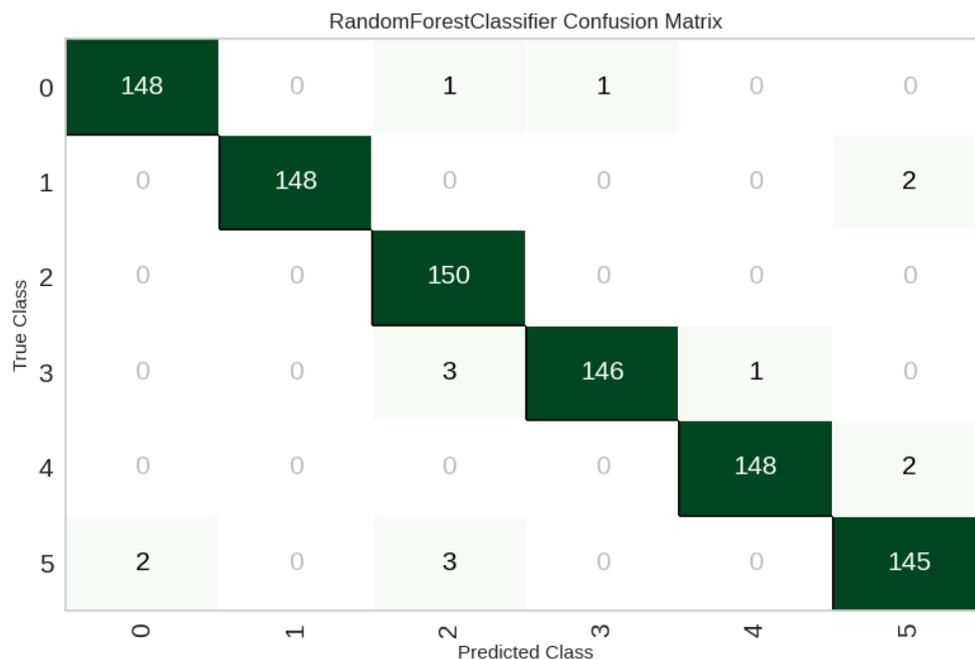
de acurácia para o modelo como um todo. A acurácia para a calibração (**Figura 49 A**) apresentou valor inicial de 95%, sendo elevada até um total de 98% após aplicação de todos os dados disponíveis. Por sua vez, a acurácia para validação teve início com 32%, sendo ampliada até 98%. Ambos os resultados são extremamente positivos e indicam uma excelente capacidade preditiva do modelo.

O estudo segue com a construção da matriz de confusão, visando identificar as espécies erroneamente classificadas pelo modelo. As classes foram organizadas da mesma forma que para o modelo de pós:

- Classe 0 – *Citronella gongonha*;
- Classe 1 – *Jodina rhombifolia*;
- Classe 2 – *Monteverdia aquifolium*.;
- Classe 3 – *Monteverdia ilicifolia* (espinheira-santa);
- Classe 4 – *Sorocea bonplandii*;
- Classe 5 – *Zollernia ilicifolia*,

A matriz de confusão podendo ser encontrada na **Figura 51**.

Figura 51 – Matriz de confusão do modelo *machine learning* para amostras na forma de pó



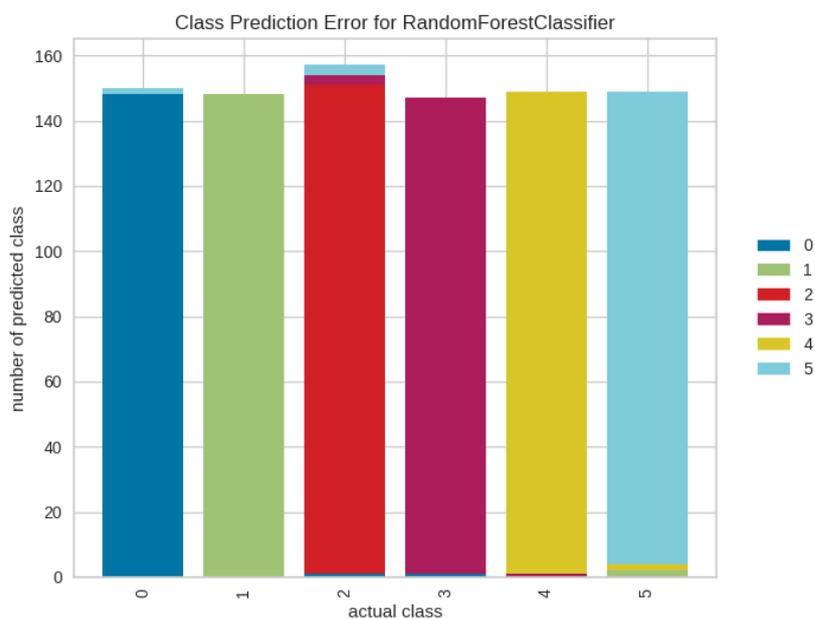
Fonte: O autor (2024)

Em relação ao modelo para folhas, observa-se uma redução significativa no número de falsos positivos, os quais se mostraram em uma quantidade extremamente reduzida. O número de falsos positivos de cada classe foram os seguintes:

- *Classe 0 – Citronella gongonha* = 2 resultados;
- *Classe 1 – Jodina rhombifolia* = 2 resultados;
- *Classe 2 – Monteverdia aquifolium*; = 0 resultados;
- *Classe 3 – Monteverdia ilicifolia* (espinheira-santa) = 4 resultados;
- *Classe 4 – Sorocea bonplandii* = 2 resultados;
- *Classe 5 – Zollernia ilicifolia* = 5 resultados.

A representação visual deste perfil de identificação para cada uma das classes também foi desenvolvida, e pode ser encontrada no gráfico de erro de classificação, encontrado na **Figura 52**.

Figura 52 – Gráfico de erro de classificação para a amostra de pó

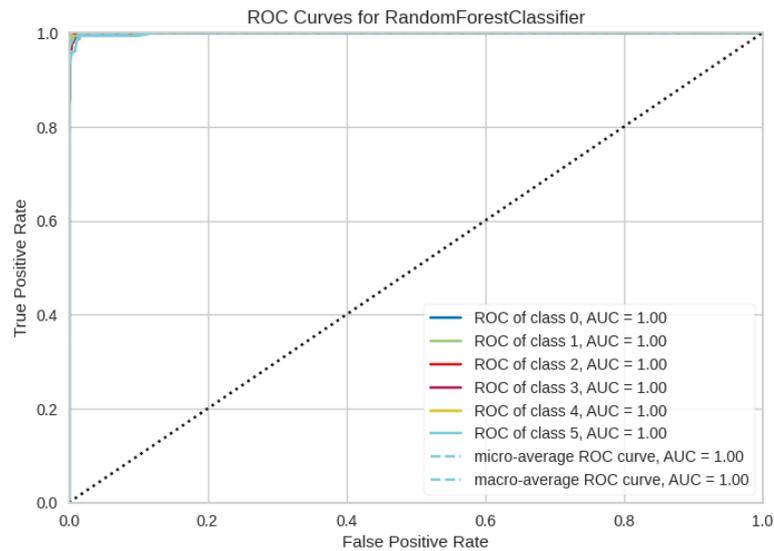


Legenda: Classe 0 – *Citronella gongonha*; Classe 1 – *Jodina rhombifolia*; Classe 2 – *Monteverdia aquifolium*.; Classe 3 – Espinheira-santa (*Monteverdia ilicifolia*); Classe 4 – *Sorocea bonplandii*; Classe 5 – *Zollernia ilicifolia*.

Fonte: O autor (2024)

A curva ROC também foi gerada para análise da acurácia do modelo e disposição das classes ao longo do gráfico (**Figura 53**).

Figura 53 – Área sob a curva ROC para a amostra de pó



Fonte: O autor (2024)

As classes estão representadas no gráfico com as mesmas cores utilizadas no gráfico do erro de classificação. Todas as classes estão praticamente sobrepostas ao vértice superior esquerdo do gráfico, indicando que pouquíssimos resultados falsos positivos foram encontrados. Consistente com resultados anteriores, a análise da curva ROC também demonstra um desempenho preditivo praticamente perfeito do modelo desenvolvido, onde as amostras foram classificadas corretamente com um mínimo de erros. A área sob a curva ROC encontrada está muito próxima de 1, valor ideal e que indica uma capacidade preditiva perfeita do modelo.

Para avaliar a qualidade do modelo foram calculadas as métricas de acurácia, *Recall*, precisão, *F1 Score*, índice *Kappa* e MCC foram calculadas 10 vezes, cujos resultados são apresentados na **Quadro 11**.

Quadro 11 – Métricas calculadas para o método *machine learning* dos pós

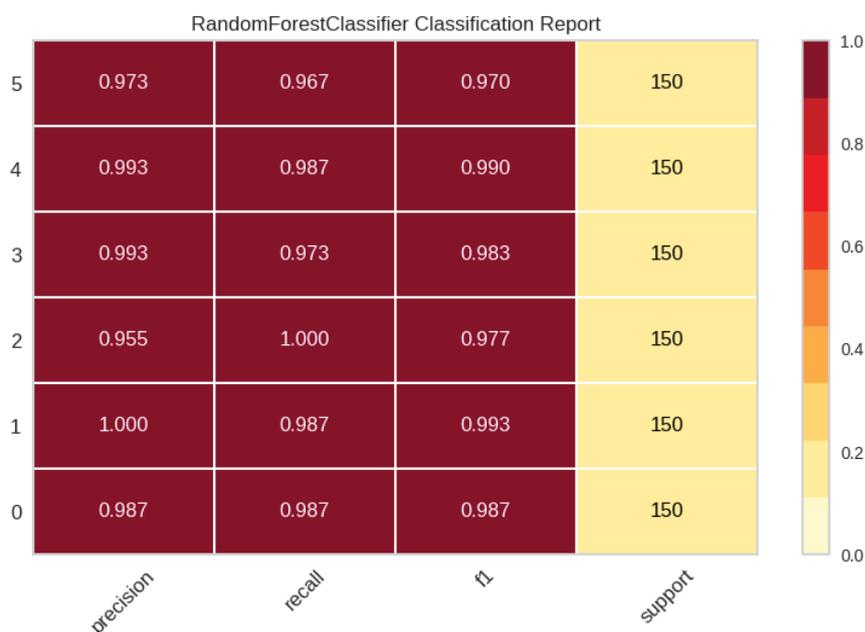
Instância	Acurácia	<i>Recall</i>	Precisão	F1	<i>Kappa</i>	MCC
1	0.9810	0.9810	0.9819	0.9811	0.9771	0.9773
2	0.9905	0.9905	0.9910	0.9905	0.9886	0.9887
3	0.9857	0.9857	0.9868	0.9858	0.9829	0.9830
4	0.9810	0.9810	0.9815	0.9809	0.9771	0.9772
5	0.9810	0.9810	0.9813	0.9810	0.9771	0.9772
6	0.9857	0.9857	0.9862	0.9858	0.9829	0.9829
7	0.9952	0.9952	0.9954	0.9952	0.9943	0.9943
8	0.9667	0.9667	0.9680	0.9667	0.9600	0.9603
9	0.9905	0.9905	0.9906	0.9905	0.9886	0.9886
10	0.9714	0.9714	0.9723	0.9713	0.9657	0.9659
<i>Mean</i>	0.9829	0.9829	0.9835	0.9829	0.9794	0.9795
Std	0.0083	0.0083	0.0080	0.0083	0.0100	0.0099
CV (%)	0.8444	0.8444	0.8134	0.8444	1.0210	1.0107

Fonte: O autor (2024)

Destacada em amarelo, a média das métricas estudadas ficou em torno de 0,98, um valor extremamente próximo ao ideal de 1,0. O CV%, por sua vez, demonstrou-se muito reduzido, indicando uma constância das métricas calculadas ao longo das instâncias de cálculo propostas para o modelo. A combinação destes dois fatores aponta para um modelo eficiente e robusto, com alto desempenho preditivo para amostras verdadeiras e falsas, positivas e negativas.

Complementando este estudo, as métricas de precisão, *Recall*, e F1 Score foram calculadas para cada uma das classes de plantas, cujos resultados estão representados no gráfico do relatório de classificação (**Figura 54**).

Figura 54 – Gráfico do relatório de classificação para a validação do modelo para amostras na forma de pó



Fonte: O autor (2024)

Para todas as classes estudadas, as métricas avaliadas mostraram-se próximas ao valor ideal de 1, como indicado pela coloração vermelha das células correspondentes no gráfico, conforme a referência cromática encontrada na lateral do relatório. O elevado valor encontrado para cada classe é consistente com resultados anteriores, demonstrando como a capacidade preditiva do modelo abrange não apenas a *M. ilicifolia*, mas todas as espécies envolvidas.

Finalizando o modelo dos pós foram verificadas quais as principais regiões do espectro que o *software* utilizou para discriminar as amostras. Diferentemente do resultado encontrado para as folhas, o modelo dos pós não destacou nenhum número de onda específico. A máquina indicou uma importância geral do espectro, sem regiões específicas para discriminação das espécies. Este perfil é congruente com o que foi observado nas amostras em pós durante o PCA e PLS-DA, revelado uma forma de apresentação mais variável, porém capaz de fornecer informações importantes que permitem diferenciar as espécies vegetais em questão.

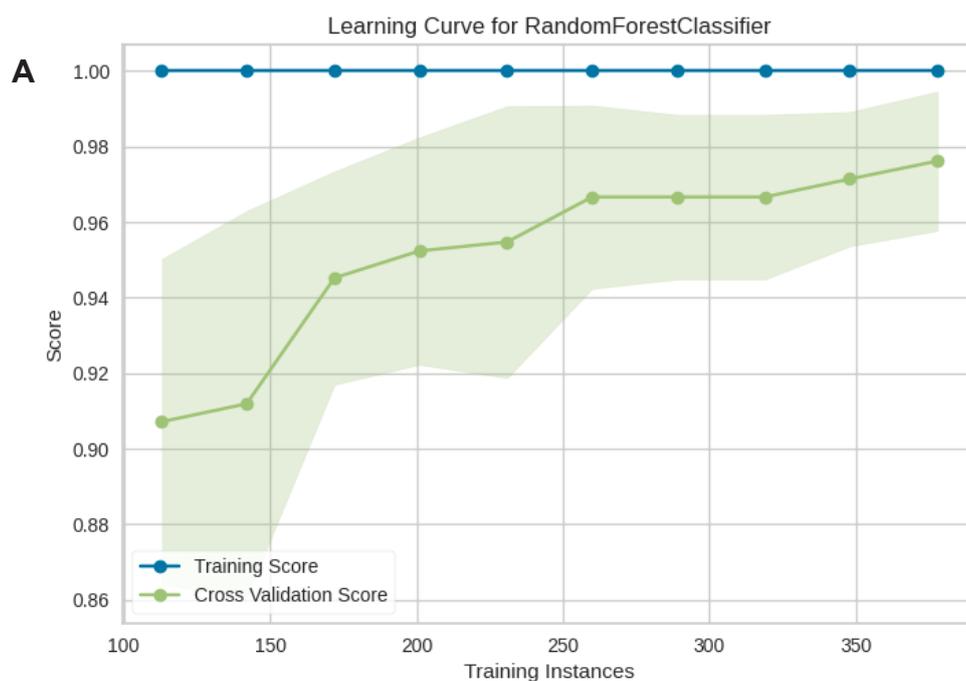
O modelo para os pós apresentou acurácias bastante positivas e de evolução significativa tanto na fase de calibração (de 95% a 98%) quanto na validação (de 32% a 98%). O número de falsos positivos foi reduzido em comparação ao de verdadeiros negativos, conforme demonstrado pela matriz de confusão e confirmado pela curva

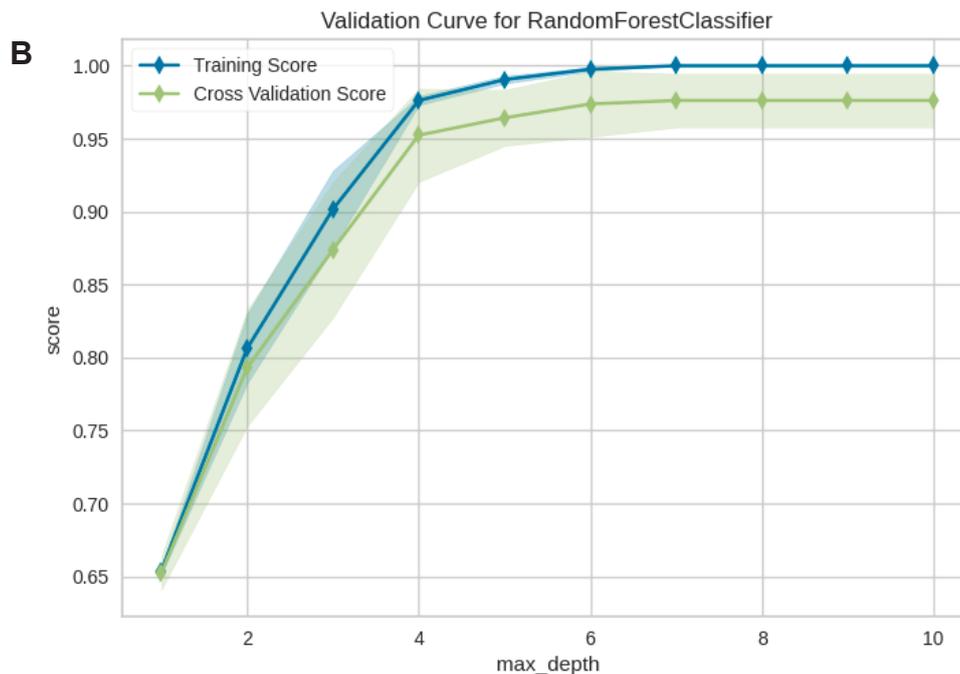
ROC, cuja área calculada se aproximou de 1. Para avaliar a qualidade do modelo, foram calculadas a acurácia, *Recall*, precisão, *F1 Score*, índice *Kappa* e *MCC*, todas com valor médio bastante elevados, próximo a 0,98. Entre essas métricas, a precisão, *Recall* e *F1 Score* foram avaliadas para todas as classes, e os resultados também indicaram um excelente desempenho preditivo do modelo, com todos próximos ao ideal. Com relação aos números de onda mais importantes, não houve destaque para nenhum em particular, sugerindo que o *software* considerou todo o espectro para a discriminação das espécies vegetais.

#### 5.4.3.3 Extratos

Como último modelo desenvolvido para a metodologia de aprendizado de máquina têm-se os extratos, o qual apresentou os melhores resultados pela técnica quimiométrica. O desenvolvimento de seu modelo seguiu o mesmo padrão dos anteriores, iniciando pela avaliação da progressão da acurácia ao longo das 10 instâncias de treino, tanto para calibração quanto validação (**Figura 55**).

Figura 55 – Progressão da acurácia da calibração e validação do modelo *machine learning* para amostras na forma de extrato





Legenda: A - Progressão da acurácia da calibração do modelo; B - Progressão da acurácia da validação do modelo

Fonte: O autor (2024)

Analisando a curva da calibração (**Figura 55 A**), observamos que inicialmente a acurácia média encontrada pelas arvores de decisão foi de 90%, aumentando para 98% ao longo das instâncias de treino. Para os dados de validação, essa evolução também foi perceptível, porém com uma diferença: sua acurácia inicial foi de 65%, o que representa a menor progressão da acurácia em comparação aos modelos anteriores, pois partiu de um valor inicial mais elevado, alcançando também 98% ao final das iterações. Ambos esses resultados contribuem para um modelo de elevado desempenho preditivo.

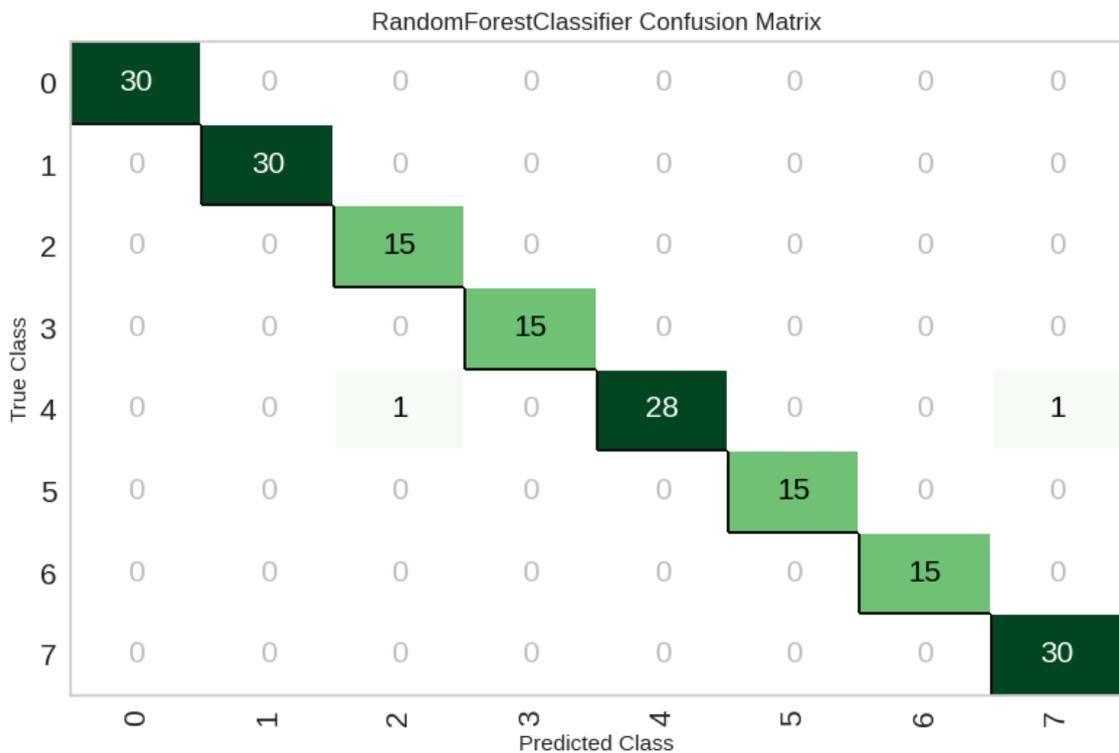
Na classificação das amostras e identificação de resultados falsos positivos, a matriz de confusão para o modelo de extratos foi desenvolvida. Em contraste com as metodologias anteriores, foram interpretadas de forma distinta as classes propostas para o estudo, dividindo em grupos diferentes aquelas espécies vegetais de múltiplas localidades. Sendo assim, para o modelo dos extratos, as classes foram as seguintes:

- *Classe 0 – Citronella gongonha;*
- *Classe 1 – Jodina rhombifolia;*
- *Classe 2 – Monteverdia aquifolium A;*
- *Classe 3 – Monteverdia aquifolium B;*

- Classe 4 - *Monteverdia ilicifolia* (espinheira-santa);
- Classe 5 – *Sorocea bonplandii* X;
- Classe 6 – *Sorocea bonplandii* Y;
- Classe 7 – *Zollernia ilicifolia*.

A matriz de confusão pode ser encontrada na **Figura 56**.

Figura 56 – Matriz de confusão do modelo *machine learning* para amostras na forma de extrato

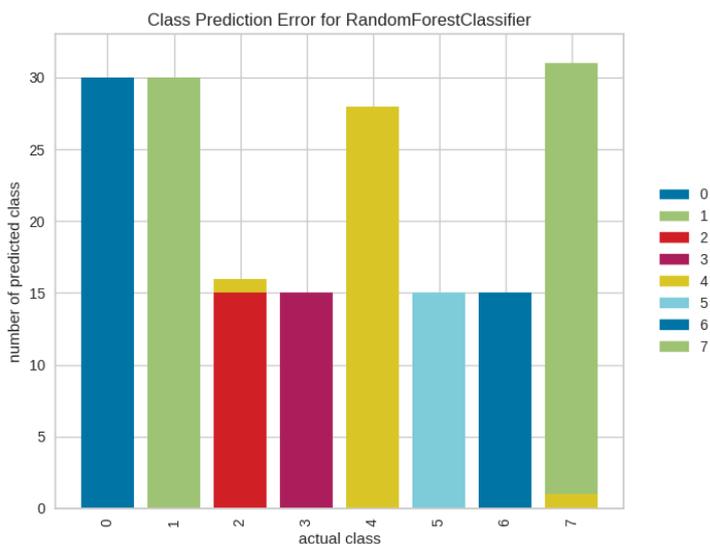


Fonte: O autor (2024)

Foram identificados apenas dois resultados erroneamente interpretados pelo modelo, ambos à mesma classe, que representa a espinheira-santa original. Como a matriz dos extratos foi construída com base em 100 varreduras, apenas 30 resultados foram utilizados para a validação, resultando em um número reduzido de verdadeiros positivos em comparação com os modelos de folhas e pós. Para as classes 2 e 3, e 5 e 6, devido a diferenciação detectada em relação a localidade de origem e tempo de colheita, respectivamente, apenas 15 resultados de validação foram alocados para cada uma delas, destacando que nenhum resultado falso positivo foi encontrado para ambas as espécies, mesmo com a separação detalhada realizada pelo *software*.

O perfil de identificação das amostras falsas positivas pode ser visualizado no gráfico de erro de classificação, na **Figura 57**.

Figura 57 – Gráfico de erro de classificação para a amostra de extrato

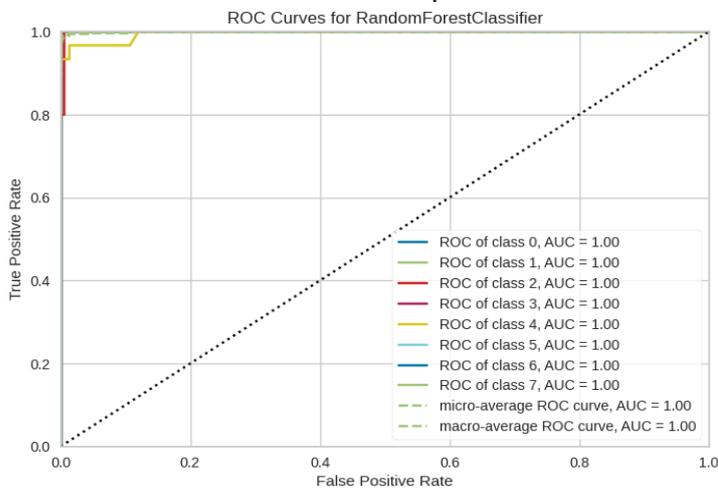


Legenda: Classe 0 – *Citronella gongonha*; Classe 1 – *Jodina rhombifolia*; Classe 2 – *Monteverdia aquifolium* A; Classe 3 – *Monteverdia aquifolium* B; Classe 4 - Espinheira-santa (*Monteverdia ilicifolia*); Classe 5 – *Sorocea bonplandii* X; Classe 6 – *Sorocea bonplandii* Y; Classe 7 – *Zollernia ilicifolia*.

Fonte: O autor (2024)

Também correlacionando a acurácia do modelo com o número de falsos positivos encontrados, têm-se a curva ROC, presente na **Figura 58**.

Figura 58 – Área sob a curva ROC para a amostra de extrato



Fonte: O autor (2024)

Da mesma forma que nos modelos anteriores, as classes estão representadas pelas mesmas cores que no gráfico do erro de classificação. Como praticamente todas as classes não apresentaram resultados faltos positivos, elas estão, em sua maioria, no vértice superior esquerdo do gráfico, apresentando uma área sob a curva ROC de 1. No entanto, para a classe da espinheira-santa, representada pela cor amarela, dois resultados errôneos foram encontrados, tornando-a a única distanciada do vértice. Nota-se, contudo, que esse distanciamento é mínimo, pois apenas dois resultados falsos positivos foram computados, resultando em uma área sob a curva ROC para essa classe também próxima a 1. Esses resultados indicam uma elevada capacidade preditiva do modelo.

A qualidade do modelo foi avaliada a partir das métricas: acurácia, *Recall*, precisão, *F1 Score*, índice *Kappa* e *MCC*. Cada uma dessas foi calculada em 10 instâncias, e ao fim suas medias foram computadas (**Quadro 12**).

Quadro 12 – Métricas calculadas para o método *machine learning* dos extratos

Instância	Acurácia	<i>Recall</i>	Precisão	<i>F1</i>	<i>Kappa</i>	<i>MCC</i>
1	0.9762	0.9762	0.9821	0.9770	0.9724	0.9730
2	0.9762	0.9762	0.9792	0.9753	0.9723	0.9729
3	0.9762	0.9762	0.9792	0.9761	0.9723	0.9730
4	0.9762	0.9762	0.9821	0.9755	0.9723	0.9730
5	0.9762	0.9762	0.9821	0.9770	0.9724	0.9730
6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0.9762	0.9762	0.9810	0.9766	0.9724	0.9730
8	0.9762	0.9762	0.9792	0.9753	0.9723	0.9729
9	0.9286	0.9286	0.9315	0.9284	0.917	0.9176
10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
<i>Mean</i>	0.9762	0.9762	0.9796	0.9761	0.9723	0.9729
Std	0.0184	0.0184	0.0178	0.0185	0.0214	0.0213
CV (%)	1.8848	1.8848	1.8170	1.8953	2.2010	2.1893

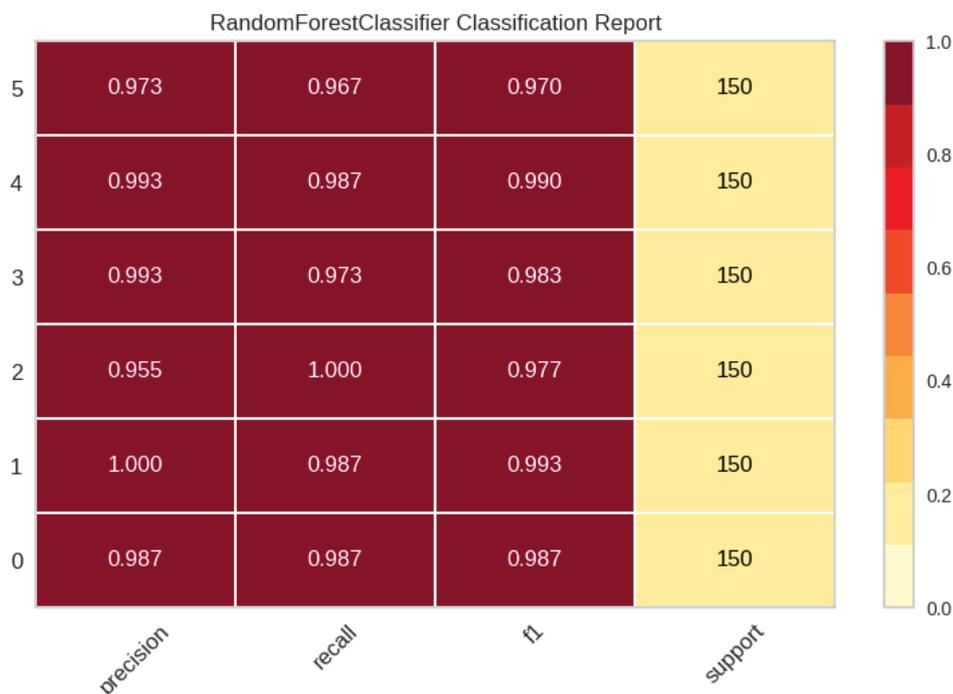
Fonte: O autor (2024)

O resultado encontrado para cada uma das métricas foi bastante elevado, com suas médias, destacadas pela coloração amarela, próximas a 0,97, indicando

proximidade ao valor ideal de 1,0. O CV% foi calculado e apresentou resultados que variam de 1,8% a 2,2%, indicando que foram pequenas as variações de cada uma das métricas ao longo das 10 instâncias de treino. Esses resultados reforçam a eficiência do modelo em prever corretamente todas as amostras.

As métricas de precisão, *Recall*, e *F1 Score* foram novamente avaliadas, desta vez para cada uma das classes de espécies vegetais estudadas, resultando na construção o gráfico do relatório de classificação (**Figura 59**).

Figura 59 – Gráfico do relatório de classificação para a validação do modelo para amostras na forma de extrato



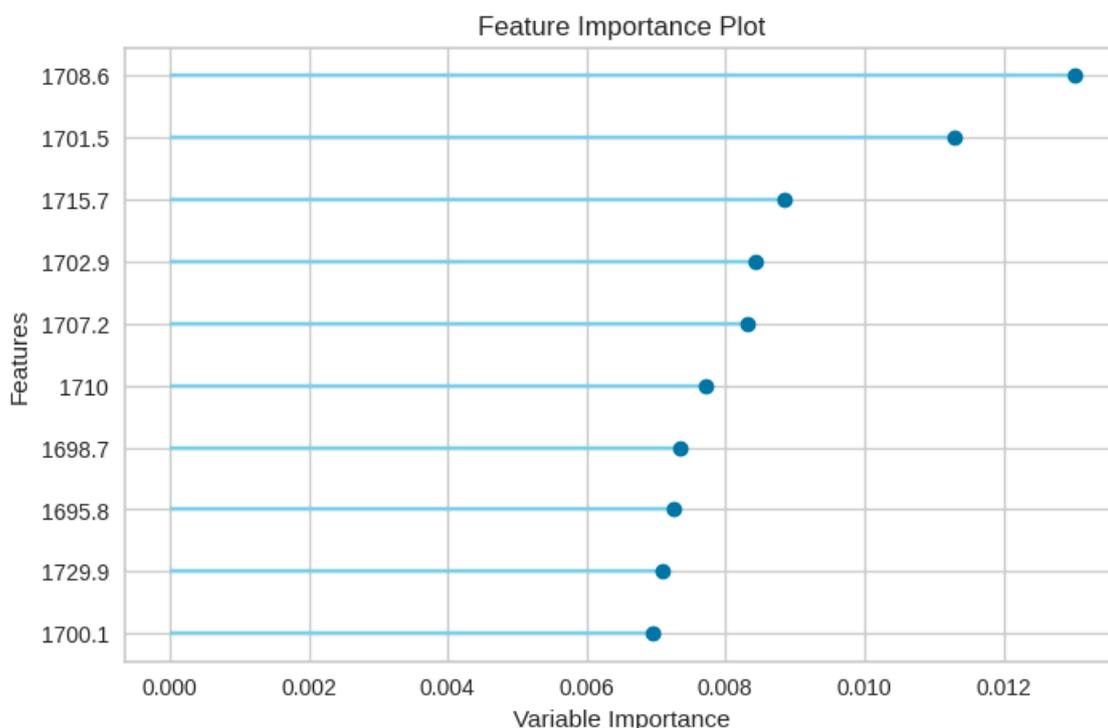
Fonte: O autor (2024)

Enquanto para a construção da matriz de confusão e curva ROC o *software* diferenciou as amostras de localidades distintas, esse mesmo tratamento não foi aplicado no relatório de classificação das amostras. Assim, a classe 1 compreende completamente a espécie *Monteverdia aquifolium*, enquanto a classe 3 engloba totalmente *Sorocea bonplandii*. Comparando os resultados encontrados para cada classe com o gradiente colorimétrico a direita do gráfico, percebe-se que todos os resultados foram próximos ao ideal e bastante positivos, com a totalidade dos valores representados em células com a coloração vermelha característica do relatório. Esta

avaliação complementa os resultados anteriores e aponta para um modelo de predição excelente.

As regiões do espectro interpretadas como mais importantes na discriminação das espécies, para o modelo de extratos, foram estabelecidos. Os 10 números de onda de maior impacto se encontram na **Figura 60**.

Figura 60 – Números de onda mais importantes na classificação das diferentes espécies vegetais na forma de extrato



Fonte: O autor (2024)

O resultado encontrado para o modelo de extratos por *machine learning* foi consistente com o do modelo PCA, ambos apontando a região de  $1750\text{ cm}^{-1}$  como a principal discriminadora das espécies envolvidas no estudo. Observa-se que todas as 10 regiões de maior impacto se encontram em números de onda adjacentes, englobando a área de  $1729\text{ cm}^{-1}$  a  $1695\text{ cm}^{-1}$ , com destaque maior para  $1708,6\text{ cm}^{-1}$  e  $1701,5\text{ cm}^{-1}$ .

A criação do modelo *machine learning* para os extratos como metodologia complementar ao PLS-DA apresentou resultados promissores. A progressão da acurácia de calibração e validação foi perceptível ao longo das 10 instâncias de treino, com a acurácia evoluindo, respectivamente, de 90% a 98%, e 65% a 98%. O número

de falsos positivos foi avaliado a partir da matriz de confusão, a qual demonstrou uma divisão das classes diferente dos demais modelos, separando aquelas que apresentaram mais de uma localidade de origem e tempo de colheita. Apenas dois resultados falsos positivos foram encontrados, fato também em evidência pelo perfil da curva ROC que, apesar dessa ocorrência, apresentou áreas sob a curva de 1 ou valores extremamente próximos ao ideal.

A avaliação por meio das métricas de acurácia, *Recall*, precisão, *F1 Score*, índice *Kappa* e MCC indicou um elevado desempenho preditivo do modelo, com as médias para essas métricas em torno de 0,97 e CV% próximo a 2,0%. O estudo das métricas de precisão, *Recall*, *F1 Score* também foi realizado para cada uma das classes envolvidas no estudo, e os resultados encontrados foram positivos, com todos os valores se aproximando do ideal e representados pela coloração vermelha típica do gráfico de relatório de classificação. Esses resultados refletem uma elevada qualidade e capacidade preditiva do modelo de extratos. As principais regiões de discriminação das espécies no espectro foram a faixa de 1729  $\text{cm}^{-1}$  a 1695  $\text{cm}^{-1}$ , com os números de onda de 1708,6  $\text{cm}^{-1}$  e 1701,5  $\text{cm}^{-1}$  sendo os mais relevantes.

## 5.5 CONCLUSÃO

Com a atual escassez de métodos analíticos capazes de identificar e discriminar a espinheira-santa de espécies similares, o presente estudo propôs um método inovador utilizando a quimiometria e *machine learning*, criando modelos de elevada eficiência para identificar corretamente a espécie vegetal em questão.

No âmbito da quimiometria, o método PCA demonstrou resultados satisfatórios para todas as formas de apresentação incluídas no estudo. Para as folhas e seus pós, foram utilizadas três componentes principais para explicar o modelo, enquanto para os extratos foram selecionadas cinco. Mesmo com diferentes perfis de apresentação das espécies na representação tridimensional, uma característica bastante positiva foi a capacidade do modelo de discriminar não apenas a espinheira-santa das demais espécies, mas também todas as plantas entre si. Enquanto o perfil dos extratos se mostrou mais homogêneo, o dos pós apresentou as amostras dispostas em longas faixas, indicando uma heterogeneidade entre os dados da matriz. Os principais números de onda para identificação das espécies mostraram semelhanças para os extratos e pós, porém, o perfil das folhas se mostrou mais inespecífico, indicando a necessidade de uma análise integral de seu espectro para identificação eficaz da espécie vegetal.

O PLS-DA apresentou resultados satisfatórios para as três formas de apresentação, mas o destaque foi o modelo desenvolvido para os extratos. Com sensibilidade, especificidade e acurácia, de 94%, 100% e 99%, respectivamente, sem *outliers* e cinco variáveis latentes, este modelo também apresentou resultados semelhantes quanto aos números de onda relevantes ao modelo PCA. Os modelos desenvolvidos para os pós e folhas também apresentaram resultados positivos, mas com uma menor capacidade preditiva quando comparados ao extrato.

Dada a limitação da quimiometria em lidar com matrizes de dados extremamente grandes e complexas, aplicou-se o *machine learning*, buscando complementar os estudos prévios. Utilizando a metodologia *Random Forest*, os resultados para as três formas de apresentação foram similares e de elevada eficiência. Todas as amostras demonstraram uma evolução significativa da acurácia ao longo das 10 instâncias de treinamento do método *10 Fold*, tanto para calibração quanto para validação. A média dessas acurácias para os três tipos de amostra foi de

96,7% e 96,3%, respectivamente. De acordo com as matrizes de confusão construídas, o número de falsos positivos foi bastante reduzido em todos os modelos, corroborando com o resultado da área sob a curva ROC, cujo resultado encontrado para todas as amostras foi extremamente próxima ao valor ideal de 1. Para avaliar a capacidade preditiva do modelo, foram verificadas as métricas de acurácia, *Recall*, precisão, *F1 Score*, índice *Kappa* e MCC, além do CV% para todas. O resultado dessas métricas foi bastante elevado para as três formas de apresentação, com todas se aproximando de seus respectivos valores ideais. Além disso, observou-se uma constância dos resultados para cada métricas ao longo das 10 instâncias de treino, indicando uma elevada regularidade dos dados que compuseram o modelo. O gráfico do relatório de classificação utilizou da precisão, *Recall* e *F1 Score* para avaliar a capacidade preditiva do modelo para cada uma das espécies envolvidas. Partindo de uma referência colorimétrica, observou-se que, para todas as amostras, os resultados foram representados pela coloração vermelha das células que os continham, demonstrando proximidade com o valor ideal. Finalmente, as principais regiões do espectro responsáveis pela discriminação das espécies foram registradas. Enquanto alguns resultados remeteram aos perfis anteriormente obtidos, como o dos extratos e folhas, a principal diferença se deu nos pós, no qual o modelo indicou a necessidade de uma análise íntegra para discriminação.

A criação dos modelos quimiométricos e de *machine learning* mostrou-se complementar e indica ser um método inovador para identificação de adulterações/falsificações no âmbito comercial. A metodologia desenvolvida neste estudo passará por análises futuras com amostras comerciais a fim de aprimorar ainda mais sua robustez e eficiência, dado seu potencial de evolução constante. Os modelos construídos apresentam potencial de aplicação em diversas valências analíticas e podem se tornar uma referência para metodologias futuras.

## 6 REFERÊNCIAS

- ALI, J., KHAN, R., AHMAD, N. Random Forests and Decision Trees. **IJCSI International Journal of Computer Science Issues**, v. 9, ed. 5, n. 3, set. 2012. Disponível em: [www.IJCSI.org](http://www.IJCSI.org).
- AL-KHAYRI, J.M.; *et al.* Flavonoids as Potential Anti-Inflammatory Molecules: A Review. **Molecules**, v. 27, 2901, 2022.
- ALVES, F. M. S.; *et al.* *Aloe vera* extract as a plasticizer agente to polymeric films: A structural and functional component. **Journal of Drug Delivery Science and Technology**, v. 99. 2024.
- ANDRADE, S. F.; *et al.* Evaluation of the Antiulcerogenic Activity of *Maytenus robusta* (Celastraceae) in Different Experiments Ulcer Models. **Journal of Ethnopharmacology**, v. 113, 2007, p. 252-257.
- ANTUNES, K. A.; *et al.* 2023a. Comprehensive High-Performance Thin-Layer Chromatography Analysis of *Monteverdia ilicifolia* Leaf and its Adulterants. **Natural Product Research**. 2023.
- ANTUNES, K. A.; *et al.* 2023b. Authentication and Quality Control of the Brazilian Traditional Herb “Espinheira-Santa” (*Monteverdia ilicifolia*) by Morpho-Anatomy and Microscopy. **Microscopy and Microanalysis**, v.29, 2023, p. 1809-1821.
- ALMEIDA, C.; *et al.* **Comercialização de Espinheira-Santa na Zona Central de Pelotas**. In. XIV ENPOS. Universidade Federal de Pelotas (UFPEL). 2012.
- ALZAND, K. I.; MOHAMED, M. A. Flavonoids: Chemistry, Biochemistry and Antioxidant Activity. **Journal of Pharmacy Research**, v. 5, ed. 8, p. 4013-4020. ago. 2012.
- ARAÚJO JÚNIOR, R. F.; *et al.* *Maytenus ilicifolia* dry extract protects normal cells, induces apoptosis and regulates Bcl-2 in human cancer cells. **Experimental Biology and Medicine**, v. 238, n. 11, p. 1251–1258, 2013.
- ASTM - AMERICAN SOCIETY FOR TESTING AND MATERIALS. ASTM E1655 - standard practices for infrared multivariate quantitative analysis. 2017.
- BADSHAH, S.L.; *et al.* **Antiviral activities of flavonoids, Biomedicine & Pharmacotherapy**. V. 140, 111596. 2021.
- BAGGIO, C.H.; *et al.* Muscarinic-dependent inhibition of gastric emptying and intestinal motility by fractions of *Maytenus ilicifolia* Mart ex. Reissek. **J Ethnopharmacol**, v. 123, p. 385–391. 2009.
- BARTHOLOMEW, D. J., KNOTT, M., MOUSTAKI, I. **Latent Variable Models and Factor Analysis: A Unified Approach**. Wiley, ed. 3. 2011.

BELTRAME, F. L.; *et al.* A quantitative validated method using liquid chromatography and chemometric analysis for evaluation of raw material of *Maytenus ilicifolia* (Schrad.) Planch., Celastraceae. **Quimica Nova**, v. 35, n. 2, p. 327–331, 2012.

BENAIGES, A.; GUILLÉN, P. Botanical Extracts. **Analysis of Cosmetics Products**. 2007, p. 345-363.

BEN-DAVID, A. About the relationship between ROC curves and Cohen's *Kappa*. **Engineering Applications of Artificial Intelligence**, v. 21, n. 6, p. 874–882, set. 2008.

BERNATONIENE, J.; KOPUSTINSKIENE, D. M. The Role of Catechins in Cellular Responses to Oxidative Stress. **Molecules**. ed. 4. v. 23. 20 abr. 2018.

BIANCOLILLO, A.; MARINI, F. **Chemometrics Applied to Plant Spectral Analysis**. 1º ed. Elsevier B.V., 2018.

BIRAL, L.; LOMBARDI, J.A. *Monteverdia aquifolium* (Mart.) Biral: *Celastraceae in Flora e Funga do Brasil*. Jardim Botânico do Rio de Janeiro. Disponível em: <<https://floradobrasil.jbrj.gov.br/FB604475>>. Acesso em: 27 set. 2024.

BIRAL, L.; LOMBARDI, J.A. *Monteverdia ilicifolia* (Mart. ex Reissek) Biral: *Celastraceae in Flora e Funga do Brasil*. Jardim Botânico do Rio de Janeiro. Disponível em: <<https://floradobrasil.jbrj.gov.br/FB605053>>. Acesso em: 27 set. 2024.

BIRAL, L.; SIMMONS, M.P.; SMIDT, E.C.; TEMROCK, L.R.; BOLSON, M.; ARCHER, R.H.; LOMBARDI, J.A. Systematics of New World *Maytenus* (Celastraceae) and a New Delimitation of the Genus. **Syst Bot.** v.42 p. 680–693, 2017.

BIRAL, L.; 2023. The correct spelling of an “espinheira-santa” species (*Monteverdia*, Celastraceae): *aquifolium* vs. *aquifolia*. **Phytotaxa**, v. 591, ed. 3, 2023. p 239-240.

BOCHNER, R.; *et al.* Problemas associados ao uso de plantas medicinais comercializadas no mercado de madureira, município do rio de janeiro, brasil. **Revista Brasileira de Plantas Mediciniais**, v. 14, n. 3, p. 537–547, 2012.

BRASIL. **Farmacopeia Brasileira 6ª Edição: Volume 1**. Agência Nacional de Vigilância Sanitária: ANVISA. Brasília. 2019. Disponível em: <https://www.gov.br/anvisa/pt-br/assuntos/farmacopeia/farmacopeia-brasileira>. Acesso em: 18 out. 2022.

BRASIL. **Farmacopeia Brasileira 6ª Edição: Volume 2 - Monografias. Plantas Mediciniais**. Agência Nacional de Vigilância Sanitária: ANVISA. Brasília. 2019. Disponível em: <https://www.gov.br/anvisa/pt-br/assuntos/farmacopeia/farmacopeia-brasileira/insumos-farmaceuticos-e-especialidades.pdf/@@download/file/INSUMOS%20FARMAC%3%8AUTICOS%20E%20ESPECIALIDADES.pdf>. Acesso em: 18 out. 2022.

BRASIL. **Hipertensão (pressão alta)**. 2023.

BRASIL. Ministério da Saúde. Portaria nº 533, de 28 de março de 2012. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 28 mar. 2012.

BRASIL. Ministério da Saúde. Portaria nº 971, de 03 de maio de 2006. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 03 maio 2006.

BRASIL. 2014a. Ministério da Saúde. Resolução-RDC nº 26, de 13 de maio de 2014. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 13 maio 2014

BRASIL. 2014b. Ministério da Saúde. Instrução Normativa nº 2, de 13 de maio de 2014. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, 13 maio 2014.

BRASIL. 2022a. **FITOTERÁPICOS**. Agência Nacional de Vigilância Sanitária: ANVISA. Brasília. Disponível em: <https://www.gov.br/anvisa/pt-br/centraisdeconteudo/publicacoes/medicamentos/publicacoes-sobre-medicamentos/folder-sobre-fitoterapicos.pdf>. Acesso em: 26 set. 2024.

BRASIL. 2022b Ministério da Saúde. **Relação Nacional de Medicamentos Essenciais 2022**. Brasília.

BREIMAN, L. **Random Forests**. 2001.

BRUYNE, S.; SPEECKAERT, M. M.; DELANGHE, J. R. Applications of Mid-Infrared Spectroscopy in the Clinical Laboratory Setting. **Critical Reviews in Clinical Laboratory Sciences**. v. 55, n. 1, 2008. p. 21.

CAPUANO, E.; VAN RUTH, S. M. **Infrared Spectroscopy: Applications**. 1º ed. Elsevier Ltd., 2015.

CASALE, E. G.; GOLANN, D. W.; LEMASTER, E. Chapter Seven - U.S. school principals and special education legal knowledge: A scoping review. In: HODAPP, R. M.; FIDLER, D. J. **International Review of Research in Developmental Disabilities**. v. 60. 2021. cap. 7. p. 213-258.

CASTRO, L. O. DE; RAMOS, R. L. D. Descrição botânica, cultivo e uso de *Maytenus ilicifolia* Mart. ex Reiss. cancorosa ou espinheira-santa. Potyó Alegre: FEPAGRO, 2003. 12p. (Circular Técnica,20). , 2003.

CASTRO, P. A. A. **Chemometrics of Biological Tissues Using Infrared Spectroscopy**. 2022. 56 f. Tese (Doutorado em Tecnologia Nuclear) - Universidade de São Paulo. São Paulo. 2022.

CATALOGUE OF LIFE CHECKLIST. *Monteverdia ilicifolia* (Mart. Ex Reissek) Biral. 2017. Disponível em: <<https://www.gbif.org/species/9798870>> Acesso em: 30 jun. 2024.

CHICCO, D.; JURMAN, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. **BMC Genomics**, v. 21, n. 1, 2 jan. 2020.

CHICCO, D.; WARRENS, M. J.; JURMAN, G. The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's *Kappa* and Brier Score in Binary Classification Assessment. **IEEE Access**, v. 9, p. 78368–78381, 2021.

CHU, K. An Introduction to Sensitivity, Specificity, Predictive Values and Likelihood Ratios. **Emergency Medicine**, v. 11, p. 175-181. 1999.

CLANCY, S. Genetic Mutation. **Nature Education**, p. 187–187. 2008.

CORMANICH, R. A.; FREITAS, M. P.; RITTNER, R. **2D Chemical Drawings Correlate to Bioactivities: MIA-QSAR Modelling of Antimalarial Activities of 2,5-Diaminobenzophenone Derivatives. Article J. Braz. Chem. Soc.** V.22, ed. 4. 2011.

DANIEL, B. **Maytenus ilicifolia Mart. Ex Reissek: Useful plants.** 2021. Disponível em: <<https://identify.plantnet.org/useful/observations/1012030483>>. Acesso em: 30 jun. 2024.

DAUSSIN, F.; HEYMAN, E.; BURELLE, Y. Effects of (-) – Epicatechin on Mitochondria. **Nutrition Reviews**. ed. 1 v. 79. 2021. p. 25-41.

DETTKE, G.A. *Jodina rhombifolia* (Hook. & Arn.) Reissek in **Flora e Funga do Brasil**. Jardim Botânico do Rio de Janeiro. Disponível em: <<https://floradobrasil.jbrj.gov.br/FB14412>>. Acesso em: 27 set. 2024

DUARTE, M.R.; DEBUR, M.C. Stem and Leaf Morphoanatomy of *Monteverdia ilicifolia*. **Fitoterapia**, ed. 1. v. 76. jan. 2005. p. 41-49.

DUCAT, G.; *et al.* Correlation among metallic ions, phenolic compounds and antimicrobial action in medicinal plants extracts. **J Food Qual**, v. 34, p. 306–314. 2011.

ECKER, A.; *et al.* Safety evaluation of suprathereapeutic dose of *Maytenus ilicifolia* Mart. ex Reissek extracts on fertility and neurobehavioral status of male and pregnant rats. **Regulatory Toxicology and Pharmacology**, v.90, p. 160–169. 2017.

ELBARBRY, F.; JONES, G.; UNG, A. Catechin Reduces Blood Pressure in Spontaneously Hypertensive Rats through Modulation of Arachidonic Acid Metabolism. **Molecules**, v. 27, 8432. 2022.

FERREIRA, K. C. V. *et al.* **Benchmarking analysis of deleterious SNP prediction tools on CYP2D6 enzyme.** 2019.

FERREIRA, M. M. C. **Quimiometria: Conceitos, Métodos e Aplicações.** Campinas, SP: Editora da Unicamp, 2015.

FIGUEREDO, C. A.; GURGEL, I. G. D.; GURGEL JUNIOR, G. D. A política nacional de plantas medicinais e fitoterápicos: Construção, perspectivas e desafios. **Physis**, v. 24, n. 2, p. 381–400, 2014.

FILHO, V. C.; ZANCHETT, C. C. C. **Fitoterapia Avançada: Uma Abordagem Química, Biológica e Nutricional**. Porto Alegre: Grupo A Educação, 2020.

FOOD and Drug Administration – FDA. **Understanding Unapproved Use of Approved Drugs “Off Label”**. 2018. Disponível em: <https://www.fda.gov/patients/learn-about-expanded-access-and-other-treatment-options/understanding-unapproved-use-approved-drugs-label>. Acesso em: 23 out. 2023.

GINWALA, R.; *et al.* Potential Role of Flavonoids in Treating Chronic Inflammatory Diseases with a Special Focus on the Anti-Inflammatory Activity of Apigenin. **Antioxidants**, v.8, 35. 2019.

GREENWOOD, M. **Intermediate Statistics with R. Montana State University**. 2023.

HASENCLEVER, L.; *et al.* The Brazilian phytotherapeutics industry: Challenges and opportunities. **Ciencia e Saude Coletiva**, v. 22, n. 8, p. 2559–2569, 2017.

HERS, T. **Maytenus ilicifolia Mart ex Reissek: World flora**. 2021 Disponível em: <<https://identify.plantnet.org/the-plant-list/observations/1010505249>>. Acesso em: 30 jun. 2024.

HOLNIK, P.R.; *et al.* Comparação do Teor de Taninos entre Duas Espécies de Espinheira-Santa (*Maytenus aquifolium* Mart. e *Monteverdia ilicifolia* Mart. ex Reissek) Cultivadas no Horto Medicinal do Refúgio Biológico Bela Vista - RBBV da Itaipu Binacional - Foz do Iguaçu, PR - Brasil. **Revista Brasileira de Plantas Mediciniais**. v. 17, n. 3, set. 2015.

HORN, R.C.; VARGAS, V.M.F. Antimutagenic activity of extracts of natural substances in the Salmonella/microsome assay. **Mutagenesis**, v. 18, p. 113–118.

HUMPHREY, A. *et al.* Machine-learning classification of astronomical sources: estimating F1-score in the absence of ground truth. **Monthly Notices of the Royal Astronomical Society: Letters**, v. 517, n. 1, p. L116–L120, 1 nov. 2022.

INSTITUTO DE FÍSICA DE SÃO CARLOS. IFSC/USP. **Infrared Spectroscopy**. 2013. Disponível em: <https://www.ifsc.usp.br/~lavfis2/BancoApostilasImagens/ApLuminescencia/Infrared%20Spectroscopy1.pdf>. Acesso em: 19 out. 2022.

INTERNATIONAL COMMISSION ON NON-IONIZING RADIATION PROTECTION. ICNIRP. **Infrared Radiation (780 nm-1mm)**. 2013. Disponível em: <https://www.icnirp.org/en/frequencies/infrared/index.html>. Acesso em: 20 out. 2022.

JACKSON, R. S. **Wine Science: Principles and Applications**. ed. 3. Academic Press. 2008.

JANIESCH, C.; ZSCHECH, P.; HEINRICH, K. Machine learning and deep learning.

PERIS-DÍAZ, M. D.; KREŽEL, A. **A guide to good practice in chemometric methods for vibrational spectroscopy, electrochemistry, and hyphenated mass spectrometry.** *TrAC - Trends in Analytical Chemistry*. Elsevier B. v. 1 fev. 2021.

JORGE, R. M.; *et al.* Evaluation of antinociceptive, anti-inflammatory and antiulcerogenic activities of *Maytenus ilicifolia*. **Journal of Ethnopharmacology**, v. 94, n. 1, p. 93–100, 2004.

JÚNIOR, R. F. DE A.; *et al.* *Maytenus ilicifolia* dry extract protects normal cells, induces apoptosis and regulates Bcl-2 in human cancer cells. **Experimental Biology and Medicine**, v. 238, n. 11, p. 1251–1258, 2013.

KASSE, C. A.; *et al.* O uso de *Maytenus ilicifolia* na prevenção da ototoxicidade induzida pela cisplatina. **Brazilian Journal of Otorhinolaryngology**, v. 74, n. 5, p. 712–717, 2008.

KHURANA, P; PULICHARLA, R.; BRAR, S. K. Antibiotic-metal complexes in wastewaters: fate and treatment trajectory. **Environ Int**, v.157, 106863. 2021.

KOHN, L.K.; *et al.* In vitro antiviral activity of Brazilian plants (*Maytenus ilicifolia* and *Aniba rosaeodora*) against bovine herpesvirus type 5 and avian metapneumovirus. **Pharm Biol**, v. 50, p. 1269–1275. 2012.

KOWALCZUK, A. P.; *et al.* Application of chemometrics for identification of psychoactive plants. **Acta Poloniae Pharmaceutica - Drug Research**, v. 72, n. 3, p. 517–525, 2015.

KUMAR, S.; PANDEY, A. K. Chemistry and Biological Activities of Flavonoids: An Overview. **The Scientific World Journal**. v. 2013, 2013, p. 16.

LARSON, A.J.; SYMONS, J.D.; JALILI, T. Therapeutic Potential of Quercetin to Decrease Blood Pressure: Review of Efficacy and Mechanisms. **Advances in Nutrition**, v.3, p. 39–46. 2012.

LAVINE, B. K.; RAYENS, W. S. Statistical Discriminant Analysis. **Comprehensive Chemometrics**, v. 3, p. 517–540, 2009.

LEME, T. S. V.; *et al.* Role of prostaglandin/cAMP pathway in the diuretic and hypotensive effects of purified fraction of *Maytenus ilicifolia* Mart ex Reissek (Celastraceae). **J Ethnopharmacol**, v. 150, p. 154–161. 2013.

LIMA, M. D. DE; BARBOSA, R. Methods of Authentication of Food Grown in Organic and Conventional Systems Using Chemometrics and Data Mining Algorithms: a Review. **Food Analytical Methods**. ed. 4. v. 12. abr. 2019. p. 887-901.

LINDON, J.; TRANTER, G. E.; KOPPENAAL, D. **Encyclopedia of Spectroscopy and Spectrometry**. ed. 3. Estados Unidos: Academic Press. 2016.

LOPES, D. C.; STEIDLE NETO, A. J. Classification and Authentication of Plants by Chemometric Analysis of Spectral Data. **Comprehensive Analytical Chemistry**, v.

80, p. 105–125, 2018.

LOPES, G. C.; *et al.* Development and validation of an HPLC method for the determination of epicatechin in *Maytenus ilicifolia* (Schrad.) Planch., Celastraceae. **Revista Brasileira de Farmacognosia**, v. 20, n. 5, p. 789–795, 2010.

MAK, S.; THOMAS, A. Steps for Conducting a Scoping Review. **Journal of graduate medical education**, v. 14, n. 5, p. 565–567, 2022.

MANSANO, V.F.; TIerno, L.R. *Zollernia ilicifolia* (Brongn.) Vogel in **Flora e Funga do Brasil**. Jardim Botânico do Rio de Janeiro. Disponível em: <<https://floradobrasil.jbrj.gov.br/FB29915>>. Acesso em: 27 set. 2024

MARIANO, L.N.B.; *et al.* Preclinical evaluation of the diuretic and saluretic effects of (-)-epicatechin and the result of its combination with standard diuretics. **Biomedicine & Pharmacotherapy**, v.107, p. 520–525. 2018.

MARIOT, M.P.; BARBIERI, R. L.; **Secondary metabolites and medicinal properties of espinheira-santa (*Maytenus ilicifolia* Mart. ex Reiss. e *Maytenus aquifolium* Mart.)**. V.9, ed. 3, p. 83-89. 2007.

MARTINS, M. F. **Caracterização Fitoquímica dos Compostos Orgânicos da Espécie Vegetal *Maytenus ilicifolia* Mart. ex Reissek por Espectroscopia no Infravermelho com Transformada de Fourier e Espectroscopia Fotoacústica**. Dissertação (Mestrado em Ciências da Saúde) - Universidade Estadual de Ponta Grossa. Ponta Grossa, p.82. 2020.

MIRANDA, R. S.; *et al.* Antiinflammatory activity of natural triterpenes—An overview from 2006 to 2021. **Phytotherapy Research**, v. 36, p. 1459–1506. 2022.

MODZELEWSKA, B.; *et al.* Quercetin relaxes human gastric smooth muscles directly through ATP-sensitive potassium channels and not depending on the nitric oxide pathway. **Neurogastroenterology & Motility**, v. 33. 2021.

MONTANARI, T.; DE CARVALHO, J. E.; DOLDER, H. Effect of *Maytenus ilicifolia* Mart. ex. Reiss on spermatogenesis. **Contraception**, v. 57, n. 5 SUPPL., p. 335–339, 1998.

MONTEIRO, S. C.; BRANDELLI, C. L. C. **Farmacobotânica: Aspectos Teóricos e Aplicação**. Porto Alegre: Artmed, 2017.

MOURA, F. B. R. DE, *et al.*; Wound Healing Activity of the Hydroethanolic Extract of the Leaves of *Monteverdia ilicifolia* Mart Ex. Reis. **Journal of Traditional and Complementary Medicine**. ed. 5. v. 11. sep. 2021. p. 446- 456.

NAKAMURA, M.; *et al.* **Anti-ulcerative drug**. 1997.

NAKAMURA, M.; *et al.* **Analgesic anti-inflammatory drug**. 1997.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION. NASA. **Infrared Waves**. 2010. Disponível em: [https://science.nasa.gov/ems/07\\_infraredwaves](https://science.nasa.gov/ems/07_infraredwaves). Acesso em: 20 out. 2022.

NICHOLS, J. A.; HERBERT CHAN, H. W.; BAKER, M. A. B. **Machine learning: applications of artificial intelligence to imaging and diagnosis**. **Biophysical Reviews**. Springer, v. 7. fev. 2019.

NIERO, R.; *et al.* A Comparative Chemical Study of *Maytenus ilicifolia* Mart. Reiss and *Maytenus robusta* Reiss (Celastraceae). **Zeitschrift fur Naturforschung C**. v. 56, fev. 2001, p. 158-161.

NOSSACK, A.C.; *et al.* Quantitative analysis of triterpenes friedelan-3-b-ol and friedelina in *Maytenus aquifolium* by high-resolution gas chromatography and high-temperature capillary gas chromatography. **Phytochemical Analysis**, v. 11, p. 243–246. 2000.

NUNEZ, M. J.; *et al.* Salvadoran Celastraceae Species as a Source of Antikinetoplastid Quinonemethide Triterpenoids. **Plants**, v. 13, ed. 3, 2024.

OKUDA, T.; ITO, H. Tannins of Constant Structure in Medical and Food Plants- Hydrolyzable Tannins and Polyphenols Related Tannins. **Molecules**. v. 16, 2011, p. 2191-2217.

OLIVEIRA, A. D.; *et al.* Use of Infrared Spectroscopy and Near Infrared Hyperspectral Images to Evaluate Effects of Different Chemical Agents on PET Bottle Surface. **Materials Research**. ed. 5. v. 21. 2018.

OLIVEIRA, M.G.M.; *et al.* Pharmacologic and toxicologic effects of two *Maytenus* species in laboratory animals. **J Ethnopharmacol**, v. 34, p. 29–41.1991.

PAGE, M. L. Mutation-inducing drugs could raise virus dangers. **New Sci**, v. 11, p. 255. 2022.

PANIKUTTIRA, B.; O'DONNELL, C. P. **Process Analytical Technology for the Fruit Juice Industry**. Elsevier Inc., 2018.

PASQUINI, C. Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications. **Journal of the Brazilian Chemical Society**. v. 14, n. 2, Abr. 2003.

PAVIA, D. L. *et al.* **Introdução a Espectroscopia**. ed. 2. São Paulo: Cengage Learning, 2015.

PEDERNEIRAS, L.C.; RIBEIRO, J.E.L.S.; MATTOS, L., *Sorocea bonplandii* (Baill.) W.C.Burger *et al.* **in Flora e Funga do Brasil**. Jardim Botânico do Rio de Janeiro. Disponível em: <<https://floradobrasil.jbrj.gov.br/FB31440>>. Acesso em: 27 set. 2024

PENN STATE. **STAT 462: Applied Regression Analysis**. 2018 Disponível em: <<https://online.stat.psu.edu/stat462/>>. Acesso em: 25 out. 2023.

PEREIRA, A.M.S.; *et al.* Evaluation of *Maytenus aquifolia* Mart. and *Monteverdia ilicifolia* Mart. Chemotypes for Tannins, Total Phenols and Triterpenes. **Revista Brasileira de Plantas Mediciniais**. v. 8, n. 1, 2005, p. 13-17.

PERIS-DÍAZ, M. D.; KREŽEL, A. **A guide to good practice in chemometric methods for vibrational spectroscopy, electrochemistry, and hyphenated mass spectrometry**. **TrAC - Trends in Analytical Chemistry**. Elsevier. 1 fev. 2021.

PETERS, M.D.J. *et al.* Guidance for conducting systematic scoping reviews. **Int J Evid Based Health**. V.13, p.141–146. 2015.

PIZZINO, G.; *et al.* Oxidative Stress: Harms and Benefits for Human Health. **Oxid Med Cell Longev**, p. 1–13. 2017.

POETSCH, A.R. The genomics of oxidative DNA damage, repair, and resulting mutagenesis. **Comput Struct Biotechnol J**, v. 18, p. 207–219. 2020.

POWERS, D.M.W. **Evaluation: from precision, recall and f-measure to roc, informedness, markedness & correlation**. v. 2, n. 1, p. 37–63, 2011.

PUGLIESE, R.; REGONDI, S.; MARINI, R. **Machine learning-based approach: Global trends, research directions, and regulatory standpoints**. **Data Science and Management**. KeAi Communications Co. 1 dez. 2021.

QUEIROGA, C.L.; *et al.* Evaluation of the antiulcerogenic activity of friedelan-3 $\beta$ -ol and friedelina isolated from *Maytenus ilicifolia* (Celastraceae). **J Ethnopharmacol**, v.72, p. 465–468. 2000.

RAYMUNDO, T. M. *et al.* Genotoxicity of the Medicinal Plant *Maytenus robusta* in Mammalian Cells in Vivo. **Genetics and Molecular Research**. v. 11, n. 3, 2012, p. 2847-2854.

REINHART, M.; *et al.* First-line diuretics versus other classes of antihypertensive drugs for hypertension. **Cochrane Database of Systematic Reviews**. 2023.

RIBEIRO, F. F.; *et al.* Autenticidade de Amostras de *Monteverdia ilicifolia* Mart. ex Reissek Comercializadas em Mercados de São Mateus, ES, Brasil. **Revista Fitos**. v. 15, n. 2, Rio de Janeiro, 2021, p. 144-152.

RÍOS, J.L.; *et al.* **Natural Triterpenoids as Anti-Inflammatory Agents**. p. 93–143. 2000.

ROY, A.; *et al.* Flavonoids a Bioactive Compound from Medicinal Plants and Its Therapeutic Applications. **Biomed Res Int**, p.1-9. 2022.

SALVADOR, M.J.; *et al.* **Formulação tópica a base de extrato seco de espinheira santa (maytenus ilicifolia) e uso**. 2020.

SANTOS, F. **Modern methods for old data: An overview of some robust methods for outliers detection with applications in osteology.** Journal of Archaeological Science: Reports. Elsevier. 1 ago. 2020.

SANTOS, R. L.; *et al.* Análise sobre a fitoterapia como prática integrativa no Sistema Único de Saúde. **Revista Brasileira de Plantas Mediciniais**, v. 13, n. 4, p. 486–491, 2011.

SARIKAYA, S. B. O. Acetylcholinesterase inhibitory potential and antioxidant properties of pyrogallol. **Journal of Enzyme Inhibition and Medicinal Chemistry**, v. 30, n. 5, p. 761–766, 2015.

SARKER, I. H. Machine Learning: Algorithms, Real-World Applications and Research Directions. **SN Computer Science**. Springer. 1 maio 2021.

SARKER, S. D.; NAHAR, L. **Applications of High Performance Liquid Chromatography in the Analysis of Herbal Products.** Elsevier Inc., 2015.

SCHINDLER, M.S.Z.; *et al.* Characterization of the chemical profile and the effects of ethanolic extracts of *Maytenus ilicifolia* Mart. ex Reissek on glucose metabolism in normal hyperglycemic rats. **J Ethnopharmacol**, v. 276, 114173. 2021.

SEEDHER, N.; AGARWAL, P. Effect of metal ions on some pharmacologically relevant interactions involving fluoroquinolone antibiotics. **Dmd**, v. 25, p. 17–24. 2010.

SHARMA, K.; *et al.* Health Effects, Sources, Utilization and Safety of Tannins: a Critical Review. **Toxin Reviews**. Sep. 2019, p. 1-13.

SHIN, M.; PARK, E.; LEE, H. Plant-Inspired Pyrogallol-Containing Functional Materials. **Advanced Functional Materials**, v. 29, n. 43, p. 1–26, 2019.

SHINDE, P. P., SHAH, S. A Review of Machine Learning and Deep Learning Applications. **2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)**. 2018.

SHURVELL, H.F. Spectra-Structure Correlations in the Mid- and Far- Infrared. **Handbook of Vibrational Spectroscopy**. 2006.

SILVA, A. D. N. D. A.; *et al.* Ações farmacológicas e aplicações clínicas da *Maytenus ilicifolia* (Espinheira Santa). **Research, Society and Development**, v. 12, n. 1, p. e4812139393, 2023.

SILVA, C. S.; *et al.* Chemometric approaches for document dating: Handling paper variability. **Analytica Chimica Acta**, v. 1031, p. 28–37, 2018.

SILVA, F.M.A.; *et al.* Flavonoid glycosides and their putative human metabolites as potential inhibitors of the SARS-CoV-2 main protease (Mpro) and RNA-dependent RNA polymerase (RdRp). **Mem Inst Oswaldo Cruz**, v. 115. 2020.

SIMÕES, C. M. O., *et al.* **Farmacognosia: do Produto Natural ao Medicamento**. Porto Alegre: Artmed, 2017.

SINGH, I.; *et al.* **Pharmaceutical Applications of Chemometric Techniques**. **ISRN Analytical Chemistry**. v. 2013, 2013. p. 13.

SOUZA, J.; FREITAS, Z.M.F.; STORPIRTIS, S. Modelos in vitro para determinação da absorção de fármacos e previsão da relação dissolução/absorção. **Revista Brasileira de Ciências Farmacêuticas**, v. 43, p. 515–527. 2007.

STEFANO, R.D. *Citronella gongonha* (Mart.) R.A.Howard **in Flora e Funga do Brasil**. Jardim Botânico do Rio de Janeiro. Disponível em: <<https://floradobrasil.jbrj.gov.br/FB6678>>. Acesso em: 27 set. 2024.

SUBRAMANIAN, A.; PRABHAKAR, V.; RODRIGUEZ-SAONA, L. **Analytical Methods: Infrared Spectroscopy in Dairy Analysis**. Elsevier, 2016.

SUTRADHAR, I.; *et al.* **Metal Ions and their Effects on Antimicrobial Resistance Development in Wastewater**. ed. 1, v. 2.

TABACH, R.; DUARTE-ALMEIDA, J.M.; CARLINI, E.A. Pharmacological and Toxicological Study of *Maytenus ilicifolia* Leaf Extract Part II-Clinical Study (Phase I), **Phytotherapy Research**, v. 31, p. 921–926. 2017.

TRICCO, A.C. *et al.* PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation, **Ann Intern Med**. v.169, p.467–473. 2018.

UFSC. Universidade Federal de Santa Catarina. **Monteverdia ilicifolia (Mart.) Biral**. 2020. Disponível em: <[https://floradigital.ufsc.br/open\\_sp.php?id=7935](https://floradigital.ufsc.br/open_sp.php?id=7935)>. Acesso em: 26 set. 2024.

ULLAH, A.; *et al.* Important flavonoids and their role as a therapeutic agent. **Molecules**, v.25. 2020.

VÉKEY, K.; TELEKES, A.; VERTES, A. **Medical Applications of Mass Spectrometry**. Holanda: Elsevier Science. 2008.

VELOSO, C. C.; *et al.* Pharmacological Potential of *Maytenus* Species and Isolated Constituents, Especially Tingenone, for Treatment of Painful Inflammatory Diseases. **Revista Brasileira de Farmacognosia**. v. 27, 2017, p. 533-540.

VEERASINGAM, S. *et al.* Contributions of Fourier Transform Infrared Spectroscopy in Microplastic Pollution Research: A Review. **Critical Reviews in Environmental Science and Technology**. ed. 22. v. 51. 2021.

VIEIRA, T. *et al.* Chemometric Approach Using ComDim and PLS-DA for Discrimination and Classification of Commercial Yerba Mate (*Ilex paraguariensis* St. Hil.). **Food Analytical Methods**. 2019.

WESTAD, F.; MARINI, F. **Validation of chemometric models - A tutorial**. **Analytica Chimica Acta**. Elsevier B. 17 set. 2015.

WONFOR, R.; *et al.* Anti-inflammatory properties of an extract of *M. ilicifolia* in the human intestinal epithelial Caco-2 cell line. **J Ethnopharmacol**, v. 209, p. 283–287. 2017.

ZHANG, J.; ONAKPOYA, I.J.; POSADZKI, P.; EDDOUKS, M. **The Safety of Herbal Medicine: From Prejudice to Evidence, Evidence-Based Complementary and Alternative Medicine**. p. 1–3. 2015.

ZHAO, Y. *et al.* Discrimination of *Gentiana rigescens* from Different Origins by Fourier Transform Infrared Spectroscopy Combined with Chemometric Methods. **Journal of AOAC International**, v. 98, n.1. 2015.

ZHU, Q. On the performance of Matthews correlation coefficient (MCC) for imbalanced dataset. **Pattern Recognition Letters**, v. 136, p. 71–80, 1 ago. 2020.