



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CENTRO DE CIÊNCIAS**  
**DEPARTAMENTO DE QUÍMICA ANALÍTICA E FÍSICO-QUÍMICA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**TAMYRIS DE AQUINO GONDIM**

**CARACTERIZAÇÃO QUÍMICA DE RESÍDUOS DO CULTIVO DE ABACAXI E DE  
SEMENTES DE GUARANÁ: CONTRIBUIÇÃO PARA MAXIMIZAR O  
APROVEITAMENTO DE BIOMASSAS NA AGROINDÚSTRIA BRASILEIRA**

**FORTALEZA**

**2024**

TAMYRIS DE AQUINO GONDIM

CARACTERIZAÇÃO QUÍMICA DE RESÍDUOS DO CULTIVO DE ABACAXI E DE  
SEMENTES DE GUARANÁ: CONTRIBUIÇÃO PARA MAXIMIZAR O  
APROVEITAMENTO DE BIOMASSAS NA AGROINDÚSTRIA BRASILEIRA

Tese apresentada ao Programa de Pós-Graduação em Química da Universidade Federal do Ceará, como parte dos requisitos para obtenção do título de Doutora em Química. Área de concentração: Química Analítica.

Orientador: Profa. Dra. Gisele Simone Lopes.

Coorientador: Pesq. Dr. Guilherme Julião Zocolo.

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

G635c Gondim, Tamyris de Aquino.

Caracterização química de resíduos do cultivo de abacaxi e de sementes de guaraná: contribuição para maximizar o aproveitamento de biomassas na agroindústria brasileira / Tamyris de Aquino Gondim. – 2024.

141 f. : il. color.

Tese (doutorado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Química, Fortaleza, 2024.

Orientação: Profa. Dra. Gisele Simone Lopes.

Coorientação: Prof. Dr. Guilherme Julião Zocolo.

1. Metabolômica. 2. Ananas comosus. 3. Paullinia cupana. 4. Cromatografia líquida. 5. Análise multivariada. I. Título.

TAMYRIS DE AQUINO GONDIM

CARACTERIZAÇÃO QUÍMICA DE RESÍDUOS DO CULTIVO DE ABACAXI E DE  
SEMENTES DE GUARANÁ: CONTRIBUIÇÃO PARA MAXIMIZAR O  
APROVEITAMENTO DE BIOMASSAS NA AGROINDÚSTRIA BRASILEIRA

Tese apresentada ao Programa de Pós-Graduação em Químicas da Universidade Federal do Ceará, como parte dos requisitos para obtenção do título de Doutora em Química. Área de concentração: Química Analítica.

Aprovada em: 26/01/2024.

BANCA EXAMINADORA

---

Dra. Gisele Simone Lopes (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Dr. Guilherme Julião Zocolo (Coorientador)  
Empresa Brasileira de Pesquisa Agropecuária (Embrapa)

---

Dr. André Henrique Barbosa de Oliveira  
Universidade Federal do Ceará (UFC)

---

Dr. Alberto Jose Cavalheiro  
Universidade Estadual Paulista (UNESP)

---

Dra. Lívia Paulia Dias Ribeiro  
Universidade da Integração Internacional da Lusofonia Afro-Brasileira (UNILAB)

---

Dra. Tigressa Helena Soares Rodrigues  
Universidade Estadual Vale do Acaraú (UVA)

Aos meus pais, Vera e Gondim pelo apoio incondicional.

## AGRADECIMENTOS

A Deus, pela vida, saúde e a Nossa Senhora pela proteção.

Aos meus pais, Vera e Gondim, pela educação, amor, por sempre me ajudarem ao longo dessa jornada, pelo respeito às minhas escolhas e por estarem sempre presentes nos bons e maus momentos.

Ao meu irmão, Júnior, pela ajuda, apoio e torcida e meu sobrinho, João Pedro.

Ao meu namorado, Jhonyson pelo amor, por estar ao meu lado em todos os momentos e por ter sido meu maior apoiador e motivador ao longo dessa jornada, pela compreensão na ausência e por me incentivar sempre. Além dos conhecimentos compartilhados, me ajudou com os resultados desse trabalho sem você não seria possível. Obrigada!

À Dra. Gisele Simone Lopes, minha orientadora, pela amizade, pela compreensão, pelos conselhos e disposição durante todos os anos de orientação, pelo conhecimento e experiência generosamente compartilhadas. Todos esses atributos constituíram para mim um grande exemplo profissional a ser seguido, como professora, orientadora e pesquisadora.

Ao Dr. Guilherme Julião Zocolo, meu coorientador, pela oportunidade, acolhimento e por acreditar no meu potencial, pela disponibilidade de sempre compartilhar valiosos conhecimentos, apoio incondicional e grandiosos incentivos. Além disso, agradeço por ter sido um exemplo inspirador de humanidade, profissionalismo e pesquisador.

À profa. Dra. Wladiana, pelo acolhimento, conselhos e conhecimentos compartilhados deste o Mestrado.

Ao Dr. Alberto, Dra. Livia, Dra. Tigressa e ao Dr. André pela disponibilidade de participar da banca examinadora, e pelas valiosas contribuições para melhoria deste trabalho.

Aos meus companheiros de LEQA, que de alguma forma colaboraram para realização deste trabalho, pela amizade, pelas discussões científicas, incentivo, descontração e pela agradável convivência. Em especial Victor, Jane, Renato, Nandressa, Fabia e Luan.

Aos meus amigos do Sukita, pelas importantes horas de descontração, caminhadas e jogatinas.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Ao CNPq, por subsidiar o projeto de pesquisa.

À Embrapa Agroindústria Tropical, pelo suporte e apoio na realização das atividades de pesquisa deste trabalho, em especial, aos membros do LMQPN, que contribuíram e sempre se prontificaram a ajudar ativamente para a execução desse trabalho.

À Universidade Federal do Ceará (UFC), onde passei boa parte do tempo durante o curso, sendo a mesma como uma segunda casa, onde adquiri bastante conhecimento apesar das enormes dificuldades. Além de todos os professores que de alguma forma contribuíram para minha formação.

Coração bobo, coração bola,  
Coração balão, coração São João.  
A gente se ilude dizendo  
Já não há mais coração  
(Alceu Valença)



## RESUMO

A necessidade do agronegócio de suprir a crescente demanda por alimento resulta na pressão para aumentar e maximizar sua produtividade paralelo ao desenvolvimento de novos produtos. Nesses aspectos, fatores como a sustentabilidade e a seleção de melhores clones com diferentes finalidades específicas são de grande relevância. A sustentabilidade é um quesito de suma importância, principalmente visando o reaproveitamento das milhões de toneladas de resíduos agroindustriais gerados continuamente durante a produção agrícola. Entretanto, para que os resíduos sejam reaproveitados de maneira eficiente, é imprescindível que sejam determinadas as características químicas e biológicas dos resíduos. Nesse contexto, este trabalho tem como objetivo estabelecer os perfis metabólicos e minerais de folhas de abacaxizeiro (*Ananas comosus*) de sete variedades comerciais de abacaxi e uma avaliação de citotoxicidade. Os perfis metabólicos das folhas abacaxizeiro foram estabelecidos usando UPLC-QTOF-MS<sup>E</sup>, onde vinte e oito metabólitos foram anotados. Por outro lado, os minerais foram avaliados através do ICP-OES, onde foi possível a determinação e quantificação de Zn, Cr, Cd, Mn, P e Fe. Na análise quimiométrica concebida por meio da PCA e da HCA, foi possível constatar que as folhas de abacaxizeiro possuem semelhanças e diferenças em relação a sua composição química. Os extratos hidroetanólicos das folhas de abacaxizeiro avaliados apresentaram baixos níveis de citotoxicidade. Esse fato aliado a composição química que foi determinada pode corroborar para a prospecção de novos usos e aplicações desse coproduto em diferentes tipos de indústrias, como farmacológica, cosmética e alimentícia. Além dos abacaxizeiros, estudaram-se as sementes de diferentes clones de guaraná (*Paullinia cupana*) atribuídos a elevados teores de cafeína em sua composição. Sendo assim, foram avaliadas as sementes de cinquenta e seis clones de guaranazeiro distintos. Com isso, o perfil metabólico das sementes de guaraná foi traçado através do UPLC-QTOF-MS<sup>E</sup>, o qual levou à anotação de dezenove metabólitos especializados. Além disso, também foi utilizada metabolômica direcionada, levando à identificação e quantificação de metabólitos por RMN. Também foi concebida uma análise multivariada, elucidando as semelhanças e diferenças entre as sementes de guaraná avaliadas, principalmente no que diz respeito aos níveis de concentração dos metabólitos. Assim, conclui-se que avaliar e determinar as especificidades metabólicas de diferentes clones de guaraná permite sua aplicação

no desenvolvimento de produtos com diferentes teores de metabólitos específicos, como a cafeína. Isso pode atender a diferentes propósitos na indústria alimentícia, como por exemplo o desenvolvimento de produtos com baixo ou alto teor de cafeína. De modo geral, podemos verificar que as folhas de abacaxizeiro e as sementes de guaraná podem se estabelecer, respectivamente como coproduto e um produto de interesse agroindustrial, seja como uma fonte de compostos bioativos ou no desenvolvimento de novos produtos alimentícios.

**Palavras-chave:** metabolômica; *Ananas comosus*; *Paullinia cupana*; cromatografia líquida; análise multivariada.

## ABSTRACT

The need for agribusiness to meet the growing demand for food results in pressure to increase and maximize its productivity alongside the development of new products. In these aspects, factors such as sustainability and the selection of the best clones with different specific purposes are of great relevance. Sustainability is an extremely important issue, mainly aimed at reusing the millions of tons of agro-industrial waste generated continuously during agricultural production. However, for waste to be reused efficiently, it is essential that the chemical and biological characteristics of the waste are determined. In this context, this work aims to establish the metabolic and mineral profiles of pineapple leaves (*Ananas comosus*) of seven commercial pineapple varieties and an assessment of cytotoxicity. The metabolic profiles of pineapple leaves were established using UPLC-QTOF-MS<sup>E</sup>, where twenty-eight metabolites were noted. On the other hand, the minerals were evaluated using ICP-OES, where it was possible to determine and quantify Zn, Cr, Cd, Mn, P and Fe. In the chemometric analysis designed using PCA and HCA, it was possible to verify that Pineapple leaves have similarities and differences in relation to their chemical composition. The hydroethanolic extracts of pineapple leaves evaluated showed low levels of cytotoxicity. This fact combined with the chemical composition that was determined can support the prospect of new uses and applications of this co-product in different types of industries, such as pharmacology, cosmetics and food. In addition, seeds of different clones of guaraná (*Paullinia cupana*) were studied, attributed to high levels of caffeine in their composition. Therefore, the seeds of fifty-six different guarana clones were evaluated. With this, the metabolomic profile of guaraná seeds was traced through UPLC-QTOF-MS<sup>E</sup>, which led to the annotation of nineteen specialized metabolites. Furthermore, targeted metabolomics was also used, leading to the identification and quantification of metabolites by NMR. A multivariate analysis was also designed, elucidating the similarities and differences between the guaraná seeds evaluated, mainly with regard to the concentration levels of the metabolites. Thus, it is concluded that evaluating and determining the metabolic specificities of different guaraná clones allows their application in the development of products with different levels of specific metabolites, such as caffeine. This can serve different purposes in the food industry, such as the development of products with low or high caffeine content. In general, we can verify that pineapple leaves and guaraná seeds can be

established, respectively, as a co-product and a product of agro-industrial interest, either as a source of bioactive compounds or in the development of new food products.

**Keywords:** metabolomics; *Ananas comosus*; *Paullinia cupana*; liquid chromatography; multivariate analysis.

## LISTA DE FIGURAS

Figura 1	– Plantação de abacaxizeiros.....	21
Figura 2	– Diferentes cultivares de abacaxi.....	24
Figura 3	– Abacaxi com fusariose.....	25
Figura 4	– Planta do guaranazeiro.....	28
Figura 5	– Diferentes rotas biossintéticas responsáveis pela formação de uma grande variedade de metabólitos especializados.....	31
Figura 6	– Estrutura básica dos flavonoides e das suas classes mais comuns.....	36
Figura 7	– Estrutura química de flavan-3-ols, monômeros e procianidinas.....	38
Figura 8	– Alguns terpenoides formados a partir de unidades de isopreno.....	39
Figura 9	– Etapas e processos gerais envolvidos nas análises metabolômicas	42
Figura 10	– Figure 10. Representative chromatograms (ESI <sup>+</sup> ) of the analysis of the seven commercial varieties of pineapple leaves: (a) BRS Vitória; (b) Gold; (c) BRS Ajubá; (d) Smooth cayenne; (e) BRS Imperial; (f) Pérola; (g) Perolera.....	65
Figura 11	– Figure 11. Chemometrics analysis: (a) scores coordinate system PC1 × PC2; (b) scores coordinate system PC3 × PC4; (c) loadings coordinate system PC1 × PC2; (d) loadings coordinate system PC3 × PC4. In the loading graphs, the signals p1 to p34 represent the peaks referring to the secondary metabolites annotated and described in Table 2.....	74
Figura 12	– Result of the clusters presented as a heat map, showing the variability of metabolites and minerals.....	76
Figura 13	– Fig. 13. Chemical structures of some polyphenols (procyanidins) annotated in guaraná seeds.....	95

Figura 14 – Fig. 14. Representative chromatograms of guaraná seed samples, together with the indication of the specialized metabolites annotated (Table 1): (a) positive ionization mode (ESI <sup>+</sup> ); (b) negative ionization mode (ESI <sup>-</sup> ).....	99
Figura 15 – Fig. 15. Principal component analysis of guaraná seed samples: (a) scores with outliers; (b) Hotelling's $T^2$ from scores with outliers; (c) loadings; (d) scores without outliers; (e) Hotelling's $T^2$ from scores without outliers.....	103
Figura 16 – Fig. 16. Representative <sup>1</sup> H NMR spectrum ( $\delta$ 0.0 to 9.0 ppm) of guarana seed powder (600 MHz, CD <sub>3</sub> OD- <i>d</i> <sub>4</sub> + D <sub>2</sub> O + EDTA).....	105
Figura 17 – Fig. 17. PCA results using the <sup>1</sup> H NMR dataset ( $\delta$ 0.7 and 8.7) of the guaraná seed powder: a) PC1 × PC2 scores coordinate system from the total samples; b) PC1 loadings plotted in lines form of the total samples; c) PC1 × PC2 scores coordinate system excluding the outliers' samples (13, 14, 15 and 17); d) PC1 and PC2 loadings plotted in lines form excluding the outlier samples.....	109
Figura 18 – Fig. 18. Quantification using <sup>1</sup> H NMR: acetic acid (a), alanine (b), caffeine (c), gallic catechin (d), malic acid (e), sucrose (f), and total catechin (g) in guaraná seed powder from different clones.....	112
Figura 19 – Fig. S1. <sup>1</sup> H- <sup>13</sup> C HSQC contour maps (600/150 MHz, CD <sub>3</sub> OD- <i>d</i> <sub>4</sub> +D <sub>2</sub> O + EDTA) of guaraná seeds.....	116
Figura 20 – Fig. S2. <sup>1</sup> H- <sup>1</sup> H COSY contour maps (CD <sub>3</sub> OD- <i>d</i> <sub>4</sub> +D <sub>2</sub> O + EDTA) of guaraná seeds.....	117
Figura 21 – Fig. S3. <sup>1</sup> H- <sup>13</sup> C HMBC contour maps (600/150 MHz, CD <sub>3</sub> OD- <i>d</i> <sub>4</sub> +D <sub>2</sub> O + EDTA) of guaraná seeds.....	118
Figura 22 – Fig. S4. (a) Caffeine molecule - arrows indicate the Key HMBC correlations (H → C) observed in a representative sample of guaraná seeds, (b) expansion of PRESAT <sup>1</sup> H NMR spectrum ( $\delta$ 3.30 to 3.90 ppm) and (c) Expansion of <sup>1</sup> H- <sup>13</sup> C HSQC contour map selected region ( $\delta$ 3.1 to 4.4 ppm).....	119

Figura 23 – Fig. S5. Expansion of the representative $^1\text{H}$ NMR spectrum ( $\delta$ 5.70 to 7.10 ppm) highlighting catechin and gallocatechin A and B-ring signals.....	119
Figura 24 – Fig. S6. (a) Numbered catechin and gallocatechin structure - arrows indicate the Key HMBC correlations observed in a representative sample, (b) $^1\text{H}$ NMR spectrum expansion ( $\delta$ 5.70 to 7.10 ppm) and $^1\text{H}$ - $^{13}\text{C}$ HSQC heteronuclear correlation map in the region at $\delta$ 5.80 – 6.30 and 6.60 – 7.10 ppm regions (c and d, respectively). In the quantification step, the signals in ring B $\delta$ 6.86 and 6.99 ppm for I and II, respectively, were considered.....	120
Figura 25 – Fig. S7. Expansion of the representative $^1\text{H}$ - $^{13}\text{C}$ HSQC contour map of guaraná seeds ( $\delta$ 2.1 – 3.1 ppm) highlighting the equatorial (eq) and axial (ax) signals of I and II (C ring).....	121
Figura 26 – Fig. S8. A representative $^1\text{H}$ - $^1\text{H}$ COSY contour map of guaraná seeds highlighting specific correlations between caffeine (H-8 and H-7') and sucrose (H-1 and H-2) protons.....	121

## LISTA DE TABELAS

Tabela 1 – Principais vias metabólicas do metabolismo primário.....	32
Tabela 2 – Table 2. Metabolites annotated in the samples of seven varieties of pineapple leaves, in negative (ESI <sup>-</sup> ) and positive (ESI <sup>+</sup> ) modes.....	60
Tabela 3 – Table 3. Quantification of minerals in samples of commercial pineapple leaves.....	70
Tabela 4 – Table 4. Cytotoxic activity of the hydroethanolic extract of pineapple leaves of different commercial varieties determined by MTT assay after 72 h of incubation at a concentration of 100 µg mL <sup>-1</sup> (% inhibition ± SD*).....	80
Tabela 5 – Table 5. Annotation of metabolites in guaraná seeds, positive and negative ionization modes.....	96
Tabela 6 – Table 6. Proton and carbon chemical shifts, coupling constant, multiplicity and long-range heteronuclear <sup>1</sup> H- <sup>13</sup> C HMBC correlations gathered for the twelve identified guaraná metabolites.....	106



## SUMÁRIO

1	INTRODUÇÃO .....	16
2	OBJETIVOS .....	20
2.1	Objetivos específicos .....	21
3	REVISÃO DA LITERATURA .....	21
3.1	Abacaxizeiro.....	21
3.2	Guaranazeiro.....	27
3.3	Metabolismo vegetal.....	30
3.3.1	<i>Metabolismo primário</i> .....	31
3.3.2	<i>Metabólitos especializados</i> .....	33
3.3.2.1	<i>Compostos fenólicos</i> .....	34
3.3.2.2	<i>Terpenoides</i> .....	37
3.3.2.3	<i>Alcaloides</i> .....	40
3.4	Metabolômica.....	40
3.4.1	<i>Coleta e preparo da amostra</i> .....	42
3.4.2	<i>Técnicas analíticas utilizadas em estudos metabólicos</i> .....	45
3.4.3	<i>Análise estatísticas em estudos metabolômicos</i> .....	47
4	AVALIAÇÃO DO PERFIL METABÓLICO, MINERAL E CITOTÓXICO DE FOLHAS DE ABACAXIZEIRO DE DIFERENTES VARIEDADES COMERCIAIS: UMA NOVA FONTE ECOLOGICAMENTE CORRETA E BARATA DE COMPOSTOS BIOATIVOS .....	50
5	ABORDAGENS METABOLÔMICAS PARA EXPLORAR A QUIMIODIVERSIDADE EM SEMENTES DE GUARANÁ ( <i>Paullinia cupana</i> ) UTILIZANDO UPLC-QTOF-MS <sup>E</sup> E RMN.....	81
6	CONCLUSÃO .....	122
	REFERÊNCIAS .....	124

## 1 INTRODUÇÃO

Atualmente, o agronegócio enfrenta crescente pressão para aumentar a produtividade agrícola de forma sustentável, minimizando os impactos negativos ao meio ambiente juntamente com a ampliação de impactos positivos na sociedade e na economia (Vaz Júnior, 2020). Somado a isso, tem-se também a busca crescente por alimentos funcionais, os quais são alimentos não se limitam apenas à função de nutrição padrão, mas que também promovem alguma ação benéfica. Em geral, dentre as possíveis ações, podemos citar o fortalecimento do sistema imunológico e a melhoria da saúde geral do consumidor. Essas melhorias potenciais na saúde podem contribuir para uma redução no risco de doenças, aumentando assim o valor percebido e a demanda por esses produtos no mercado alimentício. Portanto, o agronegócio não apenas enfrenta o desafio de aumentar a produtividade de forma sustentável, mas também deve adaptar-se para atender a essas demandas emergentes por alimentos que proporcionem benefícios adicionais à saúde.

Nesse contexto, o guaraná se destaca como alimento funcional por apresentar excelentes qualidades nutricionais em suas sementes, sendo um estimulante energético devido à cafeína, além de contribuir na prevenção de diversas doenças devido à presença de compostos bioativos, tais como fenólicos, e polifenóis. Seu valor econômico torna a exploração dessa espécie uma atividade altamente rentável, sendo, portanto, de grande importância a obtenção de variedades mais produtivas ou com teores menores / maiores de algum metabolito. Assim, a avaliação dos metabólitos especializados no guaraná é crucial para analisar seu valor nutricional, bem como agregar valor ao guaraná como um alimento funcional e aprimorar a genética das culturas para especificações de plantio, visando a produção de guaraná com melhores características agrícolas, sensoriais e quimiopreventivas (Corbo *et al.*, 2014; Santana; Macedo, 2018).

Como mencionado anteriormente, é fundamental que o agronegócio aumente a produtividade agrícola de maneira sustentável. Nesse sentido, dado que a agricultura tem a responsabilidade principal de atender à crescente demanda populacional por alimentos, também deve concentrar esforços na criação de abordagens e estratégias para mitigar os impactos ambientais associados à produção agrícola. Sendo assim, uma maneira de atenuar os impactos é através da redução na geração de resíduos, por meio do reaproveitamento de todos os subprodutos e/ou

coprodutos gerados durante os processamentos agroindustriais (Vaz Júnior, 2020). Nessas circunstâncias, o abacaxi é uma das frutas tropicais mais populares do mundo, principalmente pelo seu sabor marcante, aroma atrativo e valores nutricionais. Contudo, a extensa produção resulta na geração significativa de milhões de toneladas de resíduos agrícolas. Portanto, é crucial buscar soluções e alternativas para a gestão eficiente desses resíduos, aproveitando-os como matéria-prima para diferentes finalidades, por exemplo como fonte de compostos bioativos, e promovendo estratégias de recuperação baseadas em tecnologias verdes e sustentáveis (Rico *et al.*, 2020).

Estudos sobre a composição química dos resíduos agroindustriais do abacaxi são escassos, especialmente no que diz respeito aos polifenóis bioativos, os quais têm potencial interesse em setores como a indústria farmacêutica, cosmética e medicina. A falta de pesquisa nesse campo ressalta a necessidade de explorar mais profundamente os resíduos gerados na produção de abacaxis, visando o reaproveitamento dos mesmos. Além disso, para maximizar a redução de resíduos, minimizar a poluição ambiental e promover a valorização desses resíduos agroindustriais, é crucial identificar e caracterizar adequadamente os resíduos gerados durante o processamento do abacaxi. A caracterização detalhada desses resíduos não só pode abrir novas oportunidades de utilização, como também contribuir para práticas sustentáveis na cadeia produtiva do abacaxi. Isso inclui o desenvolvimento de novos produtos, tecnologias e processos que possam aproveitar integralmente os recursos disponíveis, alinhando-se assim com as demandas crescentes por sustentabilidade e inovação na indústria agroindustrial.

Sendo assim, uma forma moderna de avaliar os metabólitos de um dado organismo biológico é através da abordagem metabolômica. Nos últimos anos, a metabolômica se consolidou como uma ferramenta essencial na pesquisa e na avaliação de organismos biológicos, visando a obtenção de informações sobre impactos ambientais, compreender as funções genéticas e caracterizar os processos celulares, tudo com o propósito de aprimorar a qualidade de vida humana. Esse campo da ciência envolve um conjunto de estratégias analíticas, que de forma resumida abrange coleta e preparo das amostras, aquisição e tratamento de dados, análises estatísticas e interpretação química / biológica (Funari *et al.*, 2013; Guedes, 2018; Liang *et al.*, 2024). Além disso, a metabolômica se constitui como uma análise abrangente, tanto qualitativa quanto quantitativa, dos metabólitos, com o objetivo

principal de obter a máxima quantidade de informações metabólicas de um organismo ou sistema biológico (Belinato *et al.*, 2019; Ellis *et al.*, 2007).

A avaliação do perfil metabólico é uma das aplicações mais comuns da metabolômica na análise de plantas. Normalmente, essa avaliação é realizada com o propósito de estabelecer uma impressão digital química que represente o estado bioquímico de um organismo ou amostra em um ponto específico no tempo. Esses metabólitos podem ser os produtos intermediários ou finais das vias metabólicas das plantas, além disso, podem ser produzidos mediante à influência de condições ambientais externas. De modo geral, essa abordagem para análise de metabólitos é amplamente utilizada na pesquisa básica de plantas e em diferentes cenários em saúde, agricultura e ciências alimentares (Chin *et al.*, 2006; Gobbo-Neto; Lopes, 2007; Liang *et al.*, 2024; Shen *et al.*, 2023). A abordagem metabolômica também pode ser utilizada para distinguir plantas ou variedades alimentares, explorar as características e padrões evolutivos do acúmulo de metabólitos em diferentes plantas, avaliar a atividade biológica de extratos específicos, identificar biomarcadores ou compostos bioativos com potenciais benefícios para a saúde e construir redes metabólicas reguladoras espaço-temporais e teciduais específicas durante o desenvolvimento da planta, bem como monitorar a mudança de estresse (Gobbo-Neto; Lopes, 2007; Shen *et al.*, 2023).

Com base no exposto, o presente trabalho adota a abordagem metabolômica para uma avaliação detalhada e determinação do perfil metabólico de sementes de guaranáis provenientes de diferentes clones, bem como de folhas de abacaxizeiros de sete variedades comerciais distintas. Esta metodologia permite não apenas explorar a diversidade química presente nos produtos agrícolas, mas também destacar o potencial de aplicação da metabolômica na valorização de produtos e subprodutos da agroindústria frutícola brasileira. Haja vista que a análise metabolômica oferece uma visão abrangente dos compostos químicos presentes nos materiais estudados, revelando informações valiosas sobre sua composição e possíveis aplicações. Ao compreender melhor os perfis metabólicos das sementes de guaraná e das folhas de abacaxizeiros, este estudo contribui para o desenvolvimento de estratégias que visam maximizar o aproveitamento desses recursos naturais. Essa abordagem não apenas promove a sustentabilidade ao reduzir o desperdício e valorizar subprodutos, mas também fomenta a inovação na agroindústria ao identificar novas oportunidades de uso e desenvolvimento de produtos derivados.

Sendo assim, a utilização da metabolômica como ferramenta neste estudo não só ressalta a importância da pesquisa científica na caracterização detalhada dos recursos agroindustriais, mas também reforça seu papel fundamental na promoção de práticas sustentáveis e na ampliação do conhecimento sobre os potenciais benefícios econômicos e ambientais dos produtos da agroindústria frutícola brasileira.

## 2 OBJETIVOS

Realizar estudo metabolômico de sementes de cinquenta e seis clones diferentes de guaraná (*Paullinia cupana* Kunth) e folhas de sete variedades comerciais distintas de abacaxizeiro (*Ananas comosus* (L.) Merril.) para a valorização de produtos e coprodutos / subprodutos / resíduos da agroindústria brasileira.

### 2.1 Objetivos específicos

- Identificar os metabolitos e estabelecer os perfis químicos dos extratos hidroetanólicos das folhas de abacaxizeiro de sete variedades comerciais (Pérola, Smooth Cayenne, Perolera, Gold, BRS Ajubá, BRS Vitória e BRS Imperial), através do UPLC-QTOF-MS<sup>E</sup>;
- Avaliar a composição mineral das variedades comerciais de folhas de abacaxizeiro através do ICP-OES;
- Correlacionar os perfis metabólicos e minerais das diferentes variedades de folhas de abacaxizeiro através ferramentas quimiométricas (PCA e HCA);
- Determinar as variáveis discriminantes, metabolitos ou minerais, que influenciam na diferenciação entre as folhas de abacaxizeiro;
- Avaliar a atividade citotóxica *in vitro* dos extratos hidroetanólicos das variedades de abacaxizeiro frente às linhagens tumorais HL60 (leucêmica), HCT-116 (colo humano), PC3 (próstata), SNB19 (astrocitoma), MCF-7 (mama), B16F10 (melanoma) e HeLa (colo do útero);
- Identificar os metabolitos e determinar os perfis químicos dos extratos hidroetanólicos de sementes de guaraná em cinquenta e seis clones distintos, por meio do UPLC-QTOF-MS<sup>E</sup>;
- Através da análise metabolômica alvo, quantificar os metabolitos nas sementes de guaraná, por meio da RMN;
- Correlacionar os perfis metabólicos de sementes de clones de guaraná através de análise quimiométricas (PCA);

### 3 REVISÃO DA LITERATURA

#### 3.1 Abacaxizeiro

A espécie *Ananas comosus* (L.) Merr., conhecida como abacaxizeiro, Figura 1, é uma frutadeira originária da América Central e América do Sul e pertencente à família Bromeliaceae e tem o seu fruto abacaxi como único representante da sua espécie utilizado como fonte de alimento (Banerjee *et al.*, 2018; Barros *et al.*, 2020; Manetti; Delaporte; Laverde Jr., 2009).

Figura 1 – Plantação de abacaxizeiros.



Fonte: <https://www.embrapa.br/busca-de-imagens/-/midia/1750001/lavoura-de-abacaxi>.

De acordo com o site do Banco de Dados Estatísticos Corporativos da *Food and Agriculture Organization of the United Nations* (FAOSTAT), a produção mundial de abacaxi em 2021 foi de quase 29 milhões de toneladas. Sendo o Brasil o maior produtor, seguido de Tailândia, Filipinas, Costa Rica, Índia e Indonésia, os quais compreende quase 40% da produção mundial (FAOSTAT, 2021). No Brasil, é estimado que sejam cultivados 2,4 bilhões de frutos em 67,2 mil hectares (Barros *et al.*, 2020; FAOSTAT, 2021; Mohd Ali *et al.*, 2020).

Nesse contexto, o fruto abacaxi se constitui como uma das frutas mais populares no mundo. Seu delicioso sabor e aroma são amplamente valorizados, tanto

que é encontrado nos mais variados produtos processados, como doces, xaropes, vinagres, vinhos, licores e etc. Proporcionando uma elevada demanda no mercado global.

A composição nutricional do abacaxi vem sendo cada vez mais estudada, de modo a torná-lo um potencial alimento funcional. Em geral, um abacaxi fresco é constituído de água (80–85%), açúcares (12–15%), ácidos orgânicos (0,6%), proteínas (0,4%), cinzas (0,5%), gorduras (0,1%) e vitaminas (Banerjee *et al.*, 2018; Barros *et al.*, 2020; Manetti; Delaporte; Laverde Jr., 2009).

Desse modo, o abacaxi é considerado mais que uma fruta nutritiva, seus diversos componentes químicos possuem uma série de aplicações medicinais, cosméticas e alimentícias. Compostos voláteis são extraídos da polpa do abacaxi com o objetivo de serem usados como aromatizantes e intensificadores de aroma, bem como para a produção de essências naturais. O abacaxi também tem destaque no campo das ciências medicinais, pois estudos apontam que a fruta possui inúmeras propriedades farmacológicas, como anti-inflamatória, antioxidante, antibacteriana, antidiabética, anticancerígena, antitrombótica, anticoagulante, antimicrobiana e anti-helmínticas por conter importantes compostos bioativos, como flavonoides, polifenóis, taninos, carboidratos, glicosídeos e proteínas. Além disso, os extratos etanólicos das folhas de abacaxi têm mostrado efeitos antidiabético, anti-hiperlipidimêmico e antioxidante (Banerjee *et al.*, 2018; Guedes *et al.*, 2018; Ma *et al.*, 2007; Manetti; Delaporte; Laverde Jr., 2009; Mohd Ali *et al.*, 2020; Rodrigues *et al.*, 2020).

Além da fruta e folhas, o caule do abacaxizeiro também é importante, onde é possível realizar a extração da bromelaína, uma enzima proteolítica, que vem sendo amplamente utilizada para fins medicinais e cosméticos. Essa enzima tem a propriedade de prevenir a agregação plaquetária, um dos maiores fatores que ocasionam o acidente vascular cerebral, bem como tem comprovada eficácia no tratamento do câncer de mama, leucemia, carcinoma pulmonar, melanoma e câncer de ovário. Além disso, ela age como um forte anti-inflamatório, combatendo distúrbios como asma, inflamação do cólon, osteoartrite e artrite reumatoide. Por último, a bromelaína também tem provado ser útil como tratamento para vários distúrbios dermatológicos, como queimaduras e escaras (Banerjee *et al.*, 2018; Mohd Ali *et al.*, 2020; Rodrigues *et al.*, 2020).

Existem muitas cultivares distintas de abacaxi, mais de 100 variedades, nas quais apenas 6 a 8 variedades são de fato cultivadas comercialmente. Com tantas



variedades é de se esperar que haja variações de cores, formas, tamanhos e características sensoriais. Sendo assim, todas as variações mencionadas, somada aos diferentes estágios de maturação da fruta, impõem variações na composição química das diferentes cultivares de abacaxizeiros (Mohd Ali *et al.*, 2020; Reinhardt *et al.*, 2018).

Como relatado previamente, o Brasil é o maior produtor mundial de abacaxis. Portanto, se destaca como um centro de origem e de diversidade da cultura do abacaxizeiro. Nesse contexto, as duas principais cultivares mundiais, Pérola (cultivar tradicional do território brasileiro) e Smooth Cayenne (Figura 2) (também chamada de Havai ou havaiano, a mais utilizada mundialmente para processamento: suco, fatias, etc), são destinadas ao consumo *in natura* e à agroindústria, respectivamente, mas possuem como desvantagem a suscetibilidade à fusariose (Embrapa, 2015; Abacaxi Smooth Cayenne (havaiano)2021; Viana *et al.*, 2013). A fusariose do abacaxi (Figura 3) é causada pelo fungo *Fusarium guttiforme*, a vulnerabilidade a esse patógeno se caracteriza como a principal limitação fitossanitária à produção do abacaxi no Brasil (Fusariose do AbacaxizeiroEmbrapa, 2021).

Por outro lado, a cultivar de abacaxizeiro Gold (ou MD-2) (Figura 2) é a cultivar mais consumida na Europa e América do Norte no mercado para frutos *in natura*, com destaque na produção da América Central, mas também na África e Ásia. Essa cultivar foi desenvolvida no Hawaii (EUA), a qual teve seus frutos disponibilizados comercialmente em 1996. Contudo, assim como as cultivares Pérola e Smooth Cayenne, a cultivar Gold também é suscetível à fusariose, o que limita sua produção no Brasil (Cultivar de abacaxi MD-2 ou GoldEmbrapa, 2021).

Diante disso, o uso de cultivares resistentes se materializa como a melhor forma de controle de doenças. Sendo assim, estudos realizados pela Embrapa Mandioca e Fruticultura, revelaram que a cultivar Perolera (Figura 2) comportou-se como resistente à fusariose. Com isso, essa cultivar é recomendada para plantio em regiões onde a incidência dessa doença é elevada. A cultivar Perolera é uma variedade originária e plantada comercialmente na Colômbia e na Venezuela, adaptada a altitudes de até 1500 m. A planta apresenta folha de cor verde escuro e sem espinhos, evidenciando faixa prateada pouco pronunciada nos bordos. O fruto tem forma cilíndrica, com peso médio de 1,8 kg, de casca e polpa amarelas, acidez moderada e alto teor de ácido ascórbico (vitamina C) (Cabral; Junghans, 2003).

Figura 2 – Diferentes cultivares de abacaxi.



Pérola



Smooth Cayenne



BRS Vitória



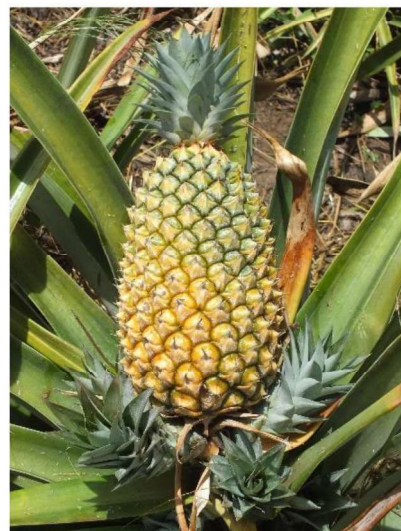
BRS Imperial



Perolera



Gold (MD-2)



BRS Ajubá

Figura 3 – Abacaxi com fusariose.



Fonte: <https://www.embrapa.br/en/busca-de-imagens/>.

Ainda no contexto de cultivares resistentes a doenças, podemos destacar os programas de melhoramento genético, os quais têm como objetivo obter híbridos superiores de abacaxizeiro resistentes à fusariose e com características físico-químicas e sensoriais que agradem aos consumidores da fruta. Esta diversificação de variedades contribui para a segurança alimentar e, conseqüentemente, reduz os custos de produção. O programa de melhoramento genético do abacaxizeiro da Embrapa, criado em 1984, tem como objetivo desenvolver cultivares resistentes à fusariose, na qual as características comerciais sejam iguais ou superiores às da cultivar Pérola e Smooth Cayenne. Como resultado deste programa de melhoramento, foram criadas e lançadas três novas cultivares: BRS Imperial, BRS Vitória e BRS Ajubá, resistentes à fusariose e que apresentam características comerciais iguais ou superiores as cultivares Pérola e Smooth Cayenne (Viana *et al.*, 2013).

A cultivar BRS Imperial (Figura 2) foi desenvolvida pela Embrapa Mandioca e Fruticultura em 2003, a qual é um híbrido resultante do cruzamento entre duas variedades: Perolera e Smooth Cayenne. O BRS Imperial é resistente à fusariose,

principal doença da cultura, gerando frutos doces e de excelente qualidade. A planta tem porte médio e apresenta folha de cor verde-escura, sem espinhos nas bordas. Os frutos são menores do que os da cultivar Pérola, têm formato cilíndrico e casca de cor amarelo-intenso na maturação. A polpa é amarela, com elevado teor de açúcar, acidez moderada, alto teor em ácido ascórbico (antioxidante) e excelente sabor nas análises sensoriais realizadas. Além disso, os frutos tem peso médio com a coroa de 1,2 kg, podendo alcançar 1,5 kg e tamanho médio de 16 cm (Cabral; Junghans, 2003; Embrapa, 2004).

Outra cultivar desenvolvida pela Embrapa Mandioca e Fruticultura, em 2006, foi o abacaxi BRS Vitória (Figura 2). Esta cultivar é um híbrido resultante do cruzamento das variedades Primavera e Smooth Cayenne. O BRS Vitória tem como características gerais: resistência à fusariose, formato cilíndrico, folha de cor verde claro, sem espinhos nas bordas, casca de cor amarela na maturação. Além disso, o fruto tem polpa branca, com elevado teor de açúcares e excelente sabor nas análises químicas e sensoriais, sugerindo que suas características relativas à acidez (0,8%) são superiores às do abacaxi Pérola e Smooth Cayenne. O fruto apresenta uma maior resistência ao transporte e em pós-colheita, o que pode facilitar a sua disseminação entre os produtores e conquistar a preferência dos consumidores. Apresenta peso médio do fruto sem coroa de 1,4 kg (Embrapa, 2006).

A cultivar de abacaxi BRS Ajubá (Figura 2), disponibilizada em 2009, assim como a cultivar BRS Imperial, também é um híbrido, resultante do cruzamento entre as cultivares Perolera e Smooth Cayenne, com resistência à fusariose. Portanto, o plantio desta variedade dispensa a utilização de fungicida para o controle da fusariose. A planta tem porte médio e apresenta folha de cor verde escuro, totalmente desprovida de espinhos. O fruto é cilíndrico, com casca de cor amarela na maturação. A polpa é amarela, com elevado teor de açúcar e acidez moderada. Apresenta peso médio do fruto sem a coroa de 1,3 kg e tamanho médio de 15,8 cm (Embrapa, 2009).

De modo geral, o abacaxizeiro (Figura 1) possui características específicas, como altura entre 75 a 150 cm, alcançando uma largura entre 90 a 120 cm. Suas folhas são longas e pontiagudas, com tamanho variando de 50 a 180 cm de comprimento e espinhos pontiagudos nas bordas. Normalmente, leva de 12 a 14 meses para florescer, que pode depender da variedade, e cerca de 6 a 8 meses para o fruto amadurecer. Gerando cerca de 30 a 50 folhas com um peso médio de 35 g por folha, o que resulta em cerca de 1 a 1,5 kg de folhas por planta deixadas no campo

após o cultivo. As folhas de abacaxi são ricas em fibras e compreendem um alto teor de celulose (75-85%) (Banerjee *et al.*, 2018; Sibaly; Jeetah, 2017).

Sendo assim, o crescimento sucessivo na produção de abacaxi a cada ano, tem como consequência uma maior geração de resíduos sólidos provenientes do plantio de abacaxizeiro (FAOSTAT, 2021). De acordo com uma pesquisa na Índia, estima-se que cada abacaxizeiro produza entre 6 e 8 kg de resíduos, incluindo folhas, caules e raízes. Se plantados 12.000 abacaxizeiro por hectare, a colheita resultará entre 76 a 92 toneladas de resíduos nos campos após a colheita do abacaxi (Banerjee *et al.*, 2018). Portanto, é crucial realizar estudos visando a valorização desses resíduos agroindustriais, os quais podem potencialmente se tornar insumos da indústria farmacêutica, cosmética e alimentícia.

### 3.2 Guaranazeiro

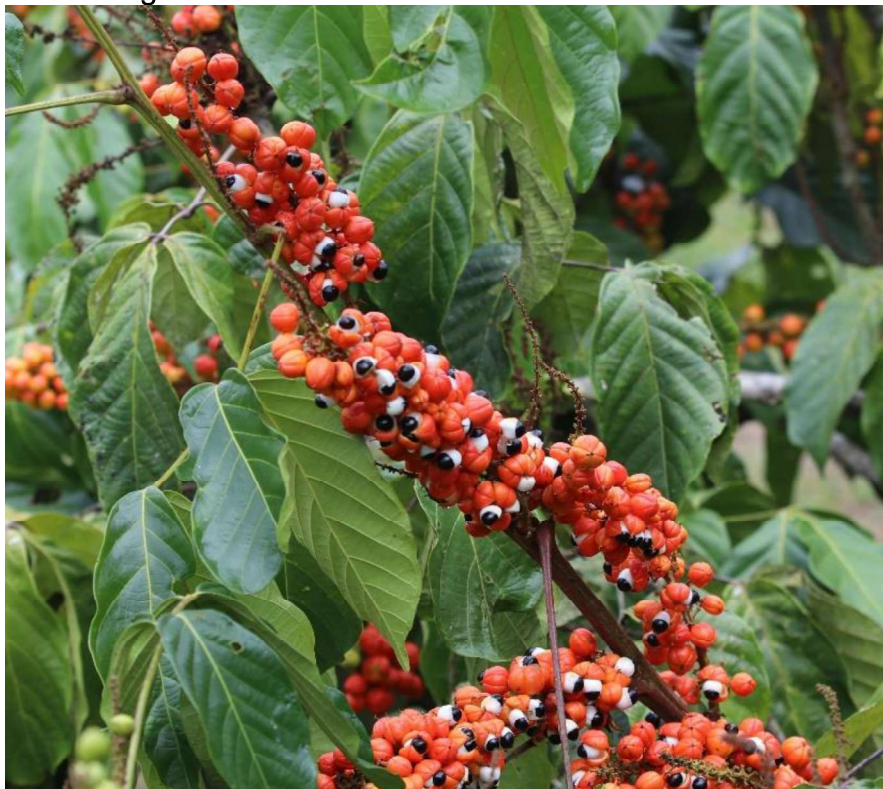
O guaranazeiro (*Paullinia cupana* Kunth), conhecido também como *guaraná-da-amazônia*, *guaranaina*, *guaranauva*, *uarana* ou *narana*, (Figura 4) é uma planta nativa da região amazônica que produz frutos guaraná. Pertencente à família Sapindaceae e à ordem Sapindales, essa planta tem sido utilizada há séculos por tribos indígenas devido às suas propriedades estimulantes, afrodisíacas e curativas para dores de cabeça utilizando suas sementes torradas. O primeiro registro do uso do guaraná como uma bebida ocorreu em 1669, durante uma expedição jesuíta à Amazônia, onde o missionário João Felipe Bettendorf observou que os índios Sateré-Mawé consumiam uma bebida estimulante com propriedades diuréticas e terapêuticas contra cefaleia, febre e cólicas (Marques *et al.*, 2016; Schimpl *et al.*, 2013).

A palavra *guaraná*, *uarana* ou *varana* significa “trepadeira” em vários dialetos indígenas, refere-se ao tipo trepadeira dessa planta perene, que possui gavinhas que podem atingir até 10m de comprimento na presença de árvores que servem de suporte. O caule é estriado e apresenta coloração marrom-amarelada quando lignificado. As folhas são alternadas e ímpares pinadas. As bainhas bem desenvolvidas têm aproximadamente 1,5 cm de comprimento. O pecíolo principal tem de 8 a 19 cm, e os pecíolos dos folíolos são muito curtos. Os folíolos têm formato aproximadamente oval e ápice serrilhado, com largura variando de 10 a 14 cm e comprimento de 27 a 33 cm. Os folhetos são bem espaçados e têm veias inferiores proeminentes. As folhas são verde-escuras com a parte superior brilhante. Na base

de cada folha há uma gema vegetativa e outra reprodutiva (Schimpl *et al.*, 2013).

O guaranazeiro é uma planta monoica com flores pequenas e zigomorfas, dispostas em toda a extensão do seu eixo podendo atingir até 13 m de altura. O cálice é constituído por cinco sépalas, duas menores e externas, e a corola formada por quatro pétalas brancas, que se unem formando um capuz internamente contendo escamas coriáceas de coloração amarelada. As flores são classificadas como pseudo-hermafroditas, contendo caracteres femininos e masculinos, mas os machos possuem estilete e estigma regredidos. Quando madura, o fruto passa de amarelo-laranja a vermelho vivo e seu pedúnculo fica evidente. Possuem entre uma e quatro sementes marrons com um arilo branco, que lembra um olho humano, sendo uma característica marcante para a identificação do guaraná. É encontrada no Brasil, Guiana, Venezuela e Equador (Marques *et al.*, 2019; Patrick *et al.*, 2019; Schimpl *et al.*, 2013).

Figura 4 – Planta do guaranazeiro.



Fonte: <https://www.embrapa.br/busca-de-imagens/-/midia/6350001/cultivar-de-guaranazeiro-brs-nocoquem>.

O Brasil é praticamente o único produtor de guaraná no mundo. Mais especificamente, o guaraná é encontrado principalmente na região sudeste do estado do Amazonas, nos municípios de Maués e Parintins. Pequenas áreas da Amazônia

venezuelana também possuem cultivos de guaraná. Nas últimas décadas, o cultivo dessa planta tem sido incentivado em outras áreas do Brasil, como nos vales dos rios Purus e Tapajós (no Amazonas), nos estados do Pará, Acre e Rondônia, na região cacaueira da Bahia (entre as cidades de Salvador e Ilhéus), no Vale do Ribeira (no estado de São Paulo) e na região de Alta Floresta, no Mato Grosso (Marques *et al.*, 2019).

Uma parcela do guaraná fabricado no Brasil é consumida pelos habitantes da Amazônia. O pó, normalmente extraído da semente triturada, é misturado com água e adoçado com açúcar ou mel. A indústria alimentícia utiliza aproximadamente 70% da produção nacional de sementes de guaraná para fazer refrigerantes famosos mundialmente, havendo também uma crescente demanda da indústria farmacêutica e cosmética. Na indústria cosmética, os extratos de sementes são utilizados em sabonetes, cremes e shampoos. Estudos mostram que as metilxantinas presentes no guaraná têm efeito redutor da celulite, o que também faz com que ele seja utilizado em outras formulações cosméticas. Os produtos derivados do guaraná estão disponíveis em diversas partes do mundo, e uma parte da produção nacional é exportada, principalmente para o Japão, Estados Unidos, Canadá e Austrália (Funasaki *et al.*, 2016; Marques *et al.*, 2019; Patrick *et al.*, 2019; Schimpl *et al.*, 2013).

O guaraná possui inúmeras propriedades farmacológicas, tais como anticarcinogênica, antiproliferativa, antimicrobiana, antidepressivas, ansiolíticas e antioxidante, melhora do estado de alerta, tempo de reação, velocidade de processamento de informações, memória, humor e desempenho em exercícios físicos, além de efeitos termogênicos associados a perda de peso. A principal substância presente na semente de guaraná é a metilxantina cafeína sendo a quantidade variando entre 2,5 e 6%, quantidade até 5 vezes maior que a presente nas sementes de café arábica, possui também outras metilxantinas como a teobromina e a teofilina, em baixas concentrações. Elas também apresentam altos níveis de polifenóis, principalmente as proantocianidinas A2, B1, B2, B3 e B4, com maior prevalência de catequinas e epicatequinas. Outros compostos bioativos encontrados incluem saponinas, polissacarídeos, proteínas, ácidos graxos e alguns elementos inorgânicos como manganês, rubídio, níquel e estrôncio. Estudos mostram que a composição do óleo de guaraná contém uma variedade de substâncias, entre elas metilbenzenos, monoterpenos cíclicos, sesquiterpenos, ácido oleico, ácido paulínico, metoxifenil propeno (Marques *et al.*, 2016, 2019; Schimpl *et al.*, 2013; Silva, F. de A.

*et al.*, 2018)

A Embrapa está conduzindo pesquisas sobre diferentes variedades de guaraná, com o objetivo de aumentar sua produção e torná-las mais resistente a doenças. Em 2013, foram lançadas variedades que são resistentes à antracnose, principal doença que danifica o guaranazeiro, como as cultivares BRS Marabitanã e guaraná BRS Saterê. Essas variedades prometem aumentar a produção de guaraná na região amazônica em até 40%, sem a necessidade de aumentar o desmatamento da floresta (Marques *et al.*, 2019).

### 3.3 Metabolismo vegetal

De maneira geral, o metabolismo pode ser definido como um conjunto de transformações pelas quais são submetidas as substâncias que compõem um dado organismo biológico. Essas transformações ocorrem através de inúmeros processos bioquímicos (Figura 5), os quais geram como produto final metabólitos de diferentes classes químicas que são responsáveis por diferentes funções fisiológicas nos organismos vivos. Desse modo, os processos bioquímicos são frequentemente descritos como uma complexa rede de reações químicas, mediadas por diversas enzimas. Os metabólitos, moléculas de baixo peso molecular, surgem como os intermediários e produtos resultantes desse metabolismo. Os organismos vivos utilizam a energia e substâncias químicas de seus ambientes para sintetizar e degradar diversos metabólitos, atendendo às necessidades fisiológicas específicas e assegurando a sobrevivência e adaptação em ambientes sob constante mudança (Fang; Fernie; Luo, 2019; Weng, 2014).

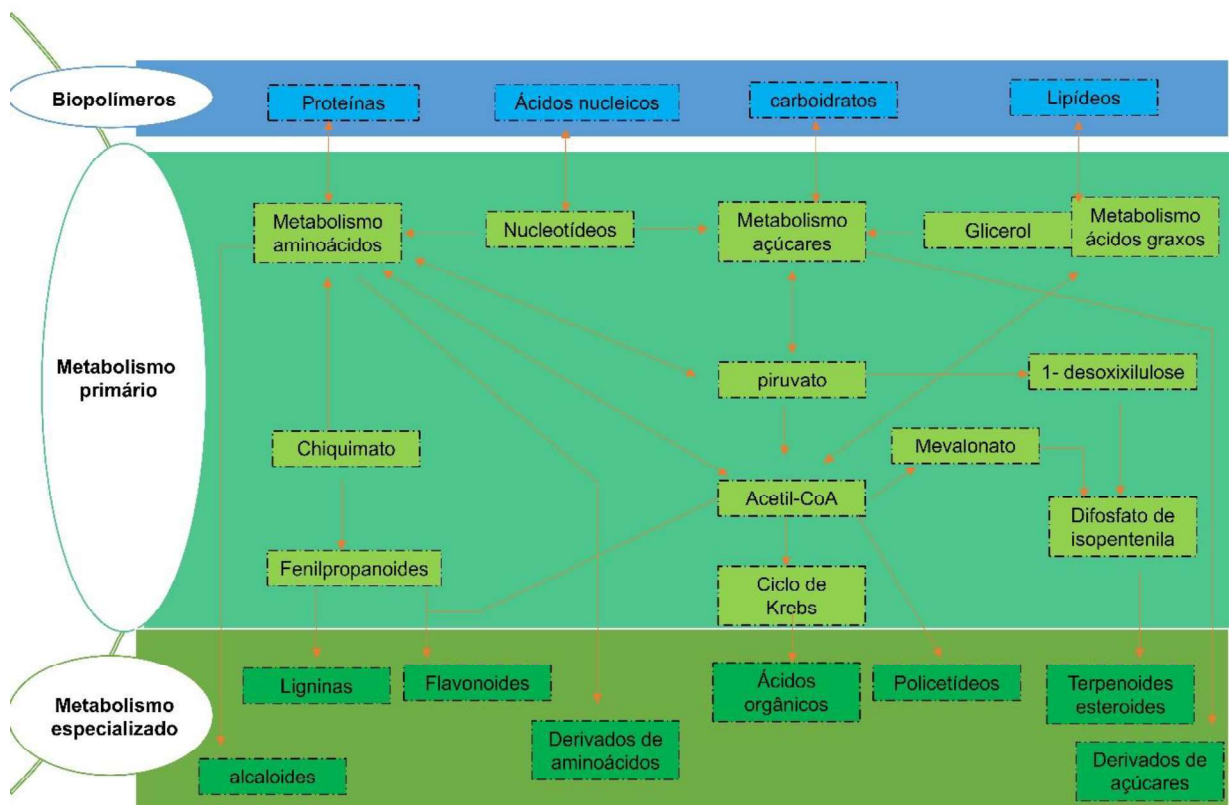
As plantas possuem a capacidade fisiológica de produzir uma grande diversidade de metabólitos. Essas substâncias desempenham um papel fundamental na manutenção do estado fisiológico adequado das plantas e no seu crescimento normal. A rede metabólica das plantas é complexa e extensa, sendo mais elaborada do que a dos demais organismos. Estima-se que o reino vegetal contenha entre 100.000 e 1 milhão de metabólitos, e cada espécie pode conter mais de 5.000 deles. Isso significa que as plantas são capazes de produzir uma quantidade significativamente maior de metabólitos em comparação com a maioria dos outros seres vivos (Alseekh; Fernie, 2018; Wang *et al.*, 2022).

Os metabólitos vegetais, além de suas funções estruturais, desempenham



funções vitais no crescimento, na reposição celular e na alocação de recursos em toda a planta, bem como na adaptação das plantas a diversos ambientes. Esses metabólitos são geralmente classificados como metabólitos primários e metabólitos especializados (metabólitos secundários), dependendo de suas funções biológicas principais. Os metabólitos primários geralmente são responsáveis por manter as atividades básicas da vida e reprodução das plantas, enquanto os metabólitos especializados são utilizados para responder às mudanças no ambiente circundante. No entanto, a distinção das funções entre metabólitos primários e especializados não é muito clara. Por exemplo, em relação aos aminoácidos, os mesmos não só fornecem blocos de construção essenciais às proteínas, mas também ligam o metabolismo central de carbono a uma variedade de metabólitos especializados. Além disso, por meio de diversas vias biossintéticas (Figura 5), algumas vezes até desconhecidas, é possível conceber uma série de combinações para essas subunidades, as quais podem fornecer uma grande diversidade de metabólitos especializados (Fang; Fernie; Luo, 2019; Simões *et al.*, 2017; Wang *et al.*, 2022).

Figura 5 – Diferentes rotas biossintéticas responsáveis pela formação de uma grande variedade de metabólitos especializados.



Fonte: Adaptado de SIMÕES *et al.*, 2017.

### 3.3.1 Metabolismo primário

Os metabólitos primários são indispensáveis para o crescimento e sobrevivência das plantas, produzidos por processos como respiração e fotossíntese, e consistem principalmente em produtos gerais como carboidratos, aminoácidos, nucleotídeos, vitaminas e ácidos graxos. As vias metabólicas e os compostos resultantes, Tabela 1, desses processos são muito semelhantes entre animais, bactérias, fungos, plantas e outros organismos, embora em alguns casos possam não ser idênticos. Por outro lado, os metabólitos secundários são geralmente específicos de cada espécie e ajudam as plantas a interagir com ambientes bióticos e abióticos. Eles são formados a partir de precursores quimicamente simples, derivados do metabolismo primário (Kessler; Kalske, 2018; Lacchini; Venegas-Molina; Goossens, 2023; Maeda, 2019; Wang *et al.*, 2022).

Tabela 1 – Principais vias metabólicas do metabolismo primário. (continua)

<b>Metabolismo</b>	<b>Biossíntese</b>	<b>Degradação</b>
Carboidratos	Fotossíntese, ciclo de Calvin, ciclo dos ácidos C-4, gliconeogênese	Clivagem hidrolítica de carboidratos, glicólise, ciclo da pentose-fosfato
Gorduras	Síntese de lipídeos, complexo-ácido graxosíntase, acilglicerídeos, fosfolipídeos, glicolipídeos, carotenoides, esteróis	Clivagem hidrolítica de lipídeos, $\beta$ -oxidação de ácidos graxos
Proteínas	Biossíntese de aminoácidos e proteínas	Clivagem hidrolítica de proteínas, degradação e conversão dos aminoácidos

Tabela 1 – Principais vias metabólicas do metabolismo primário. (conclusão)

Metabolismo	Biossíntese	Degradação
Acetil-coenzima A	Carboidratos, ácidos graxos, aminoácidos, cetogênese, terpenos, esteroides	Ciclo de Krebs, cadeia respiratória, ciclo do glioxilato
Ácidos nucleicos	Biossíntese de nucleotídeos do RNA a partir de bases púricas e pirimídicas, replicação de DNA, formação, formação de flavinas e pteridinas a partir de GTP	Clivagem de DNA e RNA

Fonte: Adaptado de SIMÕES *et al.*, 2017.

### 3.3.2 Metabólitos especializados

A inicial falta de clareza a respeito dos metabólitos especializados levou à designação de secundários, suspeitava-se que eles não tinham nenhuma função e eram apenas produtos residuais. Apesar das imprecisões e inconsistências desse termo, a intenção geral era definir compostos que estão presentes em algumas espécies de plantas, mas não em outras, e, portanto, não poderiam estar envolvidos no metabolismo primário que ocorre em todas as plantas. À medida que mais evidências sobre suas funções foram surgindo, ficou claro que a capacidade de sintetizar tais compostos evoluiu em diferentes linhagens de plantas, e que esses compostos representam adaptações a situações ecológicas específicas (Marone *et al.*, 2022; Pichersky; Lewinsohn, 2011; Rai; Saito; Yamazaki, 2017).

Os metabólitos especializados são os produtos formados pelas interações com o meio ambiente durante o crescimento e desenvolvimento das plantas. Essas substâncias são responsáveis pela proteção das plantas contra estresses abióticos

(radiação, temperatura, salinidade, hídrico e nutricional) e bióticos (agressão de patógenos e herbívoros). Além disso, alguns desses metabólitos podem desempenhar diferentes processos, como atração de polinizadores e dispersão de sementes. Acredita-se que a diversificação de metabólitos nas plantas tenha evoluído continuamente ao longo do tempo, com o objetivo de auxiliar as plantas a sobreviverem em condições adversas (Li *et al.*, 2023; Marone *et al.*, 2022; Pichersky; Lewinsohn, 2011; Rai; Saito; Yamazaki, 2017).

Enquanto a maioria dos metabólitos primários ocorre naturalmente nas células vegetais, os metabólitos especializados são encontrados apenas em certas espécies, tecidos, órgãos, estágios de desenvolvimento ou condições ambientais específicas. A relevância desses metabólitos na medicina e economia impulsionou pesquisas sobre sua biossíntese em plantas. Ao longo dos séculos, a humanidade explorou essas adaptações fisiológicas em plantas para aprimorar a biossíntese de compostos bioativos, utilizados na produção de medicamentos ou diretamente como fitoterápicos. Atualmente, aproximadamente um quarto das drogas clínicas são derivadas desses metabólitos presentes em plantas, além de seu uso em alimentos, fragrâncias, cosméticos e produtos agrícolas. (Fang; Fernie; Luo, 2019; Funari *et al.*, 2013; Liu *et al.*, 2023; Yuan; Grotewold, 2020)

Os metabólitos especializados são comumente categorizados conforme os precursores, composição química ou mecanismos enzimáticos envolvidos em sua formação. Os principais grupos desses metabólitos em plantas incluem compostos fenólicos (como ácidos fenólicos, flavonoides e taninos), terpenoides (incluindo carotenoides) e compostos nitrogenados/sulfurados (como alcaloides e glucosinolatos), cada um desempenhando funções distintas (Marone *et al.*, 2022; Singh; Agrawal; Bednarek, 2023). Esses são muito mais numerosos do que os metabólitos produzidos pelo metabolismo primário, estima-se algo de 21.000 alcaloides, 5.000 flavonoides e 22.000 terpenoides identificados até agora. No entanto, é provável que esse número seja maior, já que muitas plantas ainda possuem metabolomas não determinados (Lim; Julca; Mutwil, 2023).

### 3.3.2.1 Compostos fenólicos

Os compostos fenólicos desempenham papéis importantes em muitos processos biológicos, são os principais responsáveis para as respostas das plantas a

estresses bióticos e abióticos, como a proteção contra a radiação solar e a robustez contra danos mecânicos e na defesa mediadora contra patógenos e herbívoros. Além disso, estão envolvidos na pigmentação de flores e frutos, aspectos importantes para a reprodução e dispersão de sementes. Lignina, taninos, cumarinas, estilbenos, flavonoides e isoflavonoides, são exemplos de grupos fenólicos (Lei *et al.*, 2015; Liu *et al.*, 2023; Zaynab *et al.*, 2018).

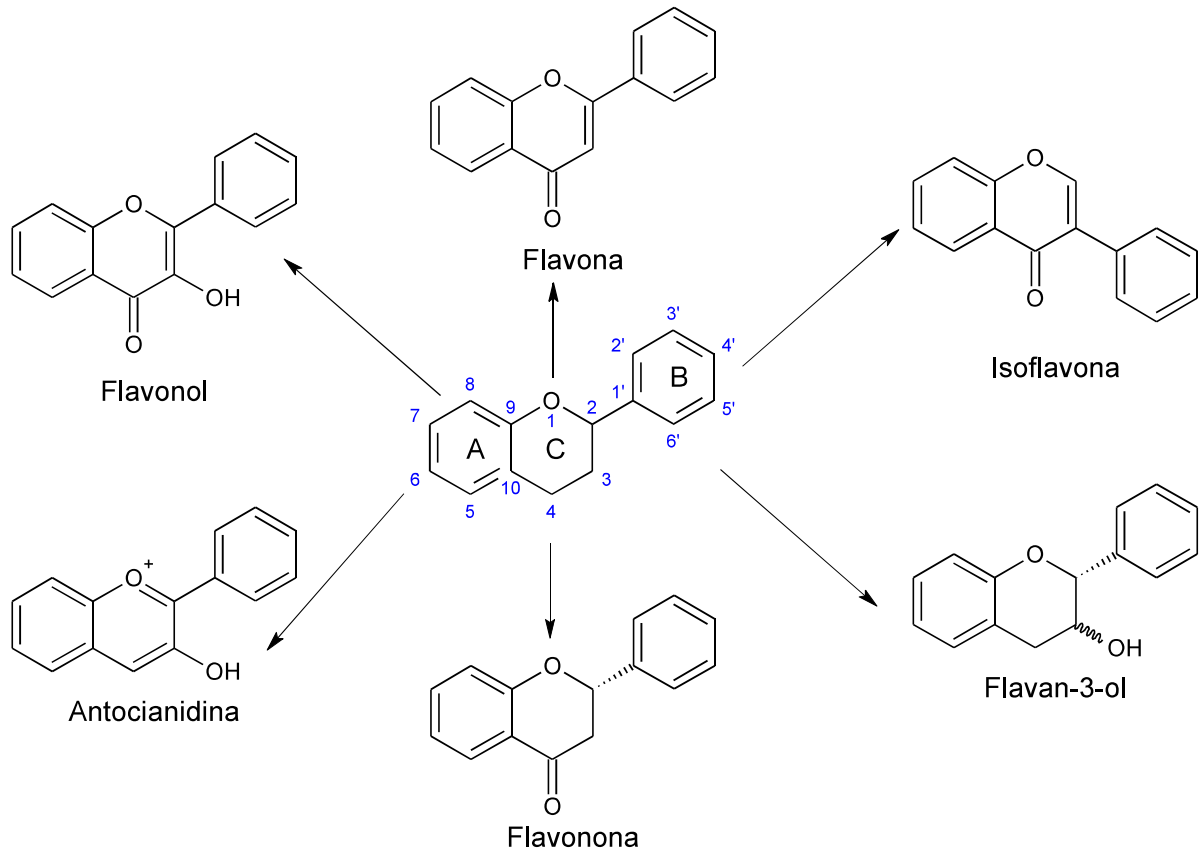
Os compostos polifenólicos consistem em múltiplos esqueletos de anéis fenólicos com grupos hidroxila ou outros substitutos, como moléculas de açúcar e ácidos orgânicos. Milhares de compostos com estrutura de polifenol foram caracterizados em plantas superiores. Devido às suas propriedades antioxidantes e antiproliferativas, eles são altamente valiosos na nutrição humana, e estudos epidemiológicos sugerem que uma alta ingestão alimentar de polifenóis está associada a uma diminuição do risco de doenças cardiovasculares e de câncer (Pott; Osorio; Vallarino, 2019).

Os flavonoides são compostos polifenólicos encontrados em todas as plantas vasculares e não vasculares. Contêm uma estrutura polifenólica característica, composta por três anéis aromáticos (C6-C3-C6). as unidades são chamadas de núcleos A, B, e C, e os átomos de carbono recebem a numeração com números ordinários para os núcleos A e C e os mesmo números seguidos de uma linha (‘) para o núcleo B. Embora não sejam essenciais para o crescimento e desenvolvimento das plantas, os flavonoides têm papéis específicos da espécie na nodulação, fertilidade, defesa e proteção UV (Peer; Murphy, 2007). Flavonoides e isoflavonoides compreendem flavonas, flavonóis, flavanonas e isoflavonas que diferem em suas estruturas de anel central (Figura 6). A complexidade dos flavonoides e isoflavonoides é ainda aumentada por diferentes padrões de hidroxilação, metilação e glicosilação. A diversidade e complexidade dessas estruturas representam um desafio substancial para estabelecer o perfil qualitativo, quantitativo e em grande escala de flavonoides e isoflavonoides em metabolômica de plantas e animais (Lei *et al.*, 2015; Simões *et al.*, 2017).

Os flavonoides são frequentemente encontrados na forma oxigenada, sendo muitos deles conjugados com açúcares. Esta forma, conhecida como heterosídeo, é referida como O-heterosídeo quando a ligação ocorre por meio de uma hidroxila e C-heterosídeo quando a ligação é estabelecida com um átomo de carbono. Quando o flavonoide está sem açúcar, é designado como aglicona ou genina (Simões

et al., 2017).

Figura 6 – Estrutura básica dos flavonoides e das suas classes mais comuns.



Fonte: Elaborado pela Autora.

Os polifenóis são metabólitos especializados que oferecem diversos benefícios à saúde, como propriedades antioxidantes, antialérgicos, antimicrobianos, anti-inflamatórios e antitrombóticos. Devido a essas propriedades, eles são amplamente incorporados em formulações de medicamentos e cosméticos. Os flavonoides encontrados nas folhas podem ser diferentes daqueles presentes nas flores, nos galhos, nas raízes ou nos frutos. Os resíduos de frutas são reconhecidos por conter níveis significativos de antioxidantes maiores quando comparados com a fruta inteira. O principal grupo de antioxidantes fenólicos inclui ácidos fenólicos, flavonóides e taninos. Compostos amplamente estudados, como quercetina, catequina e kaempferol estão entre os compostos polifenólicos que apresentam potencial na redução do risco de doenças crônicas, como câncer, diabetes, doenças cardiovasculares e obesidade (Banerjee *et al.*, 2018; Simões *et al.*, 2017).

Alguns flavonoides como flavan-3-ols catequina e epicatequina

polimerizam para formar taninos. Os taninos são um grupo de polifenóis que podem ser divididos em duas classes: taninos condensados, formados pela policondensação de duas ou mais unidades flavan-3-ol e flavan-3,4-diol e taninos hidrolisáveis que podem ser descritos como ésteres de ácido gálico com um poliol central, tipicamente  $\beta$ -D-glicose (Pott; Osorio; Vallarino, 2019; Simões *et al.*, 2017).

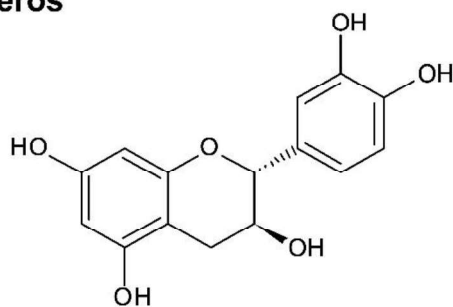
Os taninos condensados também são conhecidos como proantocianidinas. Esses metabólitos estão presentes em flores, frutas, cascas e sementes de diversas plantas, como defesa contra estressores bióticos e abióticos. A adstringência deles protege as plantas de patógenos e predadores. O grau de polimerização da proantocianidinas pode variar entre 3 e 11. As procianidinas podem ser categorizadas em tipo A e tipo B, Figura 7, dependendo da configuração estereo e da ligação entre os monômeros. As procianidinas do tipo B ( $C_{30}H_{26}O_{12}$ ) são caracterizadas por uma ligação entre o carbono C4 da unidade superior e o carbono C8 ou C6 da unidade inferior. As procianidinas do tipo B são as mais abundantes, sendo as procianidinas B1, B2, B3 e B4 as mais frequentes. As procianidinas do tipo A ( $C_{30}H_{24}O_{12}$ ) têm não apenas a ligação entre átomos de carbono, mas também uma ligação éter C2-O-C7 entre as duas unidades manoméricas. Os compostos do tipo A mais comuns são A1 e A2 (Rauf *et al.*, 2019; Rue; Rush; van Breemen, 2018).

### 3.3.2.2 Terpenoides

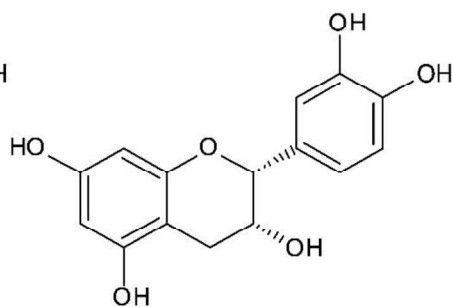
A maior classe dos metabólitos especializados são os terpenoides, cuja a estrutura deriva do 2-metilbutadieno, também denominado isopreno, são os principais componentes das emissões de compostos voláteis florais e vegetativos como  $\beta$ -cariofileno, limoneno e linalol, juntamente com os fenilpropanoides benzaldeído e ácido salicílico. Muitos terpenoides são de interesse comercial, pois são aplicados como pesticidas, agentes antimicrobianos e anticarcinogênicos. Além disso, eles também são usados como precursores para produzir produtos químicos, como vitaminas (Kessler; Kalske, 2018; Pott; Osorio; Vallarino, 2019; Simões *et al.*, 2017).

Figura 7 – Estrutura química de flavan-3-ols, monômeros e procianidinas.

### Monômeros

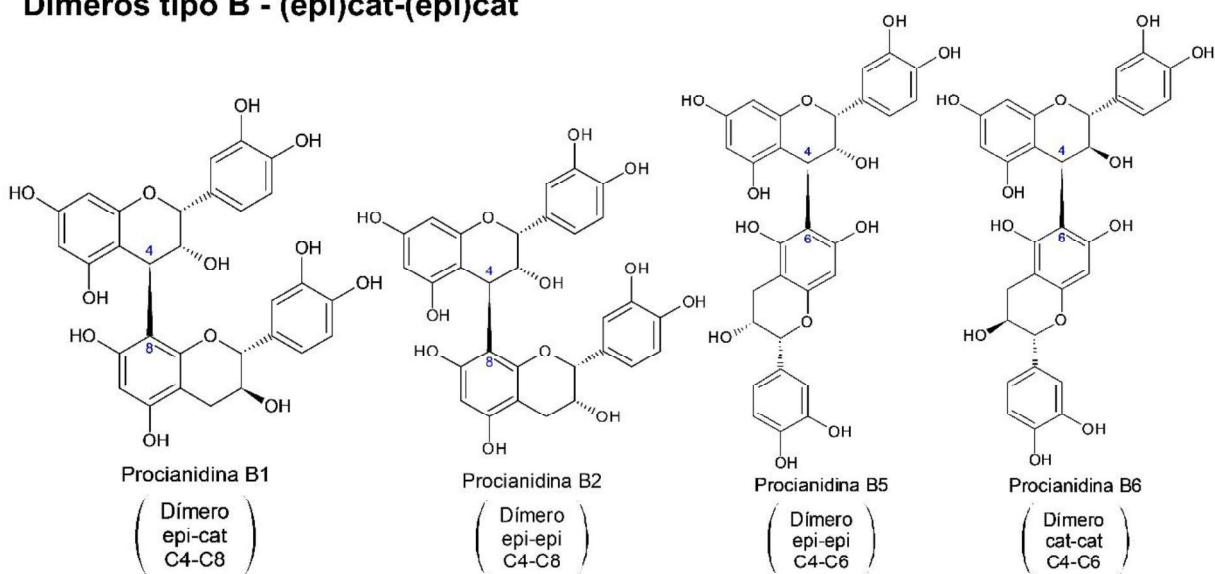


(+) Catequina

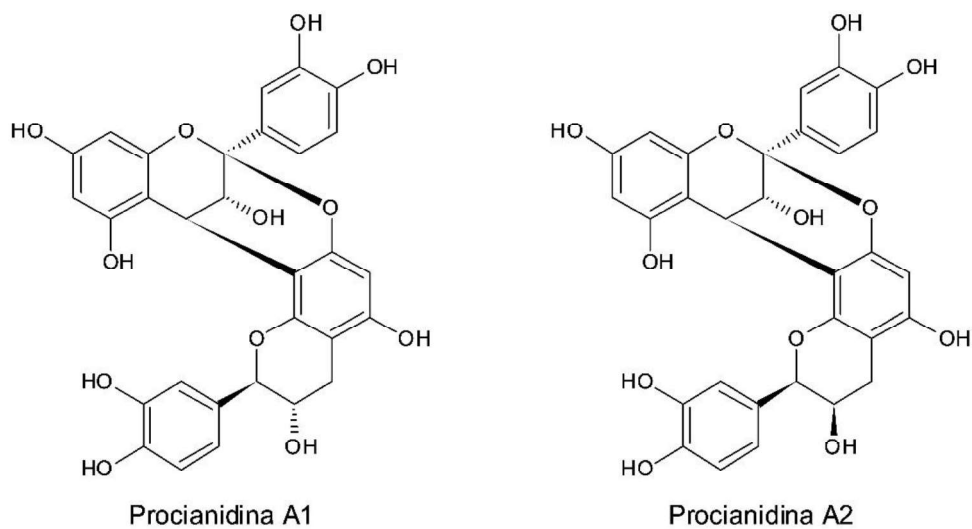


(-) Epicatequina

### Dímeros tipo B - (epi)cat-(epi)cat



### Dímeros tipo A - (epi)cat=(epi)cat



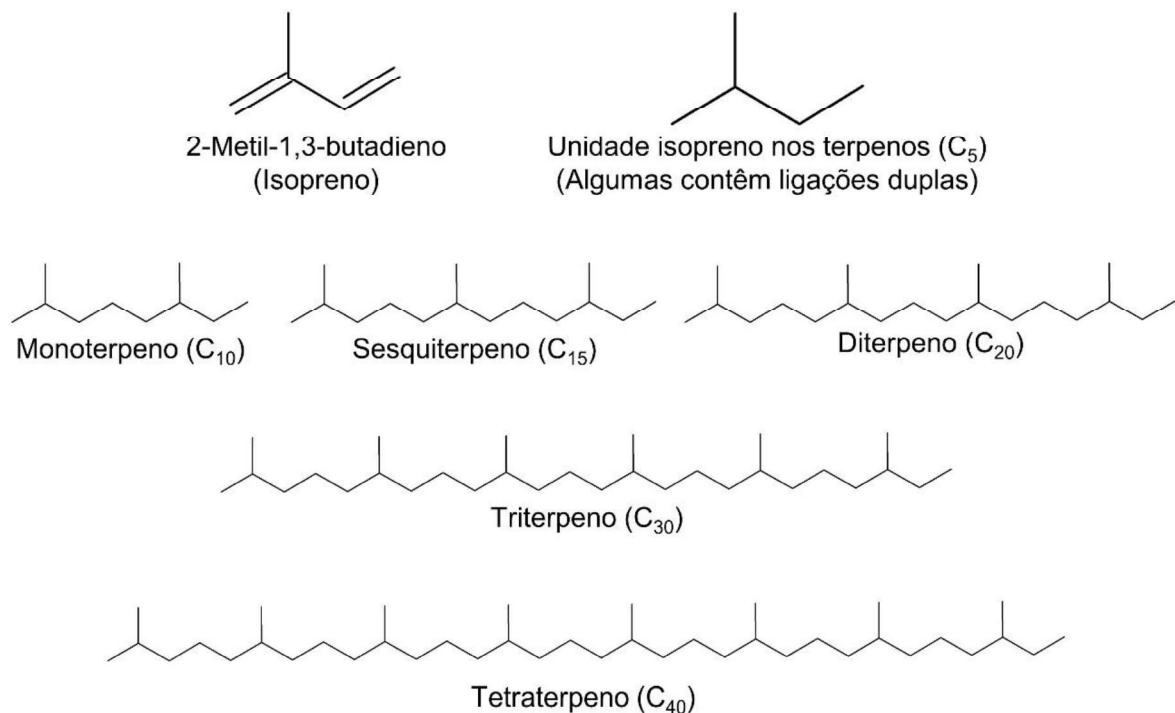
Fonte: Elaborado pela Autora.



Diferentes classes de terpenoides são formadas de um número, normalmente múltiplo de cinco, de carbonos que são os monoterpenoides (10C), sesquiterpenoides (15C) e os diterpenoides (20C). Os terpenos maiores incluem os triterpenos (C<sub>30</sub>), tetraterpenos (C<sub>40</sub>) e politerpenos ([C<sub>5</sub>]<sub>n</sub>, Figura 8. Como exemplo, o óleo essencial de frutas cítricas é formado principalmente pelo monoterpene limoneno. Seus principais usos terapêuticos são como antiespasmódicos, expectorante, atividades estomáquicas, estimuladores ou depressores do sistema nervoso central, anestésicos locais e anti-inflamatórios (Pott; Osorio; Vallarino, 2019; Simões *et al.*, 2017).

Os carotenoides são tetraterpenoides responsáveis pelos tons brilhantes e atraentes de amarelo, laranja e vermelho de muitas frutas, como tomate, abóbora, caqui. Além disso, desempenham papéis fundamentais na fotossíntese e fotoproteção, e também fornecem precursores para a biossíntese de hormônios vegetais: ácido abscísico e estrigolactonas. Além disso, eles atuam como fitonutrientes promotores da saúde e têm sido associados à prevenção de doenças cardiovasculares, câncer, diabetes, mal de Alzheimer e outras doenças relacionadas à idade (Pott; Osorio; Vallarino, 2019).

Figura 8 – Alguns terpenoides formados a partir de unidades de isopreno.



Fonte: Adaptado de Guedes, 2018.

### 3.3.2.3 Alcaloides

Os alcaloides são metabólitos especializados de ocorrência natural, são caracterizados como compostos contendo nitrogênio, onde um ou mais átomos de nitrogênio podem estar em um heterociclo (Belew *et al.*, 2022; Bhambhani; Kondhare; Giri, 2021; Liu *et al.*, 2023). Dependendo de sua estrutura química, os alcaloides apresentam uma gama de atividade farmacológicas, vão desde propriedades medicinais até toxicidade aguda. Nesse contexto, um dos primeiros alcaloides utilizados é a morfina, o qual é extraído da flor da papoula. De forma geral, apresentam características de neurotransmissores, demonstrando papel de regulação, estimulação e indução de funções (Bhambhani; Kondhare; Giri, 2021; Simões *et al.*, 2017).

Um importante alcaloide é cafeína, esse metabólito entra na composição de diversos medicamentos, tais como analgésicos, antipiréticos e antigripais, associado com o ácido acetilsalicílico, paracetamol, codeína e diidroergotamina, no alívio ou abortamento das crises de enxaqueca. Além disso, a cafeína também é usada como fármaco isolado, sobretudo na depressão respiratória em neonatos. Outro alcaloide é a teofilina que é um broncodilatador utilizado para tratamento de asma e algumas formas espásticas de pneumopatias obstrutivas, como enfisema e bronquite crônica (Simões *et al.*, 2017).

## 3.4 Metabolômica

As ciências ômicas visam compreender o funcionamento celular dos organismos e suas mudanças biológicas. Dentro desse conjunto, incluem-se a genômica (análise das alterações genéticas), transcriptômica (estudo das mudanças nos transcritos), proteômica (investigação das alterações nas proteínas) e metabolômica (exploração das mudanças nos metabólitos). No entanto, a genômica, transcriptômica e proteômica não oferecem informações estruturais sobre fitoquímicos e seus intermediários de reação, destacando a importância dos estudos metabolômicos para compreender a biossíntese de metabólitos especializados (Canuto *et al.*, 2018; Rai; Saito; Yamazaki, 2017).

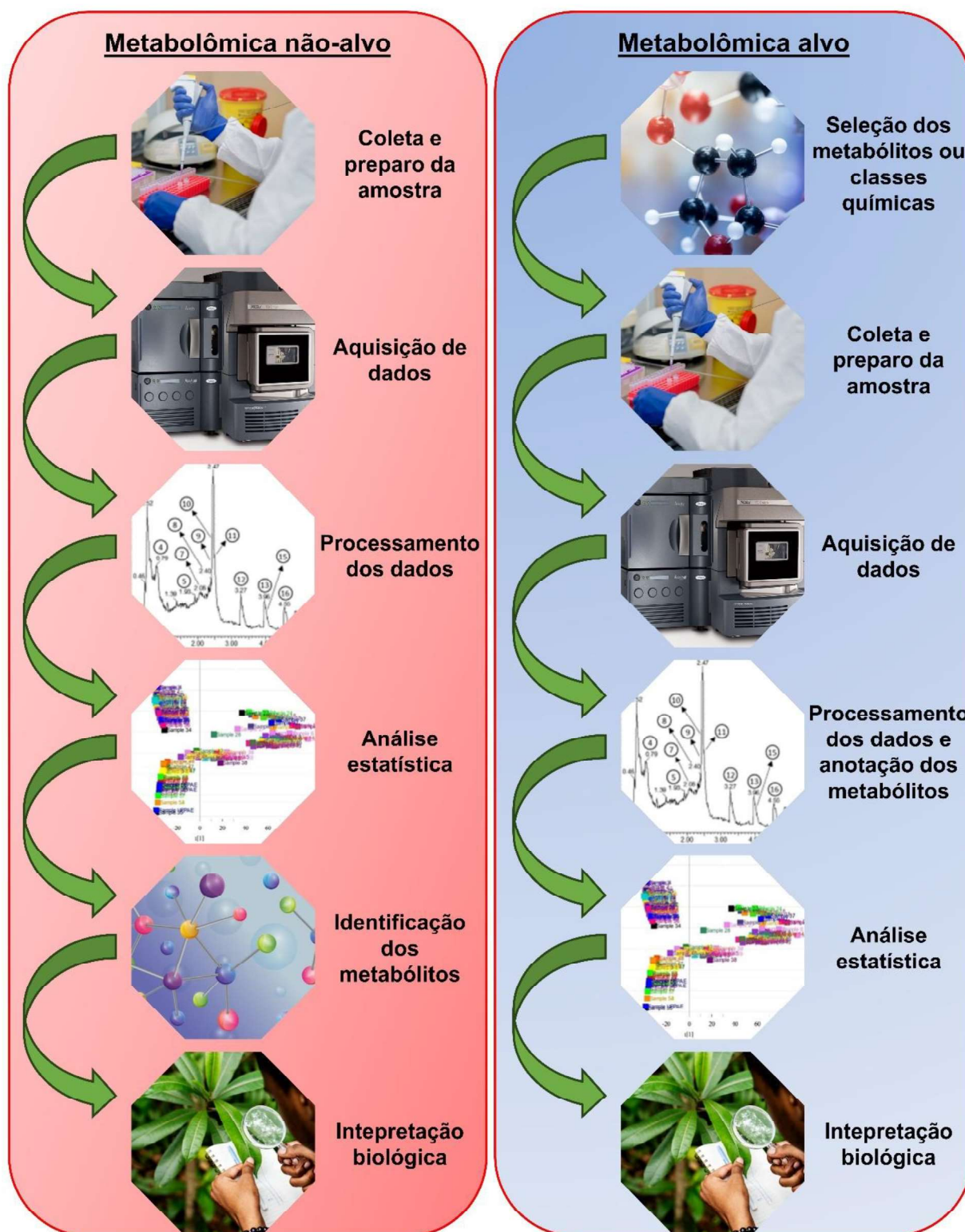
A metabolômica surgiu no final dos anos de 1990, originando-se dos estudos de genômica funcional em leveduras por Oliver e Ferenci. Em 2000, Fiehn e colaboradores do *Max-Planck Institute of Plant Physiology* apresentaram as primeiras

aplicações da metabolômica em plantas (Fiehn, 2001, 2002). No decorrer dos anos, diversas definições foram propostas para a metabolômica, abrangendo tanto o tamanho molecular das substâncias (< 1500 Da), quanto a classe estrutural. A definição mais difundida destaca a metabolômica como a ciência que investiga a expressão qualitativa e quantitativa total do metabolismo primário e secundário dos organismos. No entanto, ressalta-se que a metabolômica enfatiza a compreensão dos organismos por meio da análise comparativa de perfis metabólicos entre indivíduos e/ou populações sujeitos às diferentes condições genéticas, ambientais ou patológicas (Alseekh; Fernie, 2018; Luque de Castro; Delgado-Povedano, 2014; Pilon *et al.*, 2020).

Os fluxos de trabalho metabolômicos fundamenta-se na premissa de que o pesquisador identifica previamente o tipo de metabólitos a ser analisado. Uma abordagem metabolômica direcionada implica uma análise quantitativa (determinação das concentrações absolutas) ou uma análise semiquantitativa (determinação das intensidades relativas) de um conjunto predefinido de metabólitos, que podem estar ligados a classes químicas comuns ou a uma via metabólica específica. Por outro lado, uma abordagem metabolômica não direcionada ou global concentra-se principalmente na análise qualitativa ou semiquantitativa do maior número possível de metabólitos pertencentes a diversas classes químicas e biológicas presente em uma amostra. Sua principal vantagem reside na capacidade imparcial de examinar a relação entre metabólitos interconectados de múltiplas vias. O fluxo de trabalho metabolômico (Figura 9) compreende as etapas sequenciais de análises comparativas entre metabolômicas direcionadas e não direcionadas (Isah, 2019; Johnson; Ivanisevic; Siuzdak, 2016; Tebani; Afonso; Bekri, 2018).

Para conduzir um estudo metabolômico, é essencial inicialmente delinear previamente o problema a ser investigado, formulando uma ou mais perguntas a serem elucidadas ao término do estudo. Outro ponto crucial é a escolha da abordagem metabolômica que será empregada (direcionada ou não direcionada), pois a partir dessa será definido os procedimentos de preparo da amostra e análise. Em resumo, os processos típicos em estudos metabolômicos inclui a formulação de problema, coleta, preparação e análise da amostra, aquisição de dados e por fim, a interpretação biológica (Canuto *et al.*, 2018; Marques; Justino, 2023; Mushtaq *et al.*, 2014; Tebani; Afonso; Bekri, 2018).

Figura 9 – Etapas e processos gerais envolvidos nas análises metabolômicas.



Fonte: Elaborado pela Autora.

### 3.4.1 Coleta e preparo da amostra

A padronização durante a coleta é fundamental na metabolômica de plantas

e envolve considerações como o tipo de tecido (folhas, frutos, flores, caule, raiz), os períodos sazonais, os horários de coleta e os estágios ontogenéticos (crescimento e reprodução). Uma vez coletado o material vegetal, a inibição imediata da atividade enzimática é crucial para preservar a integridade metabólica. Essa inibição pode ser realizada de diversas maneiras, como o tratamento da amostra com metanol (quente > 80 °C ou frio -40 °C), variações de pH e/ou redução brusca da temperatura com nitrogênio líquido, provocando um choque térmico abrupto que auxilia na interrupção das reações bioquímicas (Alseekh *et al.*, 2021; Mushtaq *et al.*, 2014; Pilon *et al.*, 2020; Villate *et al.*, 2021).

Outro passo é a secagem das amostras, facilitando o armazenamento a longo prazo das amostras, inibindo a atividade enzimática e o crescimento microbiano. Além disso, a ausência de água nas amostras é crucial, pois a água pode impactar no poder de solvatação dos solventes de extração e interferir nos instrumentos utilizados nas análises das amostras. As amostras podem ser secas por meio de aquecimento, utilizando estufas, micro-ondas ou liofilização (Mushtaq *et al.*, 2014; Villate *et al.*, 2021).

As amostras precisam ser cuidadosamente homogeneizadas para garantir uma extração eficaz. À medida que a complexidade de um organismo aumenta, também aumenta sua heterogeneidade, como no caso de plantas (Mushtaq *et al.*, 2014). A trituração ou moagem se faz necessária para diminuir o tamanho dos tecidos e assim homogeneizar a amostra. A padronização desse processo (controle granulométrico) é fundamental em estudos metabolômicos, assegurando a homogeneidade na extração e proporcionalidade na área de contato entre a matriz e o solvente extrator. Embora almofariz e pistilo sejam os comumente usados na metabolômica, sua aplicação não é recomendada para estudos com muitas amostras, tendo em vista que o processo pode se tornar bastante moroso. Moinhos e homogeneizadores verticais vem sendo empregados com sucesso. A razão para o aumento da eficiência de extração com partículas de tamanho menores é a dependência da difusão de massa ou tamanho das partículas, reduzindo o tempo necessário para a difusão dos solventes e facilitando a extração direta dos metabólitos pelos solventes (Mushtaq *et al.*, 2014; Villate *et al.*, 2021).

Idealmente, o processo de coleta, secagem e moagem deve ser seguido imediatamente pela extração e análise das amostras. Entretanto, quando isso não é viável, é aconselhável armazenar as amostras em freezer a -80 °C. Embora em alguns

casos o condicionamento das amostras a temperaturas de -20 °C e 4 °C seja utilizado, não é recomendado devido à possibilidade de ocorrência de reações bioquímicas e atividades enzimáticas, mesmo em temperaturas inferiores a -20 °C, especialmente em amostras contendo sais ou solventes orgânicos (Dudzik *et al.*, 2018; Pilon *et al.*, 2020).

O processo de preparo de amostra para análise metabolômica é uma das etapas mais cruciais e que consome considerável tempo de trabalho. Isso decorre da complexidade, heterogeneidade e ampla faixa de concentração dos metabólitos presentes nas amostras. A execução desse processo está condicionado à escolha da abordagem metabolômica, ao tipo de amostra e às técnicas de análise relevantes ao estudo, garantindo assim a compatibilidade entre o preparo de amostra e a técnica analítica empregada (Canuto *et al.*, 2018).

O objetivo primordial da extração é obter os metabólitos alvo ou o número máximo de metabólitos, idealmente, todos os metabólitos presentes na amostra. Diversos métodos de extração, geralmente utilizando diferentes combinações de solventes, são empregados para alcançar esse fim. Assim, apenas um método de extração abrangente e reprodutível proporcionará dados confiáveis, uma vez que os metabólitos identificados são os que foram efetivamente extraídos e todas as conclusões serão construídas em torno dessas informações (Mushtaq *et al.*, 2014).

Para obter resultados eficazes, é necessário considerar não apenas o tipo de solvente, mas também as características físico-químicas da matriz, o efeito do pH na matriz, o tempo de contato e a compartimentalização dos metabólitos. A seleção do solvente deve contemplar aspectos como toxicidade, poder de solubilização, seletividade, taxa de dissolução, reatividade química e pH. O solvente ideal é aquele que combina a menor toxicidade com o maior poder de solubilização (Mushtaq *et al.*, 2014). Por exemplo, em um estudo não direcionado desenvolvido através de métodos cromatográficos, diferentes tipos de solventes podem ser utilizados: como solvente monofásico (água / metanol, água / acetonitrila) ou bifásico (água e metanol, frequentemente combinados com um solvente apolar como clorofórmio, diclorometano ou éter metil terc-butílico), podem ser empregados em sistemas de extração, dependendo da análise planejada. A escolha dos sistemas solventes é determinada pelo interesse em investigar moléculas polares ou apolares (Pilon *et al.*, 2020; Tebani; Afonso; Bekri, 2018).

Os métodos de preparação e introdução de amostras para análise de

amostras biológicas podem abranger diversas técnicas, incluindo, mas não se limitando a injeção direta, extração assistida por micro-ondas, extração por Soxhlet, extração líquido-líquido (LLE), extração em fase sólida (SPE) e extração com fluido supercrítico. A fase de extração determina o intervalo ou cobertura química, levando em consideração a polaridade e solubilidade dos metabólitos a serem detectados pelas técnicas analíticas. É crucial que essa etapa seja conduzida de maneira simples, rápida e eficiente, uma vez que a metabolômica envolve um considerável número de amostras e, conseqüentemente, de experimentos (Mushtaq *et al.*, 2014; Pilon *et al.*, 2020; Villate *et al.*, 2021).

### **3.4.2 técnicas analíticas utilizadas em estudos metabólicos**

Atualmente, não há uma técnica analítica que possa medir todos os metabólitos em um único experimento, devido à considerável diversidade química e concentrações variadas dessas substâncias. Por isso, a aquisição de dados em estudos metabolômicos é conduzida por meio do uso de multiplataformas de análise. Essas plataformas oferecem uma cobertura mais abrangente em termos de substâncias químicas detectadas, proporcionando, por conseguinte, um entendimento biológico mais amplo do organismo em estudo. Técnicas de análise como a ressonância magnética nuclear (RMN ou NMR, do inglês, *nuclear magnetic resonance*) e a espectrometria de massas (MS, do inglês, *mass spectrometry*), que fornecem informações estruturais de diversas classes químicas, são as técnicas analíticas mais empregadas nos estudos metabolômicos (Canuto *et al.*, 2018; Forcisi *et al.*, 2013; Pilon *et al.*, 2020; Verpoorte; Choi; Kim, 2007; Wang *et al.*, 2023; Yuliana *et al.*, 2013).

A ressonância magnética nuclear (NMR) é uma técnica que permite a identificação inequívoca de uma substância. Além disso, tem como características gerais ser uma técnica de análise robusta e abrangente que requer pouca ou nenhuma manipulação de amostra, possibilitando a análise de amostras biológicas intactas, incluindo sólidos e semissólidos. Uma vantagem é a necessidade de uma pequena quantidade de amostra, e ela não é destruída após a análise, embora seja contaminada por solventes deuterados. No entanto, apresenta desafios como baixa sensibilidade e seletividade, com regiões espectrais que podem ter sobreposição de

sinais, comprometendo a interpretação dos resultados. Além disso, requer equipamentos de alta resolução (Atanasov *et al.*, 2021; Canuto *et al.*, 2018; Kueger *et al.*, 2012; Verpoorte; Choi; Kim, 2007).

As estratégias metabolômicas que utilizam a espectrometria de massa (MS) em conjunto com técnicas de separação, como cromatografia gasosa (GC) ou líquida (LC), e até mesmo eletroforese capilar, oferecem uma sensibilidade e flexibilidade substancialmente superiores. Isso resulta em abordagens baseadas em MS que proporcionam informações mais abrangentes e detalhadas sobre o inventário metabólico de uma amostra biológica, embora a identificação precisa dos compostos ainda seja um desafio significativo (Kueger *et al.*, 2012).

A GC-MS é amplamente utilizada em estudos metabolômicos devido à sua elevada robustez, precisão, sensibilidade e exatidão nas análises. A análise de compostos voláteis por *headspace* tem sido empregada para investigar a fração volátil do metaboloma. No entanto, muitos metabólitos necessitam de derivatização para tornarem-se voláteis a baixas temperaturas. Esse processo é tedioso, podendo introduzir erros por volatilização durante o procedimento e limitar o número de amostras processadas simultaneamente. Apesar disso, o uso de GC-MS na metabolômica oferece uma vantagem significativa, dada as bibliotecas de espectros, proporcionando uma identificação confiável dos metabólitos com base na informação de tempo de retenção e padrão de fragmentação (Canuto *et al.*, 2018; Dettmer; Aronov; Hammock, 2007; Soga, 2023).

A cromatografia líquida (LC) acoplada à espectrometria de massa (MS) é uma plataforma analítica crucial para estudos metabolômicos não direcionados em larga escala devido à sua excelente combinação de robustez, sensibilidade e seletividade. A LC-MS é reconhecida como uma técnica abrangente para analisar uma ampla classe de compostos, graças à diversidade de fases estacionárias disponíveis e aos diversos modos de separação, incluindo eluição em fase reversa (essencialmente por partição), com ou sem pareamento iônico, interação hidrofílica e troca iônica. Na análise de perfis metabólicos baseados em LC-MS, os metabólitos são geralmente identificados comparando o tempo de retenção do analito e a razão massa/carga ( $m/z$ ) com padrões autênticos ou com dados cromatográficos / espectrais oriundos da literatura (desreplicação), os quais devem ter sido analisados de forma semelhantes (Canuto *et al.*, 2018; Lei *et al.*, 2015).

O UPLC tem vantagens sobre o HPLC convencional devido à sua



velocidade analítica mais rápida e maior eficiência de separação. Além disso, a combinação QTOF-MS (analisadores de massa quadrupolo-tempo de voo) fornece medições de massa de alta resolução, permitindo a identificação rápida de metabólitos e a determinação precisa de compostos em nível de baixa concentração. Usando um sistema acoplado, no qual se tem um cromatógrafo à líquido com espectrômetro de massas de alta resolução (UPLC-QTOF-MS<sup>E</sup>), a determinação de vários compostos e a concepção de um perfil metabólico de uma amostra podem ser alcançadas em uma execução cromatográfica em um curto espaço de tempo, atendendo aos requisitos de rapidez, eficiência e precisão (Fan *et al.*, 2013).

A eletroforese capilar acoplada à espectrometria de massa (CE-MS) é empregada como uma técnica complementar à LC-MS e GC-MS, destacando-se pela separação de compostos iônicos polares. Caracteriza-se por alta resolução e análises rápidas. Com volumes de injeção na ordem de nL, a CE-MS é vantajosa ao lidar com fluídos biológicos escassos, como urina de rato e saliva de animais. Entretanto, enfrenta desafios de repetibilidade e sensibilidade, este último devido à diluição das amostras pelo uso de líquido auxiliar na integração com a MS (Canuto *et al.*, 2018; Soga, 2023).

### **3.4.3 Análise estatísticas em estudos metabolômicos**

De modo geral, os dados oriundos de estudos metabolômicos podem ser usados para construir hipóteses ou para explicar observações. Os metabólitos identificados associados a uma observação podem fornecer uma visão holística sobre o sistema biológico interrogado (Tebani; Afonso; Bekri, 2018). Entretanto, é importante salientar que em estudos metabolômicos, frequentemente são gerados extensos e grandes volumes de dados de alta complexidade de interpretação, tornando praticamente impossível inferir as informações manualmente. Conseqüentemente, análises multivariadas de dados são usadas rotineiramente para inferir informações de conjuntos dados metabolômicos (Alonso; Marsal; Julià, 2015). Sendo assim, os conjuntos de dados são processados de forma a tornar mais evidente as informações e as interpretações relacionadas ao mesmo. Nesse contexto, várias ferramentas e softwares foram desenvolvidos para ajudar no processamento de dados, de modo a evitar erros e manter a integridade das variações biológicas inspecionadas (Boufridi; Quinn, 2016; Canuto *et al.*, 2018; Correia; Ferreira, 2007; Hu; Xu, 2013; Moco *et al.*,

2007; Tebani; Afonso; Bekri, 2018).

Também é importante mencionar que na maioria dos casos, os dados provenientes de equipamentos analíticos carecem de um tratamento prévio. Haja vista que os perfis metabólicos estão sujeitos a erros experimentais inerentes do método analítico, os quais devem ser corrigidos antes da análise e avaliação dos dados propriamente dita. Assim um, pré-tratamento dos dados, sejam aqueles obtidos por técnicas de separação e/ou espectroscópicas, poderão ser usadas para diferentes propósitos. Além disso, a variabilidade instrumental, como flutuações de pressão e temperatura e degradação de fase estacionária, bem como o efeito da matriz de amostras, podem intensificar a complexidade dos dados. Portanto, as etapas de pré-processamento devem ser consideradas em dados brutos, a fim de eliminar ou pelo menos mitigar as informações indesejáveis e reduzir as variações quimicamente irrelevantes antes de aplicar uma análise quimiométrica. As técnicas de pré-processamento mais comuns usadas para conjuntos de dados metabolômicos baseados em cromatografia-massa, por exemplo, são correção de linha de base, redução de ruído, alinhamento de tempo de retenção, desconvolução espectral e normalização. Combinadas, as etapas de processamento de dados ajudam a preparar os dados para uma análise multivariada que proporcione uma interpretação mais exata e confiável (Feizi *et al.*, 2021; Funari *et al.*, 2013; Paul; de Boves Harrington, 2021; Shen *et al.*, 2023).

As ferramentas quimiométricas para análise de dados metabolômicos devem ser selecionadas de acordo com o objetivo do estudo. Se o objetivo for a classificação da amostra e as informações prévias sobre a identidade da amostra forem desconhecidas, métodos não supervisionados, como análise de agrupamento hierárquico (HCA) ou análise de componentes principais (PCA) são usados. Por outro lado, a identidade da amostra é frequentemente conhecida e o objetivo do estudo é descobrir biomarcadores característicos, nesse caso, métodos supervisionados, como mínimos quadrados principais (PLS) ou método independente de analogia de classe (SIMCA), também podem ser usados (Dettmer; Aronov; Hammock, 2007).

A HCA e a PCA permitem a visualização gráfica de todo o conjunto de dados, mesmo quando o número de amostras e variáveis é elevado. O uso desses algoritmos tem como objetivo principal facilitar e aumentar a compreensão do conjunto de dados, examinando a presença ou a ausência de agrupamentos naturais entre as amostras. Além disso, é possível verificar quais dos parâmetros analisados (variáveis)

são os principais responsáveis pela formação dos grupos de amostras. Ambos são classificados como métodos exploratórios não-supervisionados, visto que nenhuma informação com relação à identidade das amostras é levada em consideração. A HCA busca agrupar as amostras em classes, baseando-se na similaridade dos participantes de uma mesma classe e nas diferenças entre os membros de classes diferentes. A representação gráfica obtida é chamada de dendrograma, um gráfico bidimensional independentemente do número de variáveis do conjunto de dados. Todavia, a utilização da PCA visa reduzir a dimensionalidade do conjunto de dados original, preservando a maior quantidade de informação (variância) possível. Essa redução é obtida por meio do estabelecimento de novas variáveis ortogonais entre si, denominadas componentes principais (PC's). Organizadas em ordem decrescente de significância, as PCs são combinações lineares das variáveis originais. Os gráficos obtidos representam as amostras em um sistema cartesiano onde os eixos são as PCs. Tanto HCA quanto PCA permitem a interpretação multivariada de conjuntos de dados grandes e complexos por meio de gráficos bi ou tridimensionais. Estes gráficos apresentam informações que expressam as inter-relações que podem existir entre as variáveis, facilitando a interpretação multivariada do comportamento das amostras (Correia; Ferreira, 2007; de León-Solis; Casasola; Monterroso, 2023; Paul; de Boves Harrington, 2021).

Por outro lado, o PLS-DA (análise discriminante por mínimos quadrados parciais) se caracteriza como uma análise supervisionada, na qual utiliza um algoritmo que combina redução da dimensionalidade com a análise discriminante de forma flexível, não ajustando os dados em uma distribuição específica. A redução da dimensionalidade é alcançada pela seleção variável em vez da combinação linear realizada na PCA. Os métodos de validação do modelo permitem a construção de um modelo preditivo e podem ser agrupados em três categorias: (a) métodos internos, como validação cruzada. (b) testes externos e (c) métodos opcionais, como ensaios de permutação. A seleção de qualquer método depende do tamanho do conjunto de dados (de León-Solis; Casasola; Monterroso, 2023; Paul; de Boves Harrington, 2021). De modo geral, esses métodos são utilizados para aprimorar a separação entre os grupos identificados nas análises exploratórias não-supervisionadas. Em resumo, o emprego de métodos supervisionados visa alcançar a máxima distinção entre as amostras dos grupos e identificar as variáveis responsáveis por essa distinção (Guedes, 2018).

## 4 AVALIAÇÃO DO PERFIL METABÓLICO, MINERAL E CITOTÓXICO DE FOLHAS DE ABACAXIZEIRO DE DIFERENTES VARIEDADES COMERCIAIS: UMA NOVA FONTE ECOLOGICAMENTE CORRETA E BARATA DE COMPOSTOS BIOATIVOS

### Assessment of metabolic, mineral, and cytotoxic profile in pineapple leaves of different commercial varieties: A new eco-friendly and inexpensive source of bioactive compounds

Tamyris de Aquino Gondim, Jhonyson Arruda Carvalho Guedes, Maria Francilene Souza Silva, Adenilton Camilo da Silva, Ana Paula Dionísio, Fernanda Vidigal Duarte Souza, Claudia do Ó Pessoa, Gisele Simone Lopes, and Guilherme Julião Zocolo

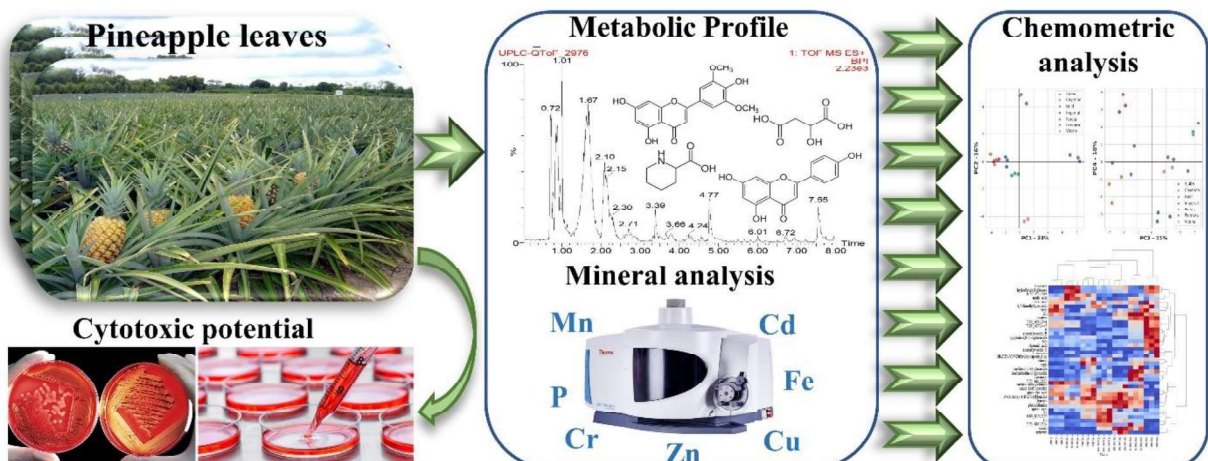
Artigo publicado em: Food Research International, Vol. 164, 2023

doi: 10.1016/j.foodres.2022.112439

Acesso via QR code:



### Graphical abstract



## ABSTRACT

Pineapple is among the most produced and consumed fruits worldwide, and consequently, its agroindustrial production/processing generates high amounts of agricultural waste, which are routinely discarded. Thus, it is crucial to seek alternatives to reuse this agricultural waste that are in high availability. Therefore, this work aims to evaluate the chemical composition of a specific residue (leaves) of seven commercial varieties of pineapples, to attribute high added value uses, and to evaluate its potential as a source of secondary metabolites and minerals. Thereby, twenty-eight metabolites were annotated by UPLC-QTOF-MS<sup>E</sup>, including amino acids, organic acids, and phenolic compounds. The following minerals were quantitatively assessed by ICP-OES: Zn (5.30 – 19.77 mg kg<sup>-1</sup>), Cr, Cd, Mn (50.80 – 113.98 mg kg<sup>-1</sup>), Cu (1.05 – 4.01 mg kg<sup>-1</sup>), P (1030.77 – 6163.63 mg kg<sup>-1</sup>) and Fe (9.06 – 70.17 mg kg<sup>-1</sup>). In addition, Cr and Cd (toxic materials) present concentration levels below the limit of quantification of the analytical method (LOQ<sub>Cr</sub> and LOQ<sub>Cd</sub> = 0.02 mg kg<sup>-1</sup>) for all samples. The multivariate analysis was conceived from the chemical profile, through the tools of PCA (principal component analysis) and HCA (hierarchical cluster analysis). The results show that pineapple leaves have similarities and differences concerning their chemical composition. In addition, the cytotoxicity assays of the extracts against tumor and non-tumor strains shows that the extracts were non-toxic. This fact can corroborate and enhance the prospection of new uses and applications of agroindustrial co-products from pineapple, enabling the evaluation and use in different types of industries, such as pharmacological, cosmetic, and food, in addition to the possibility of being a potential source of bioactive compounds.

**Keywords:** Metabolomics; *Ananas*, Chemometrics, Minerals in pineapple.

## 1 Introduction

The pineapple plant is an herbaceous monocotyledonous angiosperm that belongs to the family Bromeliaceae, genus *Ananas*, species *Ananas comosus* (L.) Merrill. Pineapple is one of the most popular tropical fruits worldwide, mostly due to its unique flavor and nutritional value. It is the only representative of the Bromeliaceae family that is widely cultivated as a food source, characterized as the most cultivated

and most economically important species of this family (Guedes *et al.*, 2018; Manetti; Delaporte; Laverde Jr., 2009; Mohd Ali *et al.*, 2020).

Native to South America, pineapple cultivation has expanded and is currently cultivated in more than eighty-two countries in different regions of the world, mainly in tropical and subtropical regions (Do *et al.*, 2020; Mohd Ali *et al.*, 2020; Rodríguez *et al.*, 2013). Based on the FAOSTAT online database (Food and Agriculture Organization of the United Nations), in 2020 world pineapple production reached almost 28 million tons with a planting area of just over 1 million ha (FAOSTAT, 2020). In this way, we can predict that tons of agricultural waste are generated simultaneously, which does not have an adequate destination.

In this context, one of the factors that contribute to increasing the tons of agricultural residues generated is the fact that the cycle of growth and production of pineapple is approximately two years. After this period, to maintain productivity, the planting must be removed completely to plant a new crop of pineapple (Chen *et al.*, 2020; Do *et al.*, 2020). In this way, the removal of agricultural residues is a critical step that can negatively impact the efficiency of pineapple productivity and must occur quickly (Chen *et al.*, 2020; Do *et al.*, 2020; Sepúlveda *et al.*, 2018). In addition, agricultural residues are sources of several different chemical species, such as pesticides, minerals, and secondary metabolites, which, depending on the circumstances, can result in water and soil contamination (Carneiro *et al.*, 2020).

Considering the continuous generation and high availability of these agricultural residues, researchers have been conducting studies to promote new uses of pineapple co-products as low-cost substrates, aiming to add value to the crop and the production of industrially important results (Kavuthodi; Sebastian, 2018; Sena Neto *et al.*, 2015, 2017).

Among the practices of elimination and/or reuse of pineapple agricultural residues, the following stand out: composting, burning, and removal before planting. The burning method is quickly and easily performed, however, it is undesirable and with a high environmental risk, such as fire outbreaks. Burning can also generate several atmospheric pollutants such as a series of incomplete combustion products. Composting takes a wide time to start, which accumulates partially decomposed waste and increases the possibility of contamination by pests and fire outbreaks (Ahmed *et al.*, 2003; Chen *et al.*, 2020; Do *et al.*, 2020). Thus, it is essential to evaluate new solutions and/or alternatives to the problem of agricultural waste, aiming to add value

and the consequent reuse of the same, based on green and sustainable technologies (Rico *et al.*, 2020).

Knowing that pineapple is recognized as a source of biologically active chemical species, we can suppose that these compounds extend to agricultural waste. In general, the chemical substances that are part of the pineapple composition are polyphenols, vitamins, flavonoids, tannins, organic acids, carbohydrates, glycosides, and proteins, possibly exercising several health benefits (Guedes *et al.*, 2018; Ma *et al.*, 2007). In addition to the compounds mentioned above, there is bromelain, a metabolite that is extracted from pineapple (especially the stem), has wide application in various food, cosmetic, and pharmaceutical industries (Difonzo *et al.*, 2019; Rodrigues *et al.*, 2020).

In recent years, studies have shown that protein components (macromolecules) of pineapple have anticancer activity. In this case, bromelain showed the ability to modulate the main pathways that support malignancy (Chakraborty *et al.*, 2021). The question would be whether the anticancer activity could be mainly related to bromelain or the effect caused by secondary metabolites (micromolecules). Therefore, we established as one of the objectives the prospection of metabolites in pineapple leaves to investigate whether these substances were also involved in anticancer activity. Because of these facts, it is of paramount importance to identify the chemical species, aiming to evaluate the waste produced to minimize environmental pollution, give new uses, and promote a high-added value to the agroindustrial residues from the pineapple plant. Under these circumstances, studies related to the chemical composition of agroindustrial pineapple residues are pretty scarce, especially concerning the mineral constitution and bioactive polyphenols that may be of great interest in the pharmaceutical, food, and cosmetics industries.

For these reasons, in this study, we performed a metabolomic workflow on the leaves of seven commercial pineapple varieties, to establish the chemical profile of these samples, enabling the determination of bioactive chemical species. The metabolic profile and mineral constitution of the samples were evaluated, respectively, by ultra-performance liquid chromatography coupled with high-resolution mass spectrometry (UPLC-HRMS) and inductively coupled plasma optical emission spectrometry (ICP-OES). The extracts from pineapple leaves used in this study were also evaluated in terms of cytotoxicity against different tumor strains and non-tumor strains. In addition, chemometric tools were used to highlight the similarities and/or

differences in the chemical composition of pineapple leaves. In this way, it was possible to assess the value of these residues generated in the fruit agroindustry.

## **2 Materials and methods**

### *2.1 Plant material*

In this study, samples of pineapple leaves (*Ananas comosus* (L.) Merr.) from seven different commercial cultivars were evaluated: Golden (MD2), Perolera (PE, AGB 049), Pérola (AGB 001), BRS Imperial, Smooth Cayenne (SC), BRS Vitória, and BRS Ajubá.

Samples of the cultivars Perolera and Pérola were collected from the AGB-Pineapple (Active Germplasm Bank of Pineapple (with more than 700 accessions, 10 plants/accession), Embrapa Cassava & Fruits, Cruz das Almas, BA, Brazil), while samples of the cultivars BRS Imperial, Golden, BRS Imperial, Smooth Cayenne, BRS Vitória, and BRS Ajubá were collected in the Experimental Fields of Embrapa Cassava and Fruits (Embrapa Cassava & Fruits, Cruz das Almas, BA, Brazil).

Samples were collected from ten different plants of each cultivar, two leaves of each plant totaling twenty leaves of each cultivar. Then, the leaves were dried in a convection oven at 40 °C for 72 h. Thereafter, the samples were ground and stored, at 28°C.

### *2.2 Reagents and chemicals*

The ultrapure water used for the mobile phase in the chromatographic analyzes and for the preparation of all the solutions necessary for the development of the work was obtained by Milli-Q system (Millipore, Bedford, MA, USA), formic acid (purity 98%), acetonitrile (LC-MS grade) was supplied by Tedia (Fairfield, Ohio, EUA), hexane (95%) and ethanol (96%) were purchased from the Tedia (Rio de Janeiro, RJ, Brazil), dimethyl sulfoxide (DMSO) and 3-(4,5- dimethyl-2-thiazol)-2,5-diphenyl-2H-tetrazolium bromide (MTT) were purchased from Sigma-Aldrich (Life Science); the tumor cells used, HCT-116 (human colon carcinoma), HL60 (leukemia), PC3 (prostate), MCF-7 (breast), SNB-19 (astrocytoma), HeLa (cervix) and L929 (mouse fibroblast, non-tumor) were donated by NCI-USA (National Cancer Institute), cultured



in RPMI 1640 medium, supplemented with 10% fetal bovine serum and 1% antibiotics, kept in an oven at 37 °C and in an atmosphere containing 5% CO<sub>2</sub>, nitric acid (65%) and hydrofluoric acid (38%) from Vetec (Rio de Janeiro, RJ, Brazil), hydrochloric acid (37%) and boric acid from Sigma-Aldrich (St. Louis, Missouri, EUA). Standard solutions were prepared from stock solutions (1000 mg L<sup>-1</sup>) of Cd, Cr, Cu, Fe, Mn, P, and Zn supplied by Acros Organics (Geel, Belgium).

### 2.3 Chemical profiling by UPLC-QTOF-MS<sup>E</sup>

#### 2.3.1 Sample preparation for UPLC analysis

The pineapple leaf extracts were obtained using a microextraction sample preparation method adapted from the literature (Chagas-Paula *et al.*, 2015; Guedes *et al.*, 2020; Nehme *et al.*, 2008). A portion of 50 mg of dried, ground, and homogenized plant material was weighed. Thereafter, 4 mL of hexane was added and vortexed for 1 min, then the sample contained in the test tube was submitted to the ultrasonic bath (fixed power of 135 W) for 20 min. Posteriorly, 4 mL of ethanol: water (7:3) solution was added, vortexed again for 1 min, and placed in the ultrasound bath for 20 min. Finally, for a complete separation of the hexane and hydroethanolic phases, the test tube containing the mixture was centrifuged for 10 min. Afterward, a 1 mL aliquot was removed from the hydroethanolic phase, filtered (0.22 µm PTFE filter), and added to vials. Subsequently, the extracts were analyzed by UPLC-QTOF-MS<sup>E</sup>.

The extraction of samples was done in quintuplet for the seven different commercial cultivars of the pineapple and five extraction blank, *N* = 40 extractions.

#### 2.3.2 Chromatographic conditions

The chromatographic separation was performed on an Acquity UPLC (Waters Corp., Milford, MA, USA) coupled to a quadrupole/time of flight (QTOF, Waters). Chromatographic runs were conducted through Waters Acquity UPLC BEH (150 mm x 2.1 mm, 1.7 µm), at a fixed temperature of 40 °C. The analysis was carried out by applying the following binary gradient (A (water containing 0.1% formic acid) and B (acetonitrile containing 0.1% formic acid)) at a flow rate of 0.4 mL min<sup>-1</sup>: 0.0–15.0 min, linear gradient from 2 to 95% B; 15.1–17.0 min, 100% B; 17.1 min, 2% B;

17.02-20.0 min, 2% B gradient. The sample injection volume was 3  $\mu\text{L}$ .

### 2.3.3 Mass spectrometry conditions and metabolite annotation

Sample analyzes were performed by a chromatographic system coupled to a mass spectrometer, interfaced by an electrospray ionization source (UPLC-ESI-QTOF-MS<sup>E</sup>). Analyzes were conducted in ESI<sup>-</sup> and ESI<sup>+</sup> ionization modes, with an acquisition range of 110-1200 in MS and 50-1200 in MS<sup>2</sup>. The desolvation gas flow was 500 L h<sup>-1</sup> (ESI<sup>-</sup>) and 350 L h<sup>-1</sup> (ESI<sup>+</sup>). The temperature of the ionization source and the desolvation gas were 120 °C and 350 °C, respectively. Leucine enkephalin was used as a lock mass. The capillary voltage was 3 kV.

The data obtained from the UPLC-ESI-QTOF-MS<sup>E</sup> was submitted to the software MS-DIAL 4.92 (Data Independent Analysis), aiming to establish the necessary parameters for untargeted metabolomics: deconvoluted spectra, peak alignment, and filtering (Lai *et al.*, 2018; Tsugawa *et al.*, 2015, 2019). Posteriorly, the unidentified metabolites can be annotated through the MS-FINDER 3.52 (Lai *et al.*, 2018; Tsugawa *et al.*, 2015). The annotation of the metabolites was conceived through the MS and MS/MS mass spectra, where the probable molecular formulas and the fragmentations were obtained, respectively. After obtaining the MS/MS spectra, the metabolite annotation was performed, comparing the data with the information from the database such as the KNApSACk Core System database, Human Metabolome Database (HMDB), Kyoto encyclopedia of genes and genome database (KEGG), SciFinder, ChemSpider, and PubChem. The annotation of metabolites was performed according to the guidelines established by MSI (Metabolic Standards Initiative) level 2.1 (Sumner *et al.*, 2007). In addition, the annotation includes the molecular formulas and the fragment ions correlated to the metabolites. The annotation of metabolites was performed considering the chemotaxonomy (family, genus, and species).

### 2.4 Mineral analysis by ICP-OES

An iCAP 6000 ICP-OES (Thermo Scientific, USA) equipped with a concentric nebulizer was used for elemental analysis. ICP operational parameters were as follows: 12 L min<sup>-1</sup> cooling gas, 1150 W RF power; 0.45 L min<sup>-1</sup> nebulizer Ar gas flow rate and sample uptake rate of 1.4 mL min<sup>-1</sup>, 0.5 L min<sup>-1</sup> auxiliary Ar gas flow

rate; The analytical wavelengths (nm) chosen were: Cd (214.4), Cr (283.5), Cu (327.3), Fe (259.9), Mn (259.3), P (177.4), Zn (219.8). The standard solutions of the analytes (Cd, Cr, Cu, Mn, Fe, P, Zn) were prepared by dilutions of 1000 mg L<sup>-1</sup> stock solutions.

The sample preparation method was used for a Tecnal TE-007D digester heating block (Piracicaba, Brazil). An accurately weighed 0.200 g of the pineapple leaves, dried, ground, and homogenized were placed into Teflon® tubes with 5 mL of a mixture of HNO<sub>3</sub> and HCl (3:1). The mixture was left overnight. Then, 2 mL of HF was added, and the tubes were placed in a heated digester block at 130 °C for 5 h. The resultant solutions were neutralized by 20 mL of boric acid (4%), filtered, and analyzed by ICP-OES. Certified reference material Tomato Leaves SRM 1573a (NIST, Gaithersburg, USA) was used to validate the accuracy of the proposed methodology. A one-way analysis of variance (ANOVA) and Tukey's test were conducted using a 95% confidence level.

### *2.5 Determination of cytotoxic potential in vitro*

Cytotoxicity evaluation was performed against different tumor cell lines: HCT-116, HL60, PC3, SNB-19, MCF-7, and HeLa. In addition, the evaluation of the extracts in terms of cellular selectivity, between healthy cells and cancer cells, was performed using the L929 (mouse fibroblast, non-tumor) cell line, where the cells were exposed to the extract for 72 h. Cells were cultured in RPMI 1640 medium, L929 was cultured in DMEM with Earle's salts, complemented with 100 IU mL<sup>-1</sup> penicillin, 10% fetal bovine serum, 2 mmol L<sup>-1</sup> L-glutamine, 100 mg mL<sup>-1</sup> streptomycin at 37 °C with 5% CO<sub>2</sub>.

In the determination of the cytotoxic potential, the cells were added to 96-well plates. The dry hydroethanolic extracts were dissolved with DMSO; the final concentration of DMSO in the culture medium was kept constant (0.1%, v/v). The cell viability was determined as Mosmann (1983) described with the method MTT. After 72 h, the plates were subjected to centrifugation and the medium was replaced with fresh medium (200 mL) containing 0.5 mg mL<sup>-1</sup> of MTT. Then, after 3 h, the MTT formazan was solubilized in DMSO (150 µL) and absorbance readings (570 nm) were taken by DTX 880 Multimode Detector (Beckman Coulter, Inc. Fullerton, California, USA).

The evaluation of the samples was carried out in triplicate and the results regarding the percentage of inhibition of cell growth (GI, %) were obtained using the

GraphPad Prism® 5.0 software.

## *2.6 Data processing and statistical analysis*

In this work, Google Colaboratory (<https://colab.research.google.com>) and pre-installed libraries were used. This software can be used free of charge as Google provides it for search activities using a 12GB Tesla K80 GPU and python language. The analysis and manipulation of the database were carried out with the PANDAS library (<https://pandas.pydata.org/>). Two pattern recognition strategies were applied: principal component analysis (PCA) and hierarchical cluster analysis (HCA). Both algorithms were performed using the SKLEARN library (<https://scikit-learn.org/>) and SCIPY (<https://www.scipy.org/scipylib/index.html>). The graphs were plotted using the graphic libraries: SEABORN (<https://seaborn.pydata.org/>) and MATPLOTLIB (<https://matplotlib.org/>).

The data sets were combined in the direction of the columns. Since the numerical value between variables differs markedly, direct comparison between variables is not feasible. Because of this, the auto-scaling transformation was applied. This transformation keeps the statistical information of the data set. Still, each variable begins to show zero mean and variance equal to one, correcting weighting effects arising from the nature of the data set. The PCA was applied with data mean centering and dimension reduction using singular value decomposition (SVD) to investigate the possible similarities and dissimilarities between the leaves of commercial pineapple varieties. A cluster analysis study was also carried out using a hierarchically grouped heat map with Euclidean distance and Ward's connection method. The contribution of the variables was standardized between zero and one (subtracting all of them by the minimum value followed by the division by the highest value found).

## **3 Results and discussion**

### *3.1 Non-targeted analysis of pineapple leaves by UPLC-QTOF-MS<sup>E</sup>*

The evaluation and determination of the metabolic profile of the seven varieties of pineapple leaves were performed using a non-targeted approach. For this purpose, the extracts obtained through microextraction were analyzed by UPLC-

QTOF-MS<sup>E</sup>, and mass spectra were acquired in positive (ESI<sup>+</sup>) and negative (ESI<sup>-</sup>) ionization modes. The evaluation of the chromatograms and the respective mass spectra was performed with the aid of the previously reported literature review based on the research of substances found in the family, genus, and species of the plant under study (Chemotaxonomy). In addition, SciFinder and other databases available in MS-DIAL and MS-FINDER software were consulted.

A total of 28 metabolites were annotated (Table 2) in the seven varieties of pineapple leaves evaluated. The annotated metabolites, based on chemical characteristics, can be summarily divided into organic acids (and their derivatives), amino acids (and their derivatives), and flavonoids in general.

In summary, it was possible to observe that the metabolic profiles of all samples show quite similarities. In addition, through the analysis of the chromatograms (Figure 10), it is possible to verify that the most significant dissimilarities between the commercial varieties of pineapple are due to different levels of concentration of metabolites in the seven commercial varieties of pineapple leaves.

### 3.1.1 Amino acids and their derivatives

The mass spectra for compound 8 suggest that the metabolite is *L*-tyrosine since the precursor ion at  $m/z$  182.0821 [M+H]<sup>+</sup> and the presence of fragment ion  $m/z$  165.0489 [(M+H)-NH<sub>3</sub>]<sup>+</sup> from the elimination of NH<sub>3</sub> of the protonated molecule [M+H]<sup>+</sup> was observed (Difonzo *et al.*, 2019; Yang *et al.*, 2018).

Tabela 2 – Table 2. Metabolites annotated in the samples of seven varieties of pineapple leaves, in negative (ESI<sup>-</sup>) and positive (ESI<sup>+</sup>) modes. (to be continued).

Peak	t <sub>R</sub> (min)	ESI <sup>-</sup>				ESI <sup>+</sup>				Molecular formula	Peak annotation	Sample*	References
		MS calculated	MS [M-H] <sup>-</sup>	MS/MS	Error (ppm)	MS calculated	MS [M+H] <sup>+</sup>	MS/MS	Error (ppm)				
1	0.72	-	-	-	-	413.2117	413.2098	-	-4.6	C <sub>28</sub> H <sub>28</sub> O <sub>3</sub>	not identified	C; Pra; G; Pla; V; A	-
2	0.88	341.1084	341.1061	89.0248 179.0485	-6.7	-	-	-	-	C <sub>12</sub> H <sub>22</sub> O <sub>11</sub>	sucrose	C; I; Pra; G; V; A	(Farag <i>et al.</i> , 2014)
3	0.89	135.0293	135.0291	90.0197 72.9977	-1.5	-	-	-	-	C <sub>4</sub> H <sub>8</sub> O <sub>5</sub>	threonic acid	C; I; Pra; G; Pla; V; A	(Maulidani <i>et al.</i> , 2019)
4	0.89	191.0556	191.0540	85.0276 101.0212	-8.4	-	-	-	-	C <sub>7</sub> H <sub>12</sub> O <sub>6</sub>	quinic acid	C; I; Pra; G; Pla; V; A	(Baskaran, Pullencheri, Somasundaram, 2016)
5	0.93	295.0665	295.0688	133.0150 115.0047	7.8	-	-	-	-	C <sub>10</sub> H <sub>16</sub> O <sub>10</sub>	malic acid hexoside	C; I; Pra; G; V; A	(Abu-Reidah <i>et al.</i> , 2015)
6	0.94	133.0137	133.0129	115.0029 89.0244	-6.0	-	-	-	-	C <sub>4</sub> H <sub>6</sub> O <sub>5</sub>	malic acid	C; I; Pra; G; Pla; V; A	(Abu-Reidah <i>et al.</i> , 2015; Oldoni <i>et al.</i> , 2019; Sun <i>et al.</i> , 2016)
7	1.01	191.0192	191.0189	111.0086 87.0109	-1.6	-	-	-	-	C <sub>6</sub> H <sub>8</sub> O <sub>7</sub>	citric acid	C; I; Pra; G; Pla; V; A	(Oldoni <i>et al.</i> , 2019; Rodríguez- Pérez <i>et al.</i> , 2015; Sun <i>et al.</i> , 2016)
8	1.01	-	-	-	-	182.0817	182.0821	165.0489 136.0611	2.2	C <sub>6</sub> H <sub>11</sub> NO <sub>3</sub>	L-tyrosine	C; I; Pra; Pla; V; A	(Difonzo <i>et al.</i> , 2019; Yang <i>et al.</i> , 2018)
9	1.03	-	-	-	-	130.0868	130.0860	84.0796	-6.1	C <sub>8</sub> H <sub>11</sub> NO <sub>2</sub>	pipecolic acid	C; I; Pra; G; V; A	(Joo <i>et al.</i> , 2020)

Tabela 2 – Table 2. Metabolites annotated in the samples of seven varieties of pineapple leaves, in negative (ESI<sup>-</sup>) and positive (ESI<sup>+</sup>) modes. (continuation).

Peak	t <sub>R</sub> (min)	ESI <sup>-</sup>				ESI <sup>+</sup>				Molecular formula	Peak annotation	Sample*	References
		MS calculated	MS [M-H] <sup>-</sup>	MS/MS	Error (ppm)	MS calculated	MS [M+H] <sup>+</sup>	MS/MS	Error (ppm)				
10	1.69	-	-	-	-	294.1558	294.1553	276.1422 258.1338 230.1355 212.1274	-1.7	C <sub>12</sub> H <sub>23</sub> NO <sub>7</sub>	N-(1-deoxy-1-fructosyl)leucine	C; I; Pra; G; Pla; V; A	(Azi <i>et al.</i> , 2020; Yang <i>et al.</i> , 2018)
11	1.69	-	-	-	-	132.1025	132.1025	86.0952 69.0337	0.0	C <sub>9</sub> H <sub>13</sub> NO <sub>2</sub>	leucine	C; I; Pra; G; Pla; V; A	(Yang <i>et al.</i> , 2018)
12	2.10	164.0712	164.0702	147.0462	-6.1	166.0868	166.0873	120.0802 103.0512 91.0528 292.1194	3.0	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	phenylalanine	C; I; Pra; G; Pla; V; A	(Jandrić <i>et al.</i> , 2014)
13	2.11	326.1240	326.1244	164.0693 147.0464 103.0521	1.2	328.1396	328.1408	264.1126 132.0810 120.0802	3.7	C <sub>15</sub> H <sub>21</sub> NO <sub>7</sub>	fructose-phenylalanine	C; I; Pra; G; Pla; V; A	(Rodríguez-Pérez <i>et al.</i> , 2018)
14	3.38	593.1506	593.1530	353.0650 383.0734 473.1075	4.0	595.1663	595.1666	577.1619 559.1481 457.1155	0.5	C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>	apigenin-6,8-C-diglucoside	C; Pra; G; Pla; V; A	(Dueñas <i>et al.</i> , 2021; Song <i>et al.</i> , 2019; Steingass <i>et al.</i> , 2015; Tsolmon <i>et al.</i> , 2020)
15	3.66	593.1506	593.1523	269.0429 311.0526 431.1013	2.9	-	-	-	-	C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>	isovitexin-7-O-glucoside or isovitexin-6"-O-glucoside	C; Pra; G; Pla; V; A	(Dueñas <i>et al.</i> , 2021; Mizuno <i>et al.</i> , 2021)
16	3.77	431.0978	431.0976	341.0677 311.0543	0.5	433.1135	433.1124	313.0741	-2.5	C <sub>21</sub> H <sub>20</sub> O <sub>10</sub>	vitexin (apigenin-8-C-glucoside)	C; Pra; G; Pla; V; A	(Dueñas <i>et al.</i> , 2021; Ren <i>et al.</i> , 2020; Yang <i>et al.</i> , 2018; Zhang <i>et al.</i> , 2016)

Tabela 2 – Table 2. Metabolites annotated in the samples of seven varieties of pineapple leaves, in negative (ESI<sup>-</sup>) and positive (ESI<sup>+</sup>) modes. (continuation).

Peak	t <sub>R</sub> (min)	ESI <sup>-</sup>				ESI <sup>+</sup>				Molecular formula	Peak annotation	Sample*	References
		MS calculated	MS [M-H] <sup>-</sup>	MS/MS	Error (ppm)	MS calculated	MS [M+H] <sup>+</sup>	MS/MS	Error (ppm)				
17	4.10	609.1456	609.1491	300.0254 301.0279	5.7	611.1612	611.1605	303.0507	-1.1	C <sub>27</sub> H <sub>30</sub> O <sub>16</sub>	quercetin-3-O-rutinoside (rutin)	C; I; Pra; Pla; V; A	(Guedes <i>et al.</i> , 2020; Kumar <i>et al.</i> , 2015; Manetti; Delaporte; Laverde Jr., 2009)
18	4.19	431.0978	431.0971	283.0604 341.0652 311.0564	1.6	433.1135	433.1127	283.0607 313.0685	-1.8	C <sub>21</sub> H <sub>20</sub> O <sub>10</sub>	isovitexin (apigenin-6-C-glucoside)	C; Pra; Pla; V; A	(Dueñas <i>et al.</i> , 2021; Mizuno <i>et al.</i> , 2021; Zhang <i>et al.</i> , 2016)
19	4.49	-	-	-	-	498.2609	498.2631	481.2277 322.2073 234.1161 177.0670	5.4	C <sub>27</sub> H <sub>35</sub> N <sub>3</sub> O <sub>6</sub>	(di- <i>E,E</i> )- <i>N,N'</i> -Diferuloyl-Ispermidine	C; I; Pra; G; Pla; V; A	(Difonzo <i>et al.</i> , 2019; Steingass <i>et al.</i> , 2015)
20	4.54	-	-	-	-	479.1190	479.1210	317.0658	4.2	C <sub>22</sub> H <sub>22</sub> O <sub>12</sub>	isorhamnetin-O-glycoside	C; I; Pra; Pla; V; A	(Frag <i>et al.</i> , 2020; Matos <i>et al.</i> , 2021) (Pacheco <i>et al.</i> , 2019;
21	4.77	503.2492	503.2495	371.1990 209.1148	0.6	-	-	-	-	C <sub>24</sub> H <sub>40</sub> O <sub>11</sub>	leeaoside	C; I; Pra; Pla; V; A	Silva, M. de F. G. da <i>et al.</i> , 2018; Zhang, W. <i>et al.</i> , 2015)
22	4.77	-	-	-	-	373.1135	373.1145	211.1686 193.1587 175.1465 135.1159	2.7	C <sub>16</sub> H <sub>20</sub> O <sub>10</sub>	hydroferuloyl-glucose	C; I; Pra; G; Pla; V; A	(Ma <i>et al.</i> , 2007; Papetti <i>et al.</i> , 2014)
23	4.81	-	-	-	-	523.1452	523.1469	361.0890	3.2	C <sub>24</sub> H <sub>26</sub> O <sub>13</sub>	ananaflavoside C	Pla; A	(Ma <i>et al.</i> , 2007)
24	5.17	-	-	-	-	345.0974	345.0982	330.0779	2.3	C <sub>18</sub> H <sub>16</sub> O <sub>7</sub>	santin	C; I; Pra	(Yan <i>et al.</i> , 2019)



Tabela 2 – Table 2. Metabolites annotated in the samples of seven varieties of pineapple leaves, in negative (ESI<sup>-</sup>) and positive (ESI<sup>+</sup>) modes. (continuation).

Peak	t <sub>R</sub> (min)	ESI <sup>-</sup>				ESI <sup>+</sup>				Molecular formula	Peak annotation	Sample*	References
		MS calculated	MS [M-H] <sup>-</sup>	MS/MS	Error (ppm)	MS calculated	MS [M+H] <sup>+</sup>	MS/MS	Error (ppm)				
25	5.65	373.0923	373.0908	329.0643 343.0411 358.0723	4.0	375.1080	375.1068	342.0729	-3.2	C <sub>19</sub> H <sub>16</sub> O <sub>6</sub>	casticin	C; I; Pra; G; Pla; V; A	(Fu <i>et al.</i> , 2020; Högner <i>et al.</i> , 2013) (Ma <i>et al.</i> , 2007;
26	6.01	-	-	-	-	537.1608	537.1600	375.1089	-1.5	C <sub>25</sub> H <sub>28</sub> O <sub>13</sub>	ananaflavoside B	C; I; Pra; G; Pla; V; A	Steingass <i>et al.</i> , 2015) (Brito <i>et al.</i> , 2020; Maia <i>et al.</i> , 2020;
27	6.37	329.0661	329.0680	299.0188 314.0463	2.4	-	-	-	-	C <sub>17</sub> H <sub>14</sub> O <sub>7</sub>	3,7-dimethylquercetin	C; Pra; G; Pla; A	Quifer-Rada <i>et al.</i> , 2015) (Fu <i>et al.</i> , 2008;
28	6.61	269.0450	269.0440	151.1072	-3.7	-	-	-	-	C <sub>15</sub> H <sub>10</sub> O <sub>5</sub>	apigenin	I	Manetti; Delaporte; Laverde Jr., 2009) (Ma <i>et al.</i> , 2017; Moheb <i>et al.</i> , 2013;
29	6.70	329.0661	329.0655	299.0187 271.0226 243.0357	-1.8	331.0819	331.0815	331.0800 315.0529 270.0568	-0.9	C <sub>17</sub> H <sub>14</sub> O <sub>7</sub>	tricin	C; I; Pra; G; Pla; V; A	Rui <i>et al.</i> , 2010)
30	6.80	819.2406	819.2399	-	-0.9	-	-	-	-	C <sub>31</sub> H <sub>48</sub> O <sub>25</sub>	not identified	C; I; Pra; G; Pla; A	-
31	7.55	-	-	-	-	455.3525	455.3541	-	3.5	C <sub>30</sub> H <sub>48</sub> O <sub>3</sub>	not identified	C; I; Pra; G; Pla; V; A	-
32	7.55	-	-	-	-	437.3478	437.3449	-	-6.6	C <sub>23</sub> H <sub>48</sub> O <sub>7</sub>	not identified	C; I; Pra; G; Pla; V; A	-
33	7.73	861.2512	861.2523	-	1.3	-	-	-	-	C <sub>33</sub> H <sub>50</sub> O <sub>26</sub>	not identified	C; I; Pra; G; Pla; V; A	-

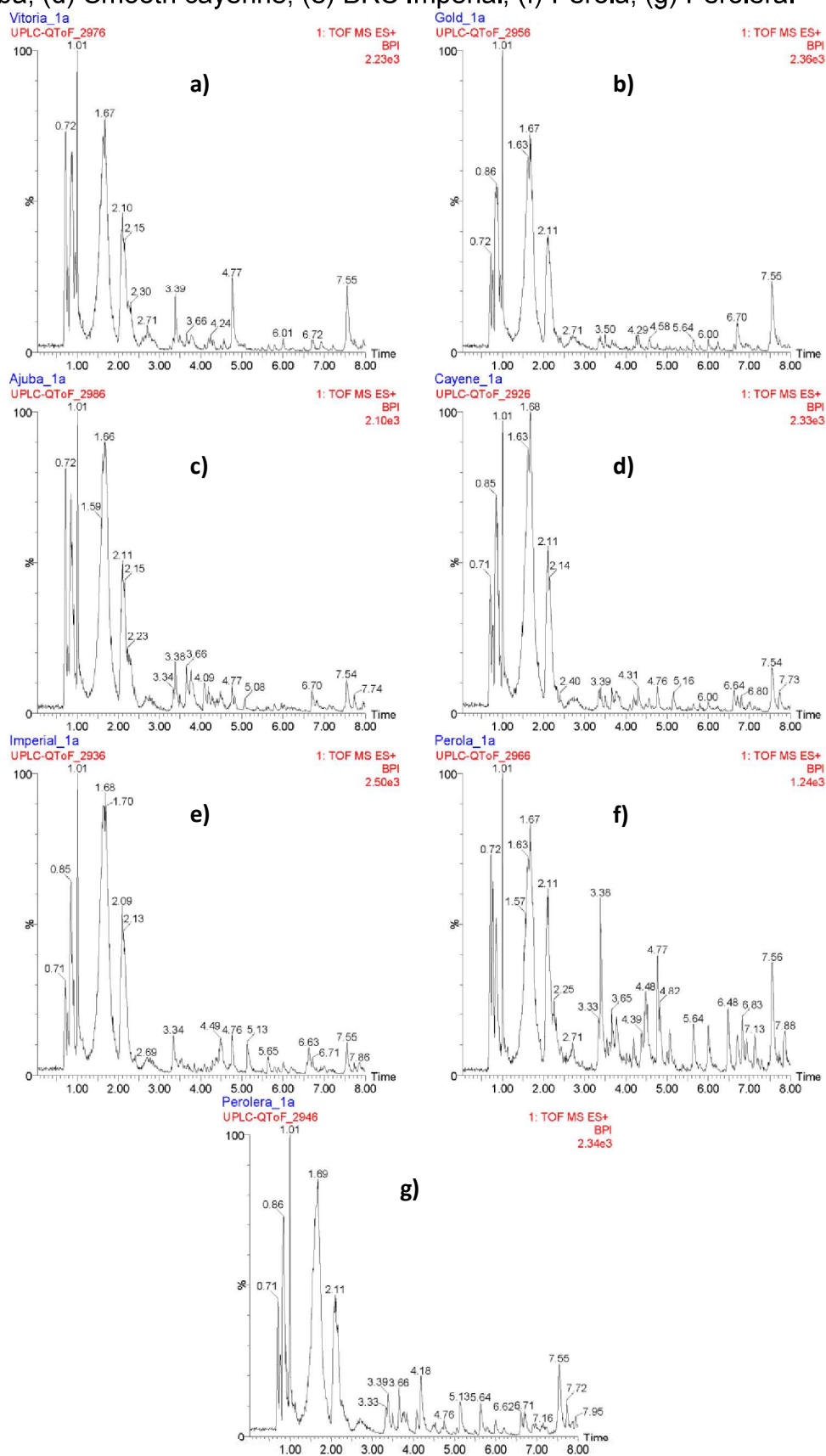
Tabela 2 – Table 2. Metabolites annotated in the samples of seven varieties of pineapple leaves, in negative (ESI<sup>-</sup>) and positive (ESI<sup>+</sup>) modes. (conclusion).

Peak	t <sub>R</sub> (min)	ESI <sup>-</sup>				ESI <sup>+</sup>				Molecular formula	Peak annotation	Sample*	References
		MS calculated	MS [M-H] <sup>-</sup>	MS/MS	Error (ppm)	MS calculated	MS [M+H] <sup>+</sup>	MS/MS	Error (ppm)				
34	7.73	-	-	-	-	885.2512	885.2504	-	-0.9	C <sub>35</sub> H <sub>48</sub> O <sub>26</sub>	not identified	C; I; Pra; G; Pla; V; A	-

Source: Prepared by the Author.

\*C: Smooth Cayenne; I: BRS Imperial; Pra: Perolera; G: Gold; Pla: Pérola; V: BRS Vitória; A: BRS Ajubá

Figura 10 – Figure 10. Representative chromatograms (ESI<sup>+</sup>) of the analysis of the seven commercial varieties of pineapple leaves: (a) BRS Vitória; (b) Gold; (c) BRS Ajubá; (d) Smooth cayenne; (e) BRS Imperial; (f) Pérola; (g) Perolera.



Source: Prepared by the Author.

At peaks 10 and 11, they showed a leucine derivative; *N*-(1-deoxy-1-fructosyl) leucine, ( $m/z$  294.1553  $[M+H]^+$ ) and leucine ( $m/z$  132.1025  $[M+H]^+$ ), respectively, were observed. In the MS/MS mass spectrum of the leucine derivative, it is possible to verify the presence of the fragment ion  $m/z$  276.1422, due to the loss of a water molecule from the precursor ion (Azi *et al.*, 2020).

Metabolites 12 and 13 were observed in both ionization modes (ESI<sup>-</sup> and ESI<sup>+</sup>), annotated as phenylalanine and fructose-phenylalanine, respectively. Phenylalanine exhibited the precursor ion  $m/z$  164.0702  $[M-H]^-$ , with fragment ion  $m/z$  147.0462. On the other hand, fructose-phenylalanine presented deprotonated ion  $m/z$  326.1244, with fragment ions  $m/z$  164.0693  $[(M-H)-162]^-$  and 147.0464  $[(M-H)-162-17]^-$ , which are due to the cleavage of the glycosidic bond followed by successive loss of a hydroxyl group (Rodríguez-Pérez *et al.*, 2018).

### 3.1.2 Organic acids and derivatives

Some organic acids were annotated: threonic acid, quinic acid, malic acid hexoside, malic acid, citric acid, and pipercolic acid.

Peak 3 exhibited the precursor ion at  $m/z$  135.0291  $[M-H]^-$  and fragment ions at  $m/z$  90.0197 and 72.9977. Therefore, the metabolite was annotated as threonic acid (Maulidiani *et al.*, 2019).

The metabolites 4 and 7 showed the precursor  $m/z$  ions 191.0540 and 191.0189, respectively. Based on this and the joint analysis with the fragment ions, mainly  $m/z$  101.0212  $[M-CO_2-H_2O-CO]^-$  (peak 4) and  $m/z$  111.0086  $[(M-H)-CO_2-2H_2O]^-$  (peak 7) they were annotated, respectively, as quinic acid and citric acid (Baskaran; Pullencheri; Somasundaram, 2016; Rodríguez-Pérez *et al.*, 2015).

Peaks 5 and 6 showed the precursor ions  $[M-H]^-$  at  $m/z$  295.0688 and 133.0129, respectively. Thus, the metabolites were annotated respectively as malic acid hexoside and malic acid. Corroborating with the characterization of malic acid and its glycoside derivative, we verified the formation of fragment ions  $m/z$  89.0244, 115.0029, and 133.0150. In which the formation of the fragment ion  $m/z$  115  $[(M-H)-H_2O]^-$  is due to the loss of a water molecule. On the other hand, the fragment ion  $m/z$  133  $[(M-H)-162]^-$  results from the loss of the hexoside group (Abu-Reidah *et al.*, 2015; Oldoni *et al.*, 2019; Sun *et al.*, 2016). Compound 9 was annotated as pipercolic acid with a signal at  $m/z$  130.0860  $[M+H]^+$  in the MS spectrum and at  $m/z$  84.0796 in the

MS/MS spectrum (Joo *et al.*, 2020).

### 3.1.3 Flavonoids

The flavonoid apigenin is classified as a flavone, which is an aglycone of several glycosylated metabolites, being present in several plant organisms. Thus, in this work, from the evaluation of pineapple leaves extracts, it was possible to annotate the aglycone and four other metabolites derived from this flavone. Compound 28 showed a precursor ion at  $m/z$  269.0440  $[M-H]^-$  and fragment ions at  $m/z$  151.1072  $[(M-H)-C_8H_6O]^-$ , being noted as the aglycone apigenin. Compounds 14 and 15 have been identified as two isomers  $m/z$  593  $[M-H]^-$ , in which they are two glycosylated derivatives of apigenin. Thus, peak 14, observed in the ESI<sup>-</sup> and ESI<sup>+</sup>, showed a precursor ion  $[M+H]^+$  at  $m/z$  595.1666 with fragmentation patterns characteristic of the apigenin-6,8-*C*-diglucoside  $m/z$  577.1619  $[(M+H)-18]^+$ ,  $m/z$  559.1481  $[(M+H)-30]^+$  and  $m/z$  457.1155  $[(M+H)-120-18]^+$  (Moraes; Tomaz; Lopes, 2007).

On the other hand, peak 15 was observed only in ESI<sup>-</sup> with deprotonated ion  $[M-H]^-$   $m/z$  593.1523 and fragment ions  $m/z$  269.0429, 311.0526, and 431.1013. This fragmentation pattern corresponds to either isovitexin-7-*O*-glucoside or isovitexin-6''-*O*-glucoside, since in the fragmentation of the substituted flavones in aglycone hydroxyl groups, it is common to observe the fragment ion resulting from the loss of a glucose moiety  $[M-H-162]^-$ . Thus, the fragment ion  $m/z$  431.1013  $[M-H-162]^-$  comes from this cleavage, where this fragment ion corresponds to isovitexin (Dueñas *et al.*, 2021). In addition, the mass spectrum shows the fragment ion  $m/z$  269.0429  $[M-H-162-162]^-$  arising from the successive losses of two units of glucose. The fragment ion  $m/z$  311.0526  $[M-H-162-120]^-$  corresponds to glucose loss, and 120 Da resulted from internal cleavages between the aglycone and the glycoside at the *C*-glycosylation position.

In the mass spectra of the peaks 16 and 18, it was found that these are two isomers (C<sub>21</sub>H<sub>20</sub>O<sub>10</sub>) which were observed in ESI<sup>-</sup> ( $m/z$  431.0976) and ESI<sup>+</sup> ( $m/z$  433.1124). The analysis of the fragment ions in MS/MS shows that they are characteristic of *C*-glycosidic metabolites. The fragment ions present in ESI<sup>-</sup> ( $m/z$  341, 311, and 283) are due to neutral losses of 90, 120, and 148 mass units, in addition to ESI<sup>+</sup> mode:  $m/z$  313  $[(M+H)-C_4H_8O_4]^+$  and  $m/z$  283  $[(M+H)-C_4H_8O_4-CH_2O]^+$ . Therefore, corroborates the annotation of two metabolites, *C*-glycosidic flavone, vitexin

(compound 16), and isovitexin (compound 18) (Dueñas *et al.*, 2021).

Compounds 17 and 27 were annotated as two quercetin derivatives. Regarding metabolite 17, there is quercetin linked to a rutinoside, where, more specifically, there is quercetin with the hydroxy group at the C-3 position substituted with glucose and rhamnose sugar groups. In this context, this metabolite was observed in both ESI<sup>-</sup> ( $m/z$  609.1491) and ESI<sup>+</sup> ( $m/z$  611.1605) ionization modes. In the MS/MS mass spectrum, referring to the deprotonated ion, the fragment ion  $m/z$  301.0279 is presented. This ion is formed from the loss of a rutinoside unit (308 Da), thus characterizing the metabolite in question as quercetin-3-O-rutinoside (rutin) (Guedes *et al.*, 2020). On the other hand, peak 27 was annotated as 3,7-dimethylquercetin, a methoxylated flavonoid. In view, that showed the precursor ion [M-H]<sup>-</sup> at  $m/z$  329.0680 and fragment ions  $m/z$  314.0463 [(M-H)-CH<sub>3</sub>]<sup>-•</sup> and 299.0188 [(M-H)-CH<sub>3</sub>-CH<sub>3</sub>]<sup>-•</sup>, which are due to the loss of methyl radical groups (Brito *et al.*, 2020; Quifer-Rada *et al.*, 2015).

Peaks 23 and 26 correspond to two flavones that were observed in the ESI<sup>+</sup> mode. The mass spectra showed the protonated precursor ions at  $m/z$  523.1469 and 537.1600. In addition, fragment ions  $m/z$  361.0890 and 375.1089 were found in MS/MS, respectively, for metabolites 23 and 26. The fragment ions in both metabolites are derived from losses of a hexoside sugar unit (162 Da), corroborating the annotation of the ananaflavoside C and B, respectively.

The mass spectra of peaks 24 and 25 show similar fragmentation patterns. Thus, it is possible to infer that these two are methoxy flavones through the analysis of the mass spectra. In the case of metabolite 24, the precursor ion at  $m/z$  345.0982 exhibits fragment ion at  $m/z$  330.0779 from the loss of radical •CH<sub>3</sub>. Therefore, the compound consists of a trimethoxy flavone, santin. As previously mentioned, metabolite 25 has a fragmentation pattern similar to metabolite 24, thus analyzing the precursor ion [M-H]<sup>-</sup> at  $m/z$  373.0908 together with the fragment ions at  $m/z$  343.0411 [(M-H)-2CH<sub>3</sub>]<sup>-•</sup> and 358.0723 [(M-H)-CH<sub>3</sub>]<sup>-•</sup>, compound 25 has been noted as casticin, in which it is chemically classified as tetramethoxyflavone (Huang *et al.*, 2015).

Metabolite 29 also showed precursor ions  $m/z$  329.0655 and fragment ions in  $m/z$  299.0187 [(M-H)-2CH<sub>3</sub>]<sup>-•</sup>,  $m/z$  271.0226 [(M-H)-C<sub>2</sub>H<sub>2</sub>O<sub>2</sub>]<sup>-</sup> and  $m/z$  243.0357 [(M-H)-C<sub>4</sub>H<sub>6</sub>O<sub>2</sub>]<sup>-</sup> consistent with a dimethoxyflavone, and may also be classified as an O-methylated flavone, in which it was noted as tricin (Li *et al.*, 2016; Rui *et al.*, 2010).

### 3.1.4 Other types of compounds

Peak 2 demonstrated the presence of the deprotonated ion at  $m/z$  341.1061, and fragment ions  $m/z$  89.0248 and 179.0485, which, according to the literature, indicates the presence of sucrose (Farag *et al.*, 2014; Guedes *et al.*, 2020). The mass spectrum from peak 19 showed a precursor ion  $[M-H]^+$  at  $m/z$  498.2277, exhibiting fragment ions at  $m/z$  481.2277, 322.2073, 234.1161, and 177.0670. Thus, based on the literature, it is suggested that the metabolite is the (di-*E,E*)-*N,N'*-diferuloylspermidine (Difonzo *et al.*, 2019; Kite *et al.*, 2013; Steingass *et al.*, 2015).

Compound 21 presented a precursor ion at  $m/z$  503.2495 and fragment ions at  $m/z$  371.1990 and 209.1148. According to the literature, the precursor ion together with fragment ions are compatible with the leaaside metabolite. Because of this, the fragment ions  $m/z$  371  $[(M-H)-132]^-$  and 209  $[(M-H)-132-162]^-$  are due to the loss of the pentoside group and the successive losses of the pentoside and glucoside groups (Silva, M. de F. G. da *et al.*, 2018; Zhang, X. *et al.*, 2015).

The mass spectrum referring to peak 22 reveals the precursor ion at  $m/z$  373.1145  $[M+H]^+$ , thus indicating that it is the hydro feruloyl glucose. Corroborating the proper identification, we observed the fragment ions at  $m/z$  211.1686, 193.1587, 175.1465, and 135.1159. The ion  $m/z$  211  $[(M+H)-162]^+$  comes from the loss of a dehydrated hexose unit, while the ion  $m/z$  193  $[(M+H)-180]^+$  corresponds to the loss of a hexose. In addition, ions in  $m/z$  193, 175, and 135 suggest the presence of a hydroxyferulic acid residue (Ma *et al.*, 2007).

### 3.2 Mineral contents in the pineapple leaves

The data referring to the quantification of the minerals are summarized in Table 4, where the average quantities may be observed together with the standard deviation of three replicates.

Tabela 3 – Table 3. Quantification of minerals in samples of commercial pineapple leaves.

Pineapple commercial varieties	Inorganic elements (mg kg <sup>-1</sup> )						
	Zn	Cr	Cu	Cd	Mn	P	Fe
Perolera	19.77 ± 1.50 <sup>a</sup>	< LQ	2.80 ± 0.03 <sup>c</sup>	< LQ	95.25 ± 2.65 <sup>b</sup>	3388.14 ± 69.59 <sup>b</sup>	36.83 ± 1.07 <sup>b</sup>
BRS Imperial	15.47 ± 1.43 <sup>b</sup>	< LQ	4.01 ± 0.30 <sup>b</sup>	< LQ	100.91 ± 9.02 <sup>b</sup>	2266.52 ± 56.56 <sup>c</sup>	42.25 ± 3.81 <sup>b</sup>
Smooth Cayenne	5.30 ± 0.51 <sup>d</sup>	< LQ	5.77 ± 0.26 <sup>a</sup>	< LQ	50.80 ± 3.21 <sup>c</sup>	1030.77 ± 24.55 <sup>d</sup>	70.17 ± 4.42 <sup>a</sup>
Pérola	16.86 ± 1.10 <sup>b</sup>	< LQ	1.05 ± 0.06 <sup>e</sup>	< LQ	218.34 ± 13.22 <sup>a</sup>	6163.62 ± 105.26 <sup>a</sup>	26.51 ± 1.78 <sup>c</sup>
BRS Vitória	8.47 ± 0.69 <sup>c</sup>	< LQ	1.61 ± 0.12 <sup>d</sup>	< LQ	113.98 ± 6.10 <sup>b</sup>	1172.13 ± 64.44 <sup>d</sup>	12.76 ± 0.47 <sup>d</sup>
Gold	16.22 ± 1.30 <sup>b</sup>	< LQ	1.73 ± 0.02 <sup>d</sup>	< LQ	99.98 ± 6.24 <sup>b</sup>	2435.22 ± 40.09 <sup>c</sup>	9.06 ± 0.73 <sup>d</sup>
BRS Ajubá	8.72 ± 0.92 <sup>c</sup>	< LQ	1.57 ± 0.14 <sup>d</sup>	< LQ	100.17 ± 2.79 <sup>b</sup>	1198.71 ± 85.00 <sup>d</sup>	41.78 ± 3.38 <sup>b</sup>

Source: Prepared by the Author.

Limit of quantification (LOQ); LOQ<sub>Cr</sub> = 0.02 mg kg<sup>-1</sup> and LOQ<sub>Cd</sub> = 0.02 mg kg<sup>-1</sup>.

Values are mean of three replicates ± standard deviation.

Mean values with the same letter (a, b, c, d, e) are not significantly different according to Tukey's test ( $p < 0.05$ ).



The bioavailability of minerals in plants depends on several factors, such as the nature of the soil, exposure to light, temperature, availability of water, fertilization practices, and the ability of plants to selectively accumulate some of these elements. Additional sources of elements for plants are plant protection agents, rain, atmospheric dust, and fertilizers that can be absorbed by leaf blades (Pytlakowska *et al.*, 2012; Yabor *et al.*, 2017). Thus, it is quite plausible to consider that the factors mentioned above interfere and corroborate to explain the differences in the levels of mineral concentration of the pineapple leaves of the seven commercial varieties evaluated in this study. The values range from 5.30 to 19.77 mg kg<sup>-1</sup> of zinc, 1.05 to 4.01 mg kg<sup>-1</sup> of copper, 50.80 to 113.98 mg kg<sup>-1</sup> of manganese, 1030.77 to 6163.63 mg kg<sup>-1</sup> for phosphorous, 9.06 to 70.17 mg kg<sup>-1</sup> of iron. These values are higher than those reported by ANSES (2020), evaluating juice, pulp, and fresh fruit. Fortunately, Cr and Cd, toxic metals, were not observed in any of the evaluated pineapple leaves, and the levels of these elements are below the limit of quantification of the method: LOQ<sub>Cr</sub> = 0.02 mg kg<sup>-1</sup> and LOQ<sub>Cd</sub> = 0.02 mg kg<sup>-1</sup> (Table 3).

All the mineral elements determined in the pineapple leaves exhibit important biological properties. Elements such as Mn and Fe can act by promoting the biosynthesis of amino acids and other organic compounds, considering that microelements generally act in catalysis as essential cofactors for many metabolic enzymes (Watanabe *et al.*, 2015). Manganese is a cofactor of classes of enzymes such as oxidoreductases, transferases, hydrolases, lyases, isomerases, ligases, lectins, and integrins (Grembecka; Szefer, 2013). Among the commercial pineapple varieties evaluated in this work, the Pérola variety stands out with a Mn content of 218.34 ± 13.22 mg kg<sup>-1</sup>, which is about two to four times greater than the amount observed in the other pineapple varieties studied. Regarding Fe, we can highlight the Smooth Cayenne variety (70.17 ± 4.42 mg kg<sup>-1</sup>) which has the highest Fe content when compared to the Gold (9.06 ± 0.73 mg kg<sup>-1</sup>) and BRS Vitória (12.76 ± 0.47 mg kg<sup>-1</sup>). It is important to point out that according to Turkey's test, the concentration levels of Fe in the Gold and BRS Vitória are not statistically different.

On the other hand, Zn is an essential micronutrient that modulates many enzymes involved in DNA transcription, proteins, nucleic acids, carbohydrates, and lipid metabolism. According to the literature, in plants with Zn deficiency, there is a substantial increase in starch and soluble sugars (Zhang, W. *et al.*, 2015). In addition, there are reports that state that Zn acts as an antioxidant (Akwu; Naidoo; Singh, 2019).

This mineral together with other secondary metabolites identified in this work, which have recognized antioxidant properties, may indicate that pineapple leaves probably have potential antioxidant activity. Based on the amounts of Zn in the pineapple varieties evaluated in this work (Table 3), we can infer that the Perolera variety stands out for its Zn content, which is about four times higher when compared to the Smooth Cayenne variety, which the lowest the amount.

Phosphor typically acts as a cofactor in several enzymatic systems involved in the metabolism of proteins, carbohydrates, and fats. It is an essential component of phospholipids in cell membranes and lipoproteins and is involved in the renal excretion of hydrogen ions. Phosphate depletion can activate gene expression of the flavonoid biosynthesis pathway (Deng *et al.*, 2019; Pereira; Dantas, 2016). Cu element showed low levels of concentration when compared to P and the other elements. This may be a good indication, considering that according to the literature, although Cu is an essential element for plant cellular metabolism, high amounts induce adverse effects on photosynthesis. Therefore, it is able to interfere with the production of carbohydrates and their intermediates, directly interfering with amino acid biosynthesis (Zhang; Tan; Li, 2014). In general, regarding the P and Cu elements (Table 3), it is interesting to observe the Pérola variety, which was previously highlighted for the amount of Mn. This variety also has a high amount of P when compared to other pineapple varieties. In addition, the lowest Cu content was determined in this variety, although the average value of the determined amount of Cu is similar to that of other varieties.

### 3.3 Chemometric analysis

The multivariate analysis was performed to evaluate the variability of the chemical composition of the pineapple leaves extracts of seven commercial varieties. In addition, this analysis was used to investigate a complex data matrix, more specifically, the similarities, dissimilarities, and relationships between the chemical composition and the samples. In this context, we use principal component analysis (PCA), an unsupervised method of multivariate analysis. The dataset is resized in a smaller number of variables without losing the information initially arranged, aiming to enable a clearer interpretation of the information. These calculations are based on singular value decomposition (SVD), and these results can be seen graphically through the scores and loadings.

Figure 11 shows the graphical results of the PCA. Four principal components (PC) were needed to identify the similarities and/or differences between the different types of pineapple leaves extracts that were investigated (explains 70% of the variation of the data). In Figure 2(a), it is possible to observe the scores of PC1 and PC2, with 33% and 16% of the variance explained, respectively; where we may verify the formation of the tendencies of four groups of scores: 1 – Pérola, 2 – Perolera, 3 – BRS Vitória and 4 – Smooth cayenne, BRS Imperial, BRS Ajubá and Gold. Other principal components with lower variance values were evaluated to find variables with relevant data to characterize the different types of pineapple leaf extracts not appropriately differentiated in PC1 and PC2. Thus, in Figure 2(b), the scores of the principal components PC3 and PC4 are presented (21% of the explained variance), where the differentiation of the types of pineapple leaves was observed: Gold, Smooth cayenne, and BRS Imperial.

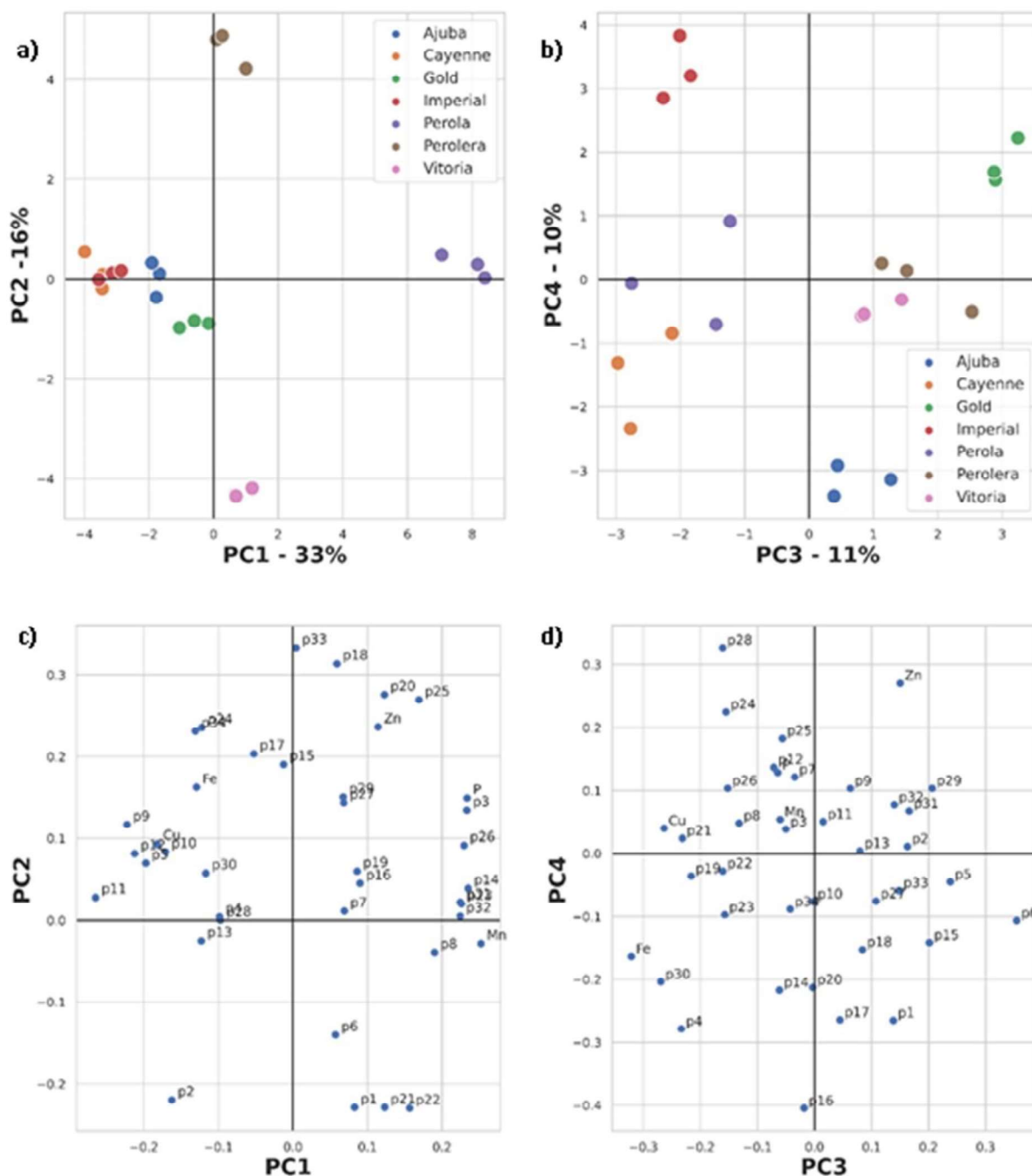
Based on the values of the loading presented in Figure 11(c) and (d), the main secondary metabolites and minerals that contributed to the differentiation of the seven commercial varieties of pineapple leaves were: Pérola (Mn, apigenin-6,8-C-diglucoside, P, threonic acid, L-tyrosine, ananaflavoside B, and ananaflavoside C), Perolera (isovitexin), BRS Vitória (sucrose, hydro feruloyl-glucose, leeaosid and malic acid), Smooth Cayenne (quinic acid and Fe), Gold (tricin), BRS Imperial (apigenin and santin), and BRS Ajubá (vitexin and rutin).

One of the metabolites identified as markers of the Pérola variety is threonic acid. According to the literature, this compound is characterized as the main product of the degradation of ascorbic acid, which it is used as a food additive (Thomas; Hughes, 1983). The compound apigenin-6,8-C-diglucoside was also identified as a metabolite associated with the Pérola variety. This flavonoid is a flavone C-glycosides that is present in foodstuffs and nutraceuticals and has been reported that it performs a wide range of biological activities (Choi *et al.*, 2014).

Other secondary metabolites were also determined as markers of the Pérola variety, among which we can highlight ananaflavoside B and ananaflavoside C. These metabolites are characteristic of the genus *Ananas* and have antioxidant, hypoglycemic, anti-inflammatory, and antiparasitic properties (Ezuruike; Prieto, 2014; Ononamadu *et al.*, 2019). Regarding the mineral Mn, according to the literature, it acts as an antioxidant. Thus, in general, the presence of this metal in pineapple leaf extracts in the seven commercial varieties – being more pronounced in the Pérola variety –

indicates that the pineapple leaves may have antioxidant activity (Akwu; Naidoo; Singh, 2019).

Figure 11 – Figure 11. Chemometrics analysis: (a) scores coordinate system PC1 × PC2; (b) scores coordinate system PC3 × PC4; (c) loadings coordinate system PC1 × PC2; (d) loadings coordinate system PC3 × PC4. In the loading graphs, the signals p1 to p34 represent the peaks referring to the secondary metabolites annotated and described in Table 2



Source: Prepared by the Author.

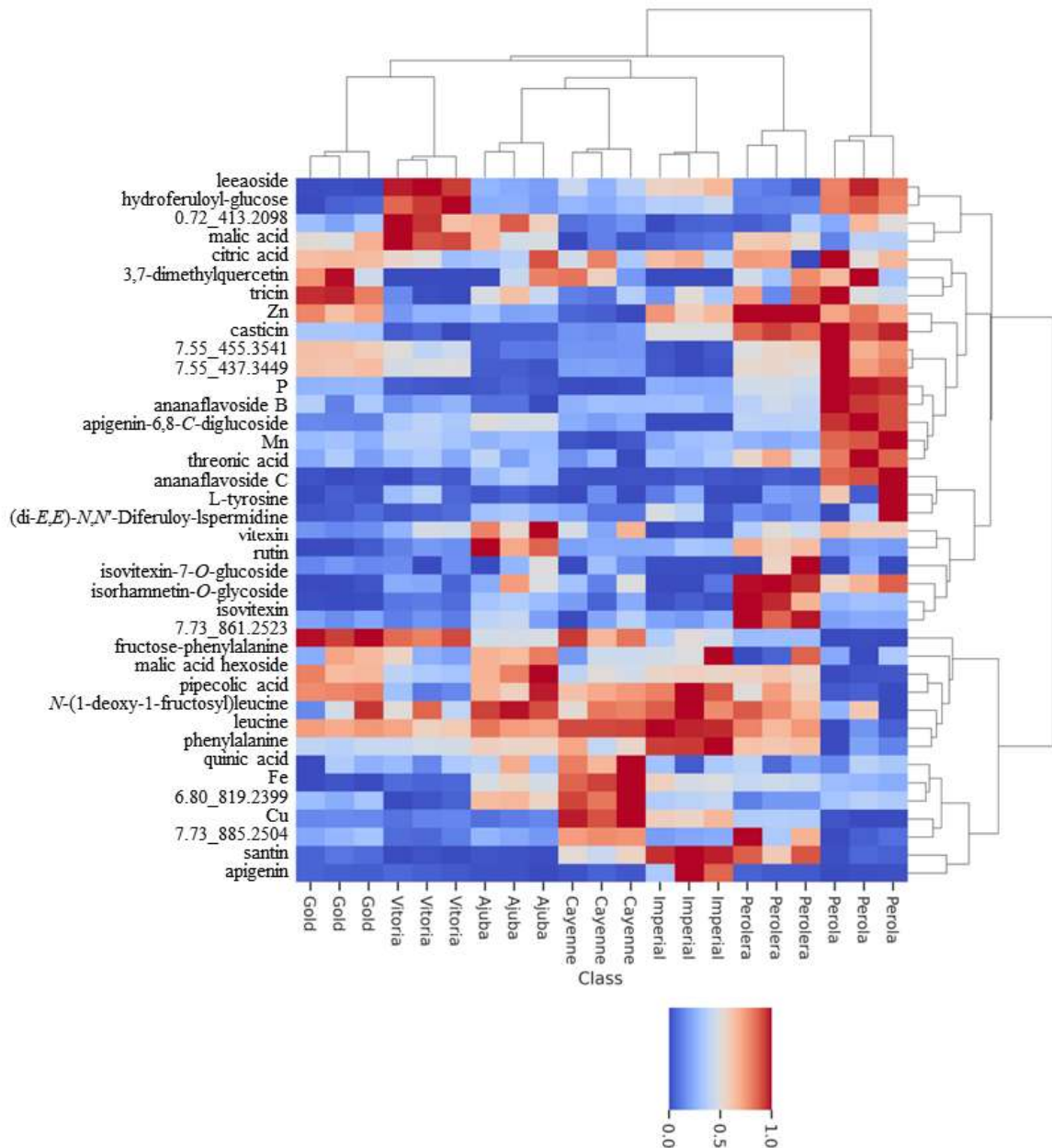
The Perolera, BRS Ajubá, and BRS Imperial varieties stand out respectively

for the presence of the metabolites isovitexin, vitexin, and apigenin. Studies have revealed that these three compounds have a wide range of bioactive properties, including anti-diabetic, antioxidant, anti-Alzheimer, anti-inflammatory, hepatoprotective, neuroprotective, and antimicrobial (Choi *et al.*, 2014). The apigenin metabolite can also be involved in other activities, such as in modulating the mRNA expression of the main enzymes responsible for obesity, helping in the Anti-obesity effects (Hassan *et al.*, 2021). It is important to report that the three metabolites mentioned above are derived from apigenin, however, they exhibit different degrees of potential in the aforementioned biological activities. Thus, it is speculated that the amount and different positions of the C-glycosides groups in apigenin affect the bioactive properties (Choi *et al.*, 2014). In addition to the apigenin metabolite, the flavonoid santin was also identified as a marker of the BRS Imperial variety. Studies reported that this flavonoid showed anti-influenza (Zhong *et al.*, 2019) and anti-parasitic activity (Mai *et al.*, 2015).

As mentioned earlier, through chemometric tools, it was found that the BRS Ajubá variety has the metabolites vitexin and rutin as markers. Rutin, also known as vitamin P, is a biflavonoid commonly found in fruit and leafy plant materials. This flavonoid has recognized antioxidant activity and may contribute to increasing UV filter activity. There is also a preliminary study on the development of a colloidal lipid system that can increase the antioxidant and photoprotective activities of rutin so that its use can be further explored by the cosmetic industry (Baldisserotto *et al.*, 2018; Kamel; Mostafa, 2015).

The extracts of pineapple leaves of the Gold and BRS Vitória varieties stand out, respectively, for the flavone triclin and the organic acid hydroferuloyl glucose (Boubakri *et al.*, 2017); both have antioxidant properties. In the case of the triclin metabolite, as well as its derivatives, their potential for application in pharmaceutical products has also been reported, due to their low toxicity and reasonable bioavailability as an anti-inflammatory and antioxidant (Li *et al.*, 2016). On the other hand, the Cayenne variety differs from the other varieties due to the presence of quinic acid and Fe. Properties such as antioxidant and antibacterial activities are attributed to quinic acid (Ramos *et al.*, 2022), while the mineral Fe is essential and necessary for oxidative metabolism and oxygen transport in the circulatory system (Akwu; Naidoo; Singh, 2019).

Figura 12 – Figure 12. Result of the clusters presented as a heat map, showing the variability of metabolites and minerals.



Source: Prepared by the Author.

A hierarchical cluster analysis (HCA) was also performed. In this analysis method, each sample starts as a separate cluster and the algorithm continues to combine them until all samples belong to a single cluster. Therefore, the HCA was graphically represented using a heat map, which was hierarchically grouped based on Ward's methods. The heat map was applied to visualize the content and distribution of differential metabolites and minerals in the seven commercial varieties of pineapple leaves (Figure 12). Each column represents a sample analyzed in this graphical

representation, and each line is a metabolite or a mineral. Concerning the colors, the red boxes indicate that the analytes are at higher concentration levels, while the blue boxes suggest lower concentration levels. Moreover, the intensity of the colors illustrates the degree of the presence or absence of the analytes in the samples. Thereby, we can observe, in general, each analyte in the seven pineapple varieties.

Evaluating the dendrogram for the commercial varieties of pineapple leaves, it is possible to verify that all of them were correctly grouped in their corresponding branches. Furthermore, by making a correlation with the PCA, it is possible to observe that the variables with a relevant contribution in the formation of the score clusters were also identified as variables that characterize the branches obtained.

### 3.4 Investigation of cytotoxic potential

The cytotoxicity of pineapple leaves was analyzed according to the percentage of cell growth inhibition, classified as low (0 – 50%), moderate (51 – 74%), and high (75 – 100%) (Guedes *et al.*, 2018; Technical Committee, 2009). Thus, the extracts exhibited a low percentage of growth inhibition for the tumor and non-tumor cell lines evaluated, ranging from 0.26 to 19.98% for HCT-116, 0.52 to 5.65% for HL60, 4.12 to 34.19% for PC3, 0.89 to 4.09% for SNB10, -4.35 to 20.70% for MCF-7, 0.41 to 2.51% for HeLa and 0.06 to 16.32% for non-tumor cell line L929 (Table 4).

In the literature, there are also reports of cell viability assays being used to investigate the cytotoxic potential of bromelain, a compound isolated from the ethanolic extract of pineapple stem and peel (Lee *et al.*, 2019). This report shows more significant cytotoxicity in tumor cells than in non-tumor keratinocyte cells. In the present work, in which the leaves of different pineapple varieties were evaluated, similar behavior can be observed when evaluating the non-tumor cell, in general, line L929 (Table 5).

In these circumstances, it is possible to verify that the hydroethanolic extracts evaluated in this study have low percentages of cytotoxicity. This fact can be verified when we observe the percentage classification mentioned above and when we compare it to the percentages reported by Guedes *et al.*, (2018) for hexane extracts. Furthermore, low percentages of cell inhibition were also observed in the non-tumor L929 strain ( $0.06 \pm 0.74\%$  to  $16.32 \pm 0.64\%$ ), which corresponds to healthy cells, corroborating to indicate a non-toxicity of the hydroethanolic extracts of the seven

commercial varieties evaluated.

An important factor that needs attention is that, among the annotated metabolites, there are several that are responsible for recognized pharmacological properties. In this sense, we can highlight apigenin-6,8-*C*-diglycoside derived from apigenin, a flavonoid known for its antiproliferative potential (Caxito *et al.*, 2015). When isolated, this compound exhibits cytotoxic activity in the concentration range of 50 – 100  $\mu\text{mol L}^{-1}$  against CCRF-CEM and CCRFADR5000 cells, human acute lymphoblastic leukemia cells (Ozarowski *et al.*, 2018). In addition to the metabolite highlighted above, other apigenin derivatives which may have similar effects were annotated (Table 2), such as isovitexin-7-*O*-glucoside or isovitexin-6''-*O*-glucoside, vitexin, isovitexin, and the aglycone apigenin.

#### 4 Conclusions

The chemical profile (metabolites and minerals) of the seven commercial varieties of pineapple leaves (BRS Ajubá, BRS Imperial, Perolera, Gold, Pérola, Smooth Cayenne, BRS Vitória, and BRS Ajubá) was evaluated for the first time. Regarding the metabolic profile, thirty-four compounds were detected by UPLC-QTOF-MS<sup>E</sup>, of which twenty-eight were duly annotated, including amino acids, organic acids, and phenolic compounds.

On the other hand, the evaluation of the mineral constitution of pineapple leaves was performed by ICP-OES, in which the minerals Zn, Cr, Cu, Cd, Mn, P, and Fe were determined and quantified. It is important to mention that the toxic metals Cr and Cd present concentration levels below the quantification limit of the method ( $\text{LOQ}_{\text{Cr}} = 0.02 \text{ mg kg}^{-1}$  and  $\text{LOQ}_{\text{Cd}} = 0.02 \text{ mg kg}^{-1}$ ) for all evaluated pineapple varieties. Additionally, corroborating with the evaluation of the cytotoxicity, the analysis of the cytotoxic potential revealed the non-cytotoxicity of the hydroethanolic extracts, indicating that, in general, the agroindustrial residues of pineapple leaves are non-toxic.

Mainly through the chemometric analysis performed through PCA and HCA, it was possible to perceive that the evaluated pineapple leaves have similarities and differences in the chemical composition of the extracts, generating information about the characterization of pineapple cultural residues.

Thus, the results obtained contribute to adding value to pineapple



agricultural residues, which are routinely discarded by the agroindustry. In addition, the results suggest that the use of this agricultural residue is a viable source of great potential for bioactive chemical species, secondary metabolites, and minerals.

### **Acknowledgments**

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001. The authors gratefully acknowledge the financial support from the CNPq, National Council for Scientific and Technological Development (303791/2016-0), INCT BioNat, National Institute of Science and Technology (grant # 465637/2014-0). We would also like to thank Embrapa (SEG 03.14.01.012.00.00).

Tabela 4 – Table 4. Cytotoxic activity of the hydroethanolic extract of pineapple leaves of different commercial varieties determined by MTT assay after 72 h of incubation at a concentration of 100 µg mL<sup>-1</sup> (% inhibition ± SD\*).

Extract	Cell proliferation inhibition (%) <sup>a</sup>						
	HCT-116	HL60	PC3	SNB19	MCF-7	HeLa	L929
Cayenne	19.98 ± 4.08	4.08 ± 3.57	12.46 ± 0.82	3.57 ± 1.04	11.24 ± 0.78	0.82 ± 1.51	13.69 ± 2.43
Perolera	13.47 ± 2.26	2.26 ± 2.83	34.19 ± 1.17	2.83 ± 1.54	20.70 ± 1.13	1.17 ± 1.97	16.32 ± 0.64
BRS Imperial	7.21 ± 0.52	0.52 ± 4.09	9.63 ± 0.47	4.09 ± 1.48	17.78 ± 0.62	0.47 ± 0.50	2.85 ± 0.87
Pérola	13.99 ± 1.04	1.04 ± 3.57	12.46 ± 2.51	3.57 ± 3.59	17.02 ± 1.37	2.51 ± 1.19	5.08 ± 0.07
BRS Vitória	19.46 ± 5.65	5.65 ± 2.94	7.43 ± 2.51	2.94 ± 3.05	17.26 ± 4.57	2.51 ± 1.91	1.10 ± 1.97
Gold	0.26 ± 1.04	1.04 ± 0.89	4.64 ± 0.41	0.89 ± 6.69	-4.35 ± 0.59	0.41 ± 1.73	0.06 ± 0.74
BRS Ajubá	4.26 ± 1.56	1.56 ± 3.62	4.12 ± 0.88	3.62 ± 1.32	0.67 ± 0.67	0.88 ± 0.26	9.38 ± 0.77
Dox <sup>a</sup> . IC <sub>50</sub> [µM]	0.21 (0.16 – 0.29)	0.02 (0.01 – 0.02)	0.76 (0.59 – 0.93)	2.07 (1.78 – 2.41)	0.15 (0.12 – 0.19)	–	1.72 (1.58 – 1.87)

Source: Prepared by the Author.

\* Results are expressed as mean percent cell growth inhibition (GI%) and standard deviation (SD) for two independent experiments in triplicate.

<sup>a</sup> Doxorubicin was the positive control. IC<sub>50</sub> is the drug concentration that caused 50% inhibition of cell growth, with the corresponding 95% confidence interval (CI 95%).

**5 ABORDAGENS METABOLÔMICAS PARA EXPLORAR A QUIMIODIVERSIDADE EM SEMENTES DE GUARANÁ (*Paullinia cupana*) UTILIZANDO UPLC-QTOF-MS<sup>E</sup> E RMN**

---

**Metabolomic approaches to explore chemodiversity in seeds of guaraná (*Paullinia cupana*) using UPLC-QTOF-MS<sup>E</sup> and NMR analysis**

Tamyris de Aquino Gondim, Jhonyson Arruda Carvalho Guedes, Elenilson de Godoy Alves Filho, Gisele Silvestre da Silva, Natasha Veruska dos Santos Nina, Firmino José do Nascimento Filho, André Luiz Atroch, Gilvan Ferreira Da Silva, Gisele Simone Lopes and Guilherme Julião Zocolo

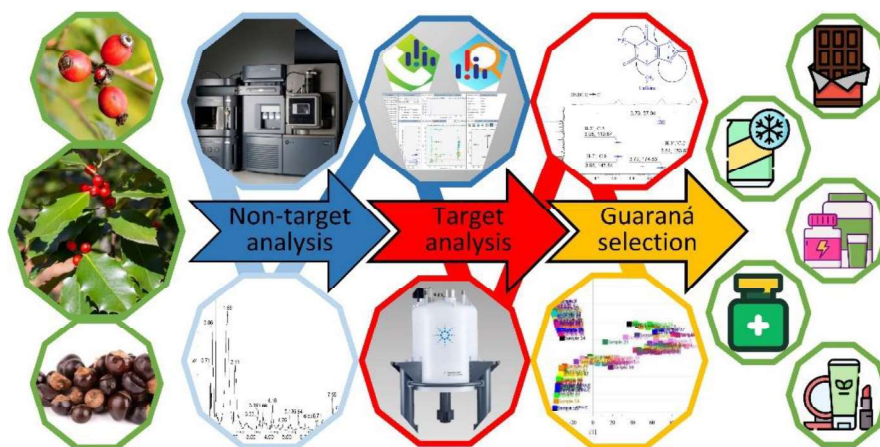
Artigo publicado em: *Analytical Methods*, Vol. 8, 2024

doi: 10.1039/D3AY01737K

Acesso via QR code:



## Graphical abstract



## ABSTRACT

The growing interest in health and well-being has spurred the evolution of functional foods, which provide enhanced health benefits beyond basic nutrition. Guaraná seeds (*Paullinia cupana*) have been widely studied and used as a functional food due to their richness in caffeine, phenolic compounds, amino acids, and other nutrients. This has established guaraná as a significant food supplement, with Brazil being the largest producer of the world. This study aims to propose a set of analytical methods to chemically evaluate fifty-six different guaraná clones, from the Guaraná Germplasm Active Bank, to accommodate the diverse requirements of the food industry. Metabolomic approaches were employed, in which a non-target metabolomic analysis via UPLC-QTOF-MS<sup>E</sup> led to the annotation of nineteen specialized metabolites. Furthermore, targeted metabolomics was also used, leading to the identification and quantification of metabolites by NMR. The extensive data generated were subjected to multivariate analysis, elucidating the similarities and differences between the evaluated guaraná seeds, particularly concerning the varying concentration levels of the metabolites. The metabolomics approach based on the combination of UPLC-QTOF-MS<sup>E</sup>, NMR and chemometric tools provided sensitivity, precision and accuracy to establish the chemical profiles of guaraná seeds. In conclusion, evaluating and determining the metabolic specificities of different guarana clones allow for their application in the development of products

with different levels of specific metabolites, such as caffeine. This caters to various purposes within the food industry. Moreover, the recognized pharmacological properties of the annotated specialized metabolites affirm the use of guarana clones as an excellent nutritional source.

## 1 Introduction

The growing concern with health and nutrition has driven the development and commercialization of functional foods. There was a greater understanding of how foods and drinks can be used to reduce the risk of diseases and consequently improve overall health. As a result, the development and consumption of functional foods have become increasingly important as they offer more than just standard nutrition (Contini *et al.*, 2023; Corbo *et al.*, 2014). In this context, for centuries, guaraná seeds (*Paullinia cupana* Kunth, family Sapindaceae) have been used by the indigenous tribes of the Amazon Forest as a stimulant and for a variety of medicinal purposes. Thus, this plant has been cultivated on a large scale by the beverage industry as a natural stimulant, with Brazil being the largest guaraná producer in the world. Currently, guaraná seeds not only represent a global trend in the soft drinks and energy drinks market but are also promising materials for the development of herbal medicines and dietary supplements (Santana; Macedo, 2018; Silva, F. de A. *et al.*, 2018).

The guaraná fruit consists of a seed partially covered by a white substance called aril, surrounded by red skin. Only the seeds are consumed and sold in powder or whole grains form. In general, guaraná seeds have a varied composition that includes 2-6% caffeine, 60% starch, 15% protein, 0.16% lipids, and 14% phenolic constituents. Among the phenolics, 13% tannins and 5.72% condensed tannins were detected (Santana; Macedo, 2018). Guaraná seeds are rich in methylxanthines, such as caffeine, theophylline, and theobromine, which have positive effects on the central nervous system, cardiovascular, gastrointestinal, respiratory, and renal

systems (Dorneles *et al.*, 2018; Santana; Macedo, 2019). In addition, they have catechins, epicatechins, and epicatechin gallate, which are antioxidant (Santana; Macedo, 2018), antimicrobial (Santana; Macedo, 2018), antiproliferative (Carla Cadoná *et al.*, 2016), antitumor (del Giglio *et al.*, 2013), and cytoprotective (Bonadiman *et al.*, 2017). These characteristics make guaraná an important commodity and a functional food ingredient (Marques *et al.*, 2016; Santana; Macedo, 2018, 2019).

The determination of the metabolic profile is an important strategy to establish an overview of a biological system, as well as a comprehensive view of the biochemical state of these organisms at a given time. Thus, an important tool to establish the metabolic profile of a given organism is the metabolomics approach. Metabolomic studies are performed through the association of analytical tools for separation and detection, known as hyphenated or coupled approaches (González-Riano *et al.*, 2020; Pilon *et al.*, 2020; Pontes *et al.*, 2017). Among the most popular combinations are liquid and gas chromatography coupled with ultraviolet detectors or mass spectrometers such as HPLC-UV-DAD, LC-MS, or GC-MS. In addition, nuclear magnetic resonance (NMR) has also been applied both for the structural elucidation of previously isolated and purified molecules and the study of enriched fractions or crude extracts of high complexity (Ouyang *et al.*, 2014; Pilon *et al.*, 2020). Nuclear Magnetic Resonance (NMR) requires little manipulation of the samples and does not require chromatography, allowing easy quantification. In addition, it offers several ways to identify metabolites, although it is usually limited to detecting the most abundant metabolites. ( $\geq 1$  mM) (Marshall; Powers, 2017). The great value of NMR spectroscopy resides in its capacity to provide information of an unambiguous identification and an absolute quantification of the detected metabolites without the need to resort to calibration curves for each analyte. This capability is due to the direct proportionality between the NMR spectrum signals and the actual molar levels of the metabolites in question (Salem *et al.*, 2020). On the other hand, LC-MS is considered the most versatile technique to cover the metabolome and the

most appropriate for the rapid dereplication of natural products in complex mixtures. Its high sensitivity together with the multiplicity of stationary phases with different chemistries, allied to the diverse sources of ionization that cover a wide variety of compounds, allows it to be optimized for almost all classes of natural products (González-Riano *et al.*, 2020; Ibáñez *et al.*, 2015; Salem *et al.*, 2020; Zhang *et al.*, 2019).

The aforementioned analytical techniques are of paramount importance in the design of targeted and non-targeted metabolomics studies. Targeted analysis is used in order to identify and quantify a specific metabolite or class of metabolites. For this, selective extractions and/or separations can be used to concentrate the desired metabolites and avoid possible interference from other compounds. This approach allows the obtainment of a sensitive and reliable identification, as well as a precise and accurate result for the selected compounds. However, this technique does not provide a complete view of the chemical composition of the sample and may leave out other compounds not considered at the beginning of the analysis (Daglia *et al.*, 2014; González-Riano *et al.*, 2020; Pilon *et al.*, 2020). On the other hand, non-target analysis comprises the use of various approaches, such as metabolic fingerprinting and metabolite profiling. Metabolic fingerprinting is employed when it is necessary to distinguish between samples without identifying particular metabolites, while metabolite profiling requires recognition and sometimes quantification of metabolites from different classes of compounds (Daglia *et al.*, 2014; González-Riano *et al.*, 2020; Pilon *et al.*, 2020).

Due to the reported above, the present work aims to select guaraná clones with the desired characteristics to meet the needs of the different types industry, such as food industry, cosmetic industry and pharmaceutical industry. In this context, once the chemical profiles of the clones have been established, the different clones can be useful for different purposes. For example, a potential source of bioactive compounds to be explored by the pharmaceutical and cosmetic industries. On the other hand, in the case of the

food industry, clones with higher or lower levels of certain metabolites can be selected, aiming to develop products to serve people of different age groups, such as children and adults. In order to achieve this, a set of different analytical methods was used, which involve sample preparation, NMR, LC-MS/MS and chemometrics tools.

In general, it is important to highlight that the combination of the set of analytical methods used, through target and non-target metabolomics approaches, provided the execution of a comprehensive chemical analysis (qualitative and quantitative) of specialized metabolites from guaraná seeds. Consequently, the methodological strategy used in this work offers advantages over other studies that use unique analytical methods. Considering that the combination of analytical methods used allows the sum of desirable and essential characteristics to obtain results with quality and analytical reliability. In this way, this combination brings together the compound separation power of chromatography, the identification capacity of mass spectrometry and nuclear magnetic resonance and tools that support the most accurate interpretation of results (MS-DIAL (Tsugawa *et al.*, 2015), MS-FINDER (Lai *et al.*, 2018; Tsugawa *et al.*, 2016), LOTUS Natural Products Online (Rutz *et al.*, 2022), NPClassifier (Kim *et al.*, 2021), ClassyFire (Djombou Feunang *et al.*, 2016) and multivariate analysis).

Thus, metabolomic fingerprinting of guaraná seed extracts was obtained using NMR and LC-MS/MS techniques, generating valuable and complex information on the clones. As to evaluate the large volume of data obtained, multivariate analyzes were applied, such as principal component analysis (PCA). These analyses made it possible to assess and identify the similarities and differences between the studied guaraná clones.

## **2 Experimental**

### **2.1 Plant material**



The guaraná (*Paullinia cupana*) samples evaluated in this study were collected from fifty-six different plants located at Embrapa Western Amazon Guaraná Germplasm Active Bank, Manaus, AM, Brazil.

Ripe fruits were collected, and the seeds were separated from the aril. Subsequently, the samples were cleaned in running water, selecting seeds ranging from dark brown color to black. Selected seeds were placed in paper bags, duly identified, and dried in an oven with forced air circulation at 45 °C. Then, the samples were crushed until the formation of a fine powder and stored.

The project has activities to access the genetic heritage of guaraná, conserved in the Embrapa Western Amazon Guaraná Germplasm Active Bank, as well as clones and progenies from genetic breeding in evaluation and selection trials. This project has received authorization from the Genetic Heritage Management Council under authorization number A1BCD7A (access registration on the SISGEN platform).

## **2.2 Reagents and chemicals**

The materials used in the analysis and the preparation of solutions necessary for the development of the work were: ultrapure water obtained by Milli-Q system (Millipore, Bedford, MA, USA); acetonitrile (LC-MS grade) supplied by Tedia (Fairfield, Ohio, EUA); hexane (95%) and ethanol (96%) purchased from Tedia (Rio de Janeiro, RJ, Brazil); and formic acid (purity 98%), deuterated methanol (99.9%), deuterated water (99.9%), and sodium-3-trimethylsilyl propionate (TMSP-d<sub>4</sub> 98%) purchased from Cambridge Isotope Laboratories (Tewksbury, MA, USA), as analytical standard: Catechin, Procyanidin B2, Caffeine, Epicatechin, Procyanidin B1 (Sigma-Aldrich Canada Ltd., Oakville, Canada), and EDTA from Vetec Quimica Fina Ltd. (Duque de Caxias, RJ, Brazil).

## **2.3 Sample preparation for UPLC analysis**

The guaraná seed samples of the 56 evaluated clones were submitted to a microextraction procedure adapted from the literature (Chagas-Paula *et al.*, 2015; Guedes *et al.*, 2020; Nehme *et al.*, 2008). Thus, 50 mg of dry, ground, and homogenized plant material samples were weighed, 4 mL of hexane were added and vortexed for 1 min. Samples were placed in an ultrasonic bath (fixed power of 135 W) for 20 min. Following this, 4 mL of a ethanol:water (7:3) solution was added, vortexed again for 1 min, and placed in the ultrasound bath for 20 min. To complete the separation of the hexane and hydroethanolic phases, the test tube containing the mixture was centrifuged for 10 min. Afterward, 2 mL aliquot was removed from the hydroethanolic phase and filtered (PTFE filter, 0.22  $\mu\text{m}$ ) before being added to vials. Seed samples of 56 different guaraná clones were extracted in biological triplicate, extraction temperature 25 to 27 °C, where each extract was analyzed only once by UPLC-QTOF-MS<sup>E</sup>. In addition, 10 blanks were extracted, totaling 178 extractions. From the sample extracts, 24 analytical quality controls (QC) were prepared, an aliquot of 10  $\mu\text{L}$  of each extraction was removed and added to a vial (24 QC).

### 2.3.1 UPLC-HRMS analysis

The chromatographic separation was performed on an Acquity UPLC (Waters Corp., Milford, MA, USA) and coupled to a quadrupole/time of flight (QTOF) mass spectrometer. Chromatographic runs were conducted through a Waters Acquity UPLC BEH 100 mm x 2.1 mm, 1.7  $\mu\text{m}$  column at a constant temperature of 40 °C, with a flow rate of 0.4 mL min<sup>-1</sup>. The analysis was carried out by applying the following binary gradient (A (water containing 0.1% formic acid) and B (acetonitrile containing 0.1% formic acid)) at a flow rate of 0.4 mL min<sup>-1</sup>: 0.0 – 10.0 min, 5 to 35% B; 10.1, 85% B; 11.0 min, 80% B; 11.1 – 16.0 min, 5% B gradient. The sample injection volume was 2  $\mu\text{L}$ .

The chromatographic system employed was a UPLC system coupled to a QTOF-MS<sup>E</sup> mass spectrometer and operated in ESI<sup>+</sup> and ESI<sup>-</sup> ionization modes. The desolvation gas flow was set at 350 L h<sup>-1</sup> (ESI<sup>+</sup>) and 500 L h<sup>-1</sup> (ESI<sup>-</sup>), while the temperature of the ionization source and the desolvation gas were set at 120 °C and 350 °C, respectively. Leucine enkephalin was used as a lock mass, and the capillary voltage was set at 3 kV. The acquisition range was set at 50 – 1500 *m/z* in MS mode and 50 – 1500 *m/z* in MS/MS mode.

The data obtained from the UPLC-ESI-QTOF-MS<sup>E</sup> was analyzed using the software MS-DIAL 4.9.221218 (Data Independent Analysis) to set up the parameters for untargeted metabolomics, including deconvoluted spectra, peak alignment, and filtering (Lai *et al.*, 2018; Tsugawa *et al.*, 2015, 2019). After this, the unidentified metabolites were annotated using the MS-FINDER 3.60 (Lai *et al.*, 2018; Tsugawa *et al.*, 2016, 2019). This annotation was done by comparing the MS and MS/MS mass spectra to databases like KNApSAcK Core System, Human Metabolome Database (HMDB), Kyoto Encyclopedia of Genes and Genomes (KEGG), SciFinder, ChemSpider, and PubChem. The annotation of metabolites was performed following the Metabolic Standards Initiative (MSI) level 2.1 guidelines (Sumner *et al.*, 2007). This annotation of metabolites also included the molecular formulas and the fragment ions related to the metabolites. Furthermore, the annotation of metabolites was done considering the chemotaxonomy (family, genus, and species).

## 2.4 NMR spectroscopy analysis

Approximately 30 mg of each sample of guaraná seed powder were directly solubilized in a solution containing 490 µL of deuterated methanol (99.9%), 210 µL of deuterated water (99.9%), 1.6 mg mL<sup>-1</sup> of sodium-3-trimethylsilyl propionate (TMSP-d<sub>4</sub>) as internal standard, and EDTA (5.6 mg mL<sup>-1</sup>). This solution was sonicated for 2 min and subsequently centrifuged for

10 min at 4,032 g (6,000 rpm in a 100 mm rotor, model 80-2B Centrifuge, Edulab, Curitiba-PR, Brazil), and the supernatant was transferred to a 5 mm NMR tube.

The NMR experiments were performed on an Agilent 600 MHz spectrometer equipped with a 5 mm inverse detection One Probe™ for high ( $^1\text{H}$ - $^{19}\text{F}$ ) and low ( $^{15}\text{N}$ - $^{31}\text{P}$ ) frequencies and actively shielded Z-gradient. The  $^1\text{H}$  NMR spectra were acquired in triplicate using the PRESAT pulse sequence for water suppression ( $\delta$  4.82 ppm). In order to ensure complete relaxation of all nuclei of the samples, the inversion recovery sequence was used after performing a  $90^\circ$  pulse calibration (7.9  $\mu\text{s}$  pulse length at 57 dB of power), and the probe was adequately tuned and matched. A 7 times T1 recycling delay between pulses was applied to ensure the full relaxation of all protons present in the samples. Therefore, a relaxation delay of 23.0 s was used, with an acquisition time of 3.32 s, 40 scans, 32 k of time-domain points with a spectral window of 16.0 ppm. The pre-fixed value for the receiver gain was achieved by comparing the spectra using same signal-to-noise ratio, which had been used for all acquisitions. The temperature was conserved at 298 K. Free induction decay was multiplied by an exponential function equivalent to 0.3 Hz line-broadening before applying Fourier transform for 16 k points. Phase correction was manually performed and the automatic baseline correction using polynomial degree 5 was applied over the entire spectral range.

Two-dimensional NMR experiments were acquired using the standard spectrometer library pulse sequences. The  $^1\text{H}$ - $^1\text{H}$  COSY experiments were obtained with a spectral width of 18,028.1 Hz in both dimensions; 1442  $\times$  200 data matrix; 32 scans per t1 increment, and a relaxation delay of 1.0 s. The one-bond  $^1\text{H}$ - $^{13}\text{C}$  HSQC experiments were acquired with an evolution delay of 1.7 ms for an average  $^1J(\text{C}, \text{H})$  of 145 Hz; 1442  $\times$  200 data matrix; 80 scans per t1 increment; spectral widths of 9615.4 Hz in f2 and 30,165.9 Hz in f1, and relaxation delay of 1.0 s. The  $^1\text{H}$ - $^{13}\text{C}$  HMBC experiments were recorded with an evolution delay of 50.0 ms for  $^{\text{LR}}J(\text{C}, \text{H})$  of 10 Hz; 1442  $\times$  200 data matrix; 180 scans per t1 increment; spectral widths of

9615.4 Hz in f2 and 30,165.9 Hz in f1, and relaxation delay of 1.0 s. Then, the constituents identification was performed through 2D-NMR analyses, using correlation spectroscopy (COSY), heteronuclear single quantum coherence (HSQC), heteronuclear multiple bond correlation (HMBC), assessments using an open-access database ([www.hmdb.ca](http://www.hmdb.ca)) (Wishart *et al.*, 2012), and literature reports (Alves Filho *et al.*, 2017; Cren-Olivé *et al.*, 2002; da Silva *et al.*, 2016, 2017). Molecular structures,  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts, multiplicity, correlations, and constant coupling are available in the Supporting Information.

#### **2.4.1 Quantification and analysis of variance**

The compounds in guaraná samples that presented high variations in chemometrics and did not exhibit overlapping resonances were quantified by the external reference method provided by the VnmJ™ program (version 4.2, Agilent). This technique is based on the principle of reciprocity and the NMR signals strengths are correlated with a reference sample. A stock solution composed of D<sub>2</sub>O (99.9%) and sucrose (5.0 mg mL<sup>-1</sup>) was used to calibrate the equipment, and the probe file was later updated with all the parameters required to determine the concentrations of other compounds.

Quantitative results were evaluated by analysis of variance (ANOVA single factor) using the Origin™ 9.4 software (OriginLab Corporation, USA) in order to statistically certify the differences among the concentrations at a significance level of 0.05. Tukey and Levene's tests were applied to assess the variance of homogeneity. The combined uncertainties were based on analytical errors and standard deviation from the triplicate of the spectra acquisitions.

#### **2.5 Multivariate statistical analysis**

A numerical matrix containing a total of 162  $^1\text{H}$  NMR spectra (56 different samples acquired in triplicate) was created using chemical shifts between  $\delta$  0.05 and 5.60. Each spectrum was converted to American Standard Code for Information Interchange (ASCII) files when imported by the same Origin™ software. The spectral region between  $\delta$  0.7 and 8.7 was selected for evaluation, excluding the influenced area according to the saturation profiling of the non-deuterated water signal (at  $\delta$  4.82) (da Silva *et al.*, 2016). Then, this numerical matrix presented a dimensionality of 1,326,864 data points (168 spectra  $\times$  7,898 variables into each spectrum), which was imported by the PLS Toolbox™ software (version 8.6.2, Eigenvector Research Incorporated, Manson, WA USA) for exploratory chemometric evaluation (unsupervised) by principal component analysis (PCA). Aiming at baseline correction and signals alignment using correlation optimized warping (COW) with a segment of 50 data points and a slack of 5 data points, algorithms were applied over the variables (Tomasi; van den Berg; Andersson, 2004), and posteriorly the samples were mean-centered, enhancing the differences between them (Beebe; Pell; Seasholtz, 1998). The singular value decomposition algorithm (SVD) was applied to decompose the matrix in scores and loading matrices. Relevant information was obtained at the first two principal components (PC axes), under a confidence level of 95%.

The multivariate analysis of guaraná seed samples from the data obtained by UPLC-QTOF-MS<sup>E</sup> was submitted and processed through the MarkerLynx XS software. For data processing, some parameters were defined, such as retention time interval, from 0.5 to 9.0 min; mass range, from 110 to 1500 Da; and mass tolerance, 0.02 Da. In addition, the data matrix was scaled with the Pareto method. Subsequently, the principal components analysis (PCA) was conceived, being described through graphs of scores and loadings.

## 2.6 Analysis of Variance

Compounds with non-overlapped signals were quantified by the external reference method provided by the VnmJ™ program (version 4.2, Agilent): a technique based on the principle of reciprocity, in which NMR signals strengths are correlated with a reference sample. A stock solution composed of D<sub>2</sub>O (99.9%) and sucrose (5.0 mg mL<sup>-1</sup>) was used to calibrate the equipment, then the probe file was updated with all the parameters required to determine the unknown concentrations of other compounds.

The quantitative results were evaluated by the analysis of variance (ANOVA single factor) using Origin™ 9.4 software in order to statistically certify the differences or equalities among the concentrations at the significance level of 0.05. Tukey and Levene's tests were applied to assess the variance in homogeneity. The combined uncertainties were based on analytical errors and standard deviation from the triplicate of the <sup>1</sup>H spectra acquisitions.

### **3 Results and discussion**

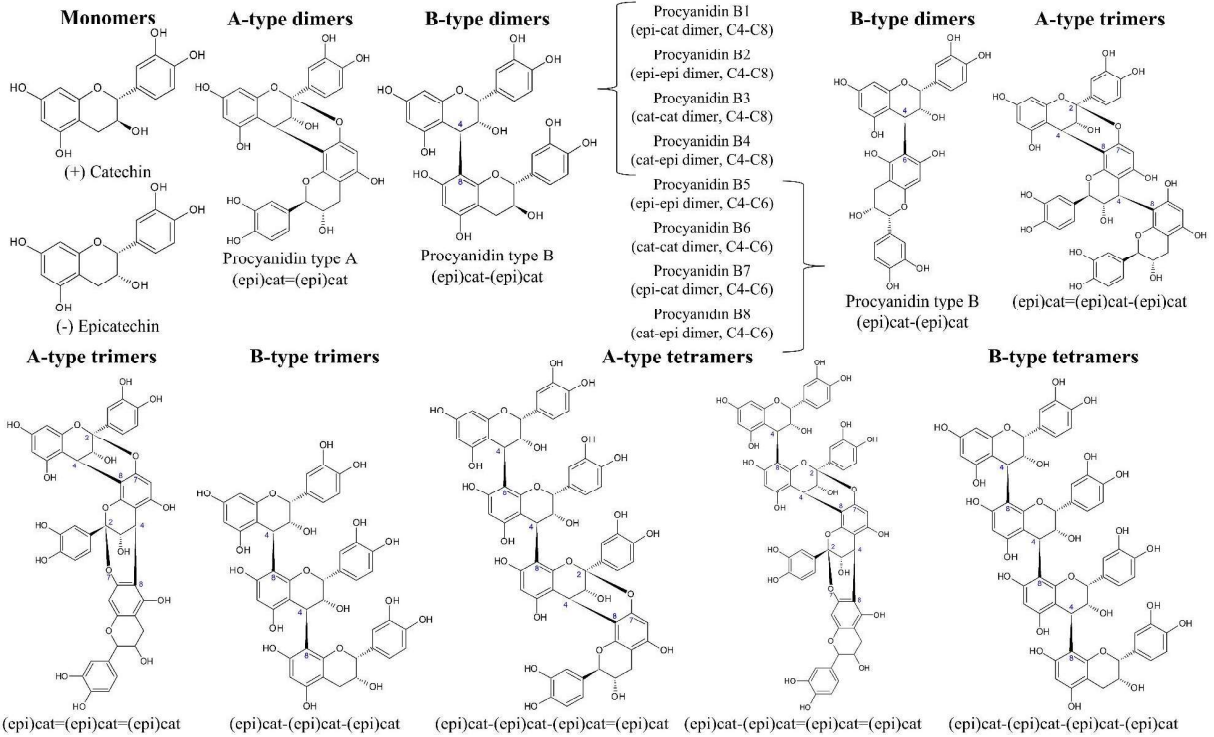
#### **3.1 Non-targeted analysis of guaraná seeds by UPLC-HRMS**

Hydroethanolic extracts from seeds of 56 guaraná clones were analyzed by UPLC-ESI-QTOF-MS<sup>E</sup>, positive (ESI<sup>+</sup>), and negative (ESI<sup>-</sup>) ionization modes. Thus, a non-targeted metabolomics approach was adopted to perform an exploratory screening of the chemical profiles of the samples. The joint evaluation of MS and MS/MS spectra allowed the observation of specialized metabolites, mainly proanthocyanidins, and methylxanthines. In general, the ESI<sup>-</sup> spectra revealed the presence of procyanidin monomers, dimers, trimers, and tetramers, which are summarized in Table 6. On the other hand, the ESI<sup>+</sup> spectra indicated the presence of methylxanthines, caffeine, and theobromine. In Fig. 13, we can observe the chemical structure of some annotated molecules. Additionally, Fig. 14 illustrates the chromatograms (ESI<sup>+</sup>

and ESI<sup>-</sup>) representative of the chemical profiles, together with the indications of the annotated specialized metabolites (Table 5), from the guaraná seeds.



Figure 13 – Fig. 13. Chemical structures of some polyphenols (procyanidins) annotated in guaraná seeds.



Source: Prepared by the Author.

Tabela 5 – Table 5. Annotation of metabolites in guaraná seeds, positive and negative ionization modes. (to be continued).

Peak	$t_R$ (min)	Negative ion mode (ESI <sup>-</sup> )				Positive ion mode (ESI <sup>+</sup> )				Molecular formula	Metabolite annotation	Chemical ontology (Kim <i>et al.</i> , 2021)	Reference
		[M-H] <sup>-</sup> observed	[M-H] <sup>-</sup> calculated	MS/MS	Error (ppm)	[M+H] <sup>+</sup> observed	[M+H] <sup>+</sup> calculated	MS/MS	Error (ppm)				
1	0.53	341.1084	341.1095	89,0281 179,0548	3.2					C <sub>12</sub> H <sub>22</sub> O <sub>11</sub>	Sucrose	Saccharides	(Farag <i>et al.</i> , 2014; Guedes <i>et al.</i> , 2020)
2	0.60	191.0192	191.0186	111,0098	-3.1					C <sub>6</sub> H <sub>8</sub> O <sub>7</sub>	Citric acid	Fatty Acids and Conjugates	(Ju <i>et al.</i> , 2021)
3	0.62	133.0137	133.036	89,0244 115,0049	-0.8					C <sub>4</sub> H <sub>6</sub> O <sub>5</sub>	Malic acid	Fatty Acids and Conjugates	(Ju <i>et al.</i> , 2021)
4	0.80					181.0726	181.0721	163,0504 138,0697 123,0474 110,0713	-2.8	C <sub>7</sub> H <sub>8</sub> N <sub>4</sub> O <sub>2</sub>	Theobromine	Pseudoalkaloids	(Novaki <i>et al.</i> , 2021)
5	1.94	577.1346	577.1359	289,0634 407,0759 425,0844	2.3	579.1503	579.1539	291,08434 09,0830	6.2	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>	B-type procyanidin dimer	Flavonoids (Proanthocyanins)	(Sang <i>et al.</i> , 2019)
6	2.02	577.1346	577.1345	289,0673 407,0540	-0.2					C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>	B-type procyanidin dimer	Flavonoids (Proanthocyanins)	(Sang <i>et al.</i> , 2019)
7	2.07	289.0712	289.0714	109,0286 125,0241 205,0476 245,0764	0.7	291,0869	291,0871	111,0791 127,0241 207,0702 247,0656 123,0450 139,0396 165,0519	0.7	C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>	Catechin	Flavonoids (Flavan-3-ols)	Analytical standard (Lv <i>et al.</i> , 2015)

Tabela 5 – Table 5. Annotation of metabolites in guaraná seeds, positive and negative ionization modes. (continuation).

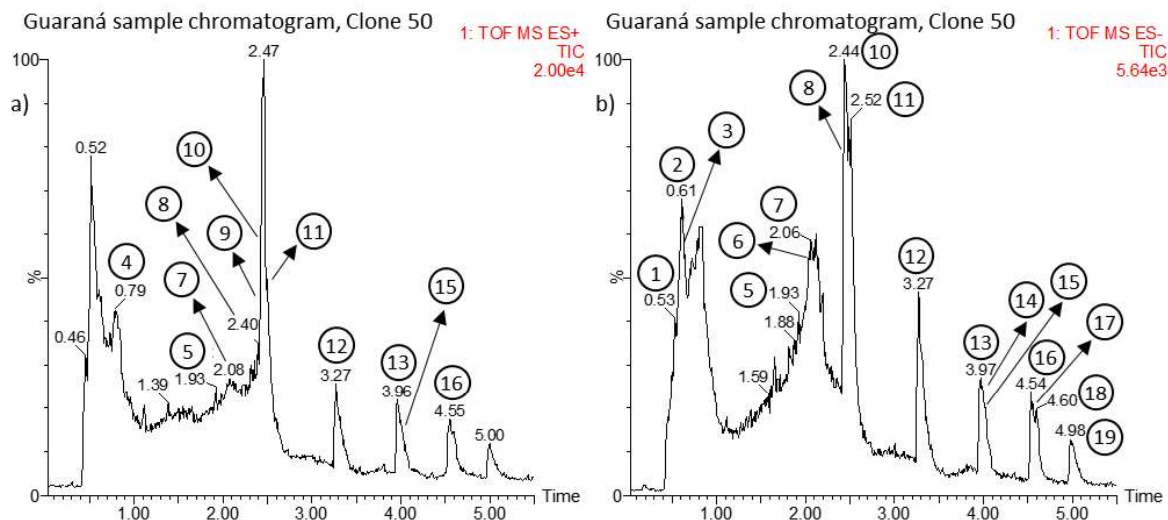
Peak	$t_R$ (min)	Negative ion mode (ESI <sup>-</sup> )				Positive ion mode (ESI <sup>+</sup> )				Molecular formula	Metabolite annotation	Chemical ontology (Kim <i>et al.</i> , 2021)	Reference
		[M-H] <sup>-</sup> observed	[M-H] <sup>-</sup> calculated	MS/MS	Error (ppm)	[M+H] <sup>+</sup> observed	[M+H] <sup>+</sup> calculated	MS/MS	Error (ppm)				
8	2.41	577.1346	577.1346	289.0697 407.0744 425.0844 451.1071	-2.9	579.1503	579.1499	291.0886 409.0918 425.1011	-0.7	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>	Procyanidin B2	Flavonoids (Proanthocyanins)	Analytical standard
9	2.43					195.0882	195.0879	138.0681 123.0456 110.0758	-1.5	C <sub>8</sub> H <sub>10</sub> N <sub>4</sub> O <sub>2</sub>	Caffeine	Pseudoalkaloids	Analytical standard
10	2.44	865.1980	865.1948	289.0636 407.0742 425.0879 577.1274	-3.7	867.2173	867.2136	291.0766 409.3626 579.1336	4.3	C <sub>45</sub> H <sub>38</sub> O <sub>18</sub>	B-type procyanidin trimer	Flavonoids (Proanthocyanins)	
11	2.51	289.0712	289.0702	245.0771 205.0440	-3.5	291.0869	291.0873	139.0384 123.0432 165.0515	1.4	C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>	Epicatechin	Flavonoids (Flavan-3-ols)	Analytical standard
12	3.28	1153.2614	1153.2646	289.0682 407.0743 577.1289 865.1855 289.0693 411.0693	2.8	1155.2770	1155.2770	291.0852 579.1578 867.2537	-6.8	C <sub>60</sub> H <sub>50</sub> O <sub>24</sub>	B-type procyanidin tetramer	Flavonoids (Proanthocyanins)	
13	3.95	863.1823	863.1770	451.0986 575.1119 711.1422 289.0690	-6.1	865.1980	865.1998	409.0936 453.1201 577.1388	2.1	C <sub>45</sub> H <sub>38</sub> O <sub>18</sub>	A-type procyanidin trimer	Flavonoids (Proanthocyanins)	
14	3.99	1151.2457	1151.2402	575.1130 863.1788	-4.8					C <sub>60</sub> H <sub>48</sub> O <sub>24</sub>	A-type procyanidin tetramer	Flavonoids (Proanthocyanins)	

Tabela 5 – Table 6. Annotation of metabolites in guaraná seeds, positive and negative ionization modes. (conclusion).

Peak	$t_R$ (min)	Negative ion mode (ESI <sup>-</sup> )				Positive ion mode (ESI <sup>+</sup> )				Molecular formula	Metabolite annotation	Chemical ontology (Kim <i>et al.</i> , 2021)	Reference
		[M-H] <sup>-</sup> observed	[M-H] <sup>-</sup> calculated	MS/MS	Error (ppm)	[M+H] <sup>+</sup> observed	[M+H] <sup>+</sup> calculated	MS/MS	Error (ppm)				
15	4.00	577.11190	577.1207	289.0708 407.0855 423.0779 449.0907 289.0695	3.0	579.1503	579.1538	291.0914 409.0931 425.0949	6.0	C <sub>30</sub> H <sub>26</sub> O <sub>12</sub>	Procyanidin B1	Flavonoids (Proanthocyanins)	Analytical standard
16	4.53	575.1190	575.1182	423.0750 449.0865 289.0700 411.0732	-1.4	577.1346	577.1383	291.0952 409.1141	6.4	C <sub>30</sub> H <sub>24</sub> O <sub>12</sub>	A-type procyanidin dimer	Flavonoids (Proanthocyanins)	
17	4.55	863.1823	863.1802	451.0945 575.1177 711.1334 289.0697	-2.4					C <sub>45</sub> H <sub>36</sub> O <sub>18</sub>	A-type procyanidin trimer	Flavonoids (Proanthocyanins)	
18	4.60	1151.2457	1151.2399	575.1180 863.1844	-5.0					C <sub>60</sub> H <sub>48</sub> O <sub>24</sub>	A-type procyanidin tetramer	Flavonoids (Proanthocyanins)	
19	4.98	575.1190	575.1210	289.0713 449.0873	3.5					C <sub>30</sub> H <sub>24</sub> O <sub>12</sub>	A-type procyanidin dimer	Flavonoids (Proanthocyanins)	

Source: Prepared by the Author.

Figure 14 – Fig. 14. Representative chromatograms of guaraná seed samples, together with the indication of the specialized metabolites annotated (Table 1): (a) positive ionization mode (ESI<sup>+</sup>); (b) negative ionization mode (ESI<sup>-</sup>).



Source: Prepared by the Author.

The identification and annotation of compounds were performed with the aid of analytical standards and an extensive review of the literature. It is important to point out that the literature review was conceived based on the research of substances found in the family, genus, and species of the plant under study (Chemotaxonomy). In addition, corroborating with the annotations of metabolites, SciFinder, LOTUS Natural Products Online (Rutz *et al.*, 2022), NPClassifier (Kim *et al.*, 2021), ClassyFire (Djombou Feunang *et al.*, 2016) and other databases available in the MS-DIAL and MS-FINDER software were consulted, as well as the *in silico* spectra available in this software. With that, a total of nineteen metabolites were noted, Table 6.

Procyanidins are oligomeric compounds formed by catechin and epicatechin monomers. These metabolites are widely found in foods and have significant medicinal properties. Since they act by exerting beneficial effects in the prevention or treatment of numerous diseases, such as cancer, diabetes, cardiovascular diseases, neurological diseases, immune imbalances, and obesity, which represent a major threat to a significant portion of the world's population (Mikyška *et al.*, 2022). Procyanidins can be

categorized into A-type and B-type, depending on the stereo configuration and the bond between the monomers. The  $m/z$  deprotonated ions 577, 865, and 1153 are intercalated by the elimination of 288 Da, which occurs by cleavage via quinone methide (QM). These ions correspond to the deprotonated molecules of the (epi)catechin dimers, trimers, and tetramers, respectively of the procyanidin series, corresponding to B-type. The A-type procyanidins found in the peaks presented ions at  $m/z$  575, 863, and 1151 and are differentiated from B-type by two units of mass. This difference is caused by the additional C-O-C bond (da Silva *et al.*, 2017; Rue; Rush; van Breemen, 2018).

Peaks 7 and 11 were annotated and identified (Table 6 and Fig. 14), respectively, as the diastereoisomers catechin and epicatechin, ESI<sup>-</sup>  $m/z$  289 and ESI<sup>+</sup>  $m/z$  291. These polyphenols are characterized as two monomers which are the building blocks of several procyanidins. The MS/MS (ESI<sup>-</sup>) spectra of these ions revealed the presence of product ions  $m/z$  109, 125, 205, and 245. The ion fragment  $m/z$  245 [M-H-44]<sup>-</sup> is due to the loss of 44 Da (CH<sub>2</sub>=CH-OH). The  $m/z$  109 fragment is a result of the loss of 180 Da. Furthermore, the  $m/z$  205 [M-H-84]<sup>-</sup> ion results from the loss of the benzene moiety, and the  $m/z$  125 [M-H-165]<sup>-</sup> ion comes from heterocyclic ring fission (HRF) (AbouZeid *et al.*, 2022; Said *et al.*, 2017).

In general, the main fragmentation pathways of procyanidins include reaction mechanisms involving cleavage by quinone methide (QM) of the inter flavonoid bond, heterocyclic ring fission (HRF), and the retro-Diels-Alder reaction (RDA). Thus, the different types of mechanisms can lead to the elimination of different amounts of masses. For example, in the case of HRF, it promotes the loss of 126 Da, while the RDA reaction leads to the elimination of 152 Da. In addition, other rearrangements may occur, as well as water loss (-18 Da) (Rue; Rush; van Breemen, 2018; Said *et al.*, 2017; Salles *et al.*, 2022). Peaks 5, 6, 8, and 11 have the same precursor ion  $m/z$  577 [M-H]<sup>-</sup>. Together with the analysis of the product ions observed in the MS/MS, this indicates the annotation of the metabolites as isomeric forms of the B-type

procyanidins. In different combinations, these metabolites are dimers formed by two monomeric units (diastereoisomers) of catechin and epicatechin (Fig. 13). These monomers are usually connected through an inter flavonoid linkage between the C4 carbon of the upper unit and the C8 carbon (C4→C8) or the C6 carbon (C4→C6) of the lower unit (Fig. 13).

Lending weight to the annotation of the B-type procyanidins dimers, the product ions  $m/z$  451, 425, 407, and 289 were observed in the MS/MS spectra. The fragmentation of the B-type procyanidin ( $m/z$  577) through HRF results in the ion  $m/z$  451 ( $[M-H-126]^-$ ). Fragment ions  $m/z$  425 and 407 result from the RDA reaction ( $[M-H-152]^-$ ) with the successive loss of a water molecule  $[M-H-152-18]^-$ . Furthermore, cleavage via quinone methide (QM) of the inter flavonoid linkage, which joins the monomeric units, produced the fragment ion at  $m/z$  289 (AbouZeid *et al.*, 2022; Qiang *et al.*, 2015; Said *et al.*, 2017; Salles *et al.*, 2022; Sui *et al.*, 2016).

The analysis of guaraná seeds also revealed the presence of oligomeric procyanidins formed by three and four monomeric units, which were annotated respectively as B-type procyanidin trimer (peaks 10,  $m/z$  865) and tetramer (peak 12,  $m/z$  1153), Table 6. The MS/MS spectra of the B-type trimers showed product ions  $m/z$  289, 407, 425, 577, 695, and 739. The fragments  $m/z$  577  $[M-H-288]^-$  and  $m/z$  289  $[M-H-288-288]^-$  are due to successive cleavages via QM. The fragment ion  $m/z$  695 ( $[M-H-152-18]^-$ ) is formed from RDA (-152 Da) with the successive loss of H<sub>2</sub>O (-18 Da). On the hand, the presence of ion  $m/z$  425 ( $[M-H-288-152]^-$ ) is the result of successive cleavages of the precursor ion ( $m/z$  865) by QM (-288 Da) and RDA (-152 Da). Furthermore, the fragment  $m/z$  425 through the loss of one water molecule forms the product ion  $m/z$  407 ( $[M-H-288-152-18]^-$ ). On the other hand, the  $m/z$  739 ion ( $[M-H-126]^-$ ) comes from an HRF (-126 Da). The procyanidin B-type tetramer ( $m/z$  1153  $[M-H]^-$ ) shows fragmentation similar to the trimer, with the addition of the product ion  $m/z$  865 which is the result of a QM cleavage (-288 Da) (da Silva *et al.*, 2017; Said *et al.*, 2017).

As shown in Table 6, dimers, trimers, and tetramers of A-type procyanidins were also annotated. Peaks 16 and 19 showed the precursor ion  $m/z$  575  $[M-H]^-$  characteristic of A-type procyanidin dimers. Fragmentation of the precursor ion led to the formation of product ions  $m/z$  449, 423, and 289. The fragment ion  $m/z$  449  $[M-H-126]^-$  is due to the HRF of the dimer by elimination of the phloroglucinol molecule (1,3,5-trihydroxybenzene). The ion  $m/z$  423  $[M-H-152]^-$  is derived from the RDA reaction. There is also the presence of the product ion  $m/z$  289, which is formed by QM fission of A-type procyanidin dimer (Salles *et al.*, 2022).

Peaks 13 and 17 suggest the presence of A-type procyanidin trimers,  $m/z$  863. Considering that the MS/MS spectra exhibited fragment ions  $m/z$  711, 575, 451, 411, and 289. In addition, we also verified the presence of A-type procyanidin tetramers (peaks 14 and 18), with precursor ion  $m/z$  1151 and with product ions  $m/z$  863, 575, and 289. The formation of these fragments ions is attributed to the mechanisms of reaction of the RDA, HRF, and QM, which represent the typical mechanisms of fragmentation of procyanidins (da Silva *et al.*, 2017; Li *et al.*, 2012; Salles *et al.*, 2022).

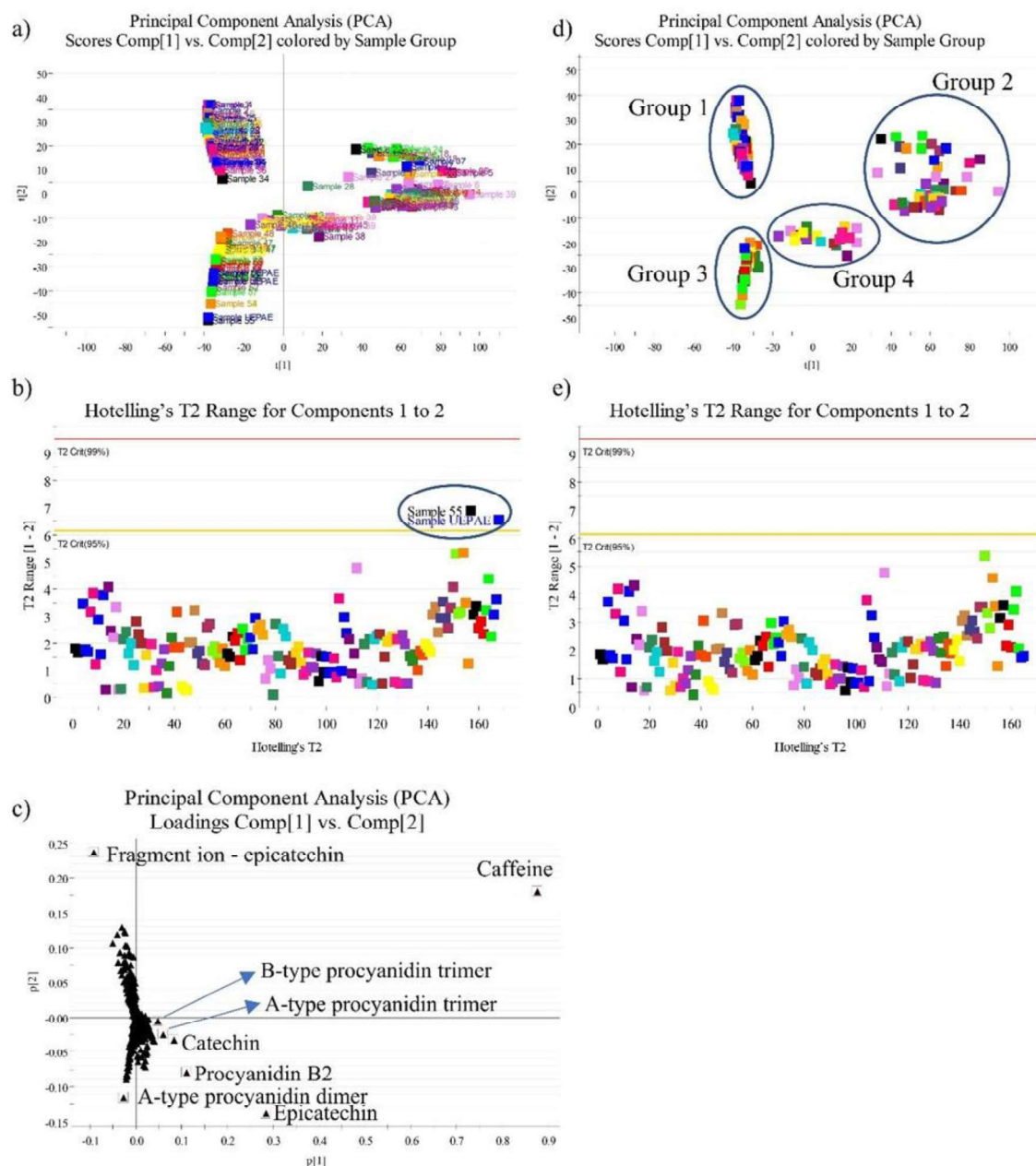
### 3.2 Multivariate analysis of data obtained by UPLC-HRMS

Principal component analysis (PCA) was used to investigate the complex data matrix with minimal loss of original information, aiming to observe trends in guaraná seed extract samples and evaluating similar characteristics, differences, and relationships between samples. In general, the numeric data matrix of guaraná seeds is 239,909, which consists of a matrix with 168 chromatograms x 1,428 variables ( $t_R$ - $m/z$  pairs). PCA was applied to the data matrix, and its graphics are shown in Fig. 15. To identify possible outliers, we used multivariate control charts of Hotelling's  $T^2$  type, as illustrated in Fig. 15(b) and (e).

Figura 15 – Fig. 15. Principal component analysis of guaraná seed samples: (a) scores with outliers; (b) Hotelling's  $T^2$  from scores with outliers; (c) loadings; (d)



scores without outliers; (e) Hotelling's  $T^2$  from scores without outliers.



Source: Prepared by the Author.

It is important to highlight that the presence of outliers is inherent to any data set with a large and complex magnitude, such as the data set evaluated in this study, which is formed by 56 subsets of guaraná seeds from different clones. In general, outliers can have a great influence on data analysis. This influence can lead to erroneous inferences about the data; in these cases, the outliers

constitute data that need to be removed. Thus, methods and criteria are needed to verify the presence of possible anomalous samples. Thereby, as mentioned previously, in this work, the evaluation of possible outliers was performed using Hotelling's  $T^2$  graphs, with 95% confidence.

Therefore, according to Hotelling's  $T^2$ , two guaraná seed clones behaved as outliers, with 95% confidence (Fig. 15 (b)). Thus, after removing the outliers from the model, the PCA was redesigned and the scores plot showed a separation of clones, as shown in Fig. 3(d). The two principal components (PC) present an accumulated explained variance of 31%. This percentage of explained variance shows great chemical similarity between the 56 samples of guaraná seeds evaluated in this study. This fact is understandable, considering that the evaluated samples are representatives of a plant organism of the same family, genus, and species.

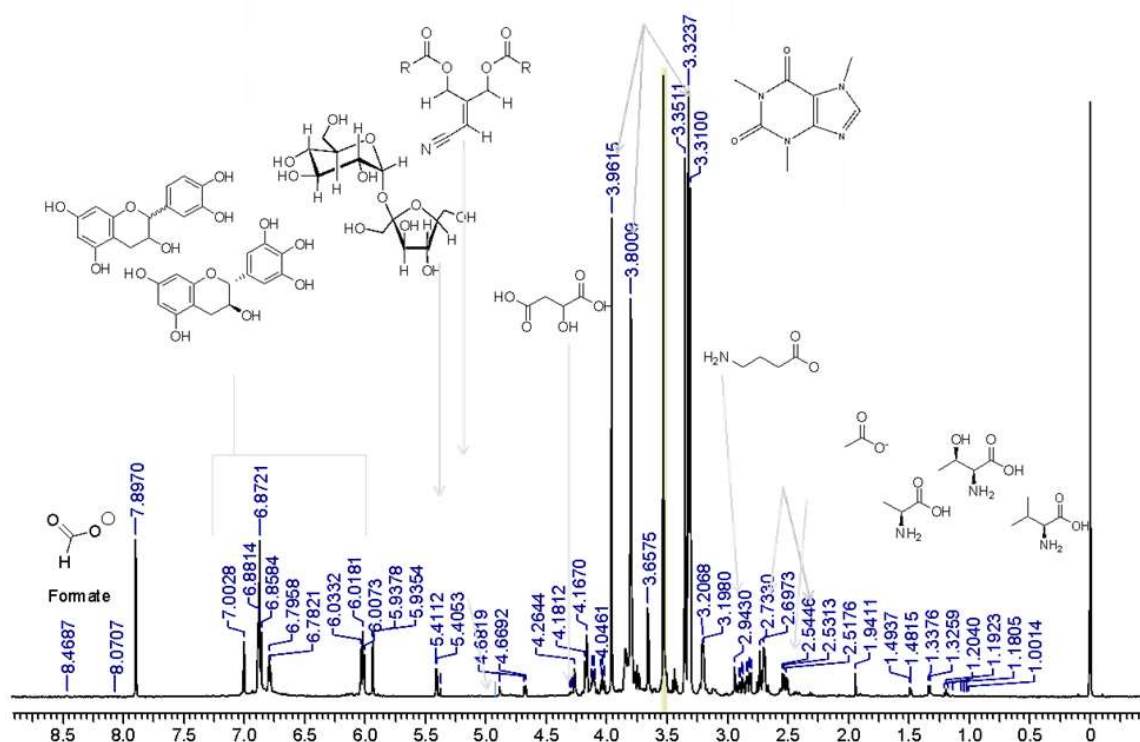
However, according to the score plot shown in Fig. 15(d), we can verify the formation tendencies of four distinct groups. The analysis of loadings (Fig. 15(c)) allowed identifying the main specialized metabolites responsible for the discrimination of the four groups. Thus, the metabolite responsible for the separation of group 1 in positive PC1 was a fragment ion of the epicatechin. The separation of Group 2 was influenced by caffeine. On the other hand, in group 3 we observed the influence of the A-type procyanidin dimer metabolite, and finally, the metabolites responsible for the separation of group 4 were attributed to catechin, procyanidin B2, and epicatechin.

### **3.3 Targeted analysis (quantification and analysis of variance)**

$^1\text{H}$  NMR spectroscopy coupled with chemometrics analysis was applied to identify the organic compounds in guaraná samples from different clones. Initially, the identification of the main organic compounds was performed in the MeOD/D<sub>2</sub>O (70:30) extract of guaraná seed powder. In general, all samples comprised a high level of caffeine. Fig. 16 illustrates a representative  $^1\text{H}$  NMR spectrum: the region between  $\delta$  0.5 and 3.0

corresponding to aliphatic hydrogen;  $\delta$  3.0 and 5.6 for carbinolic hydrogen; aromatic and carbonylic hydrogen (around  $\delta$  6.0 and 10.0). The structures of the compounds,  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts, multiplicity, and constant coupling are described in Table 7.

Figura 16 – Fig. 16. Representative  $^1\text{H}$  NMR spectrum ( $\delta$  0.0 to 9.0 ppm) of guarana seed powder (600 MHz,  $\text{CD}_3\text{OD}-d_4 + \text{D}_2\text{O} + \text{EDTA}$ ).



Source: Prepared by the Author.

Due to the high number of identified compounds and guaraná clones (total of 56), an unsupervised chemometric method by PCA was applied to explore the  $^1\text{H}$  NMR dataset, achieving the main composition variability among the samples and the relationship between the composition and the guaraná clone.

Tabela 6 – Table 6. Proton and carbon chemical shifts, coupling constant, multiplicity and long-range heteronuclear  $^1\text{H}$ - $^{13}\text{C}$  HMBC correlations gathered for the twelve identified guaraná metabolites. (to be continued).

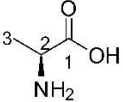
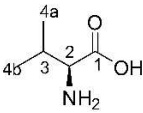
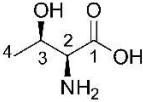
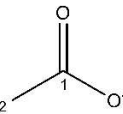
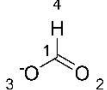
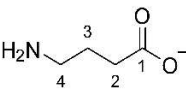
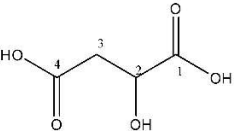
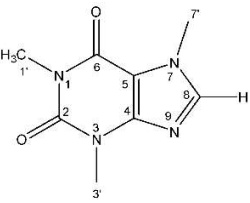
Numbered structure <sup>a</sup>	$\delta$ $^1\text{H}$ (ppm), [multiplicity, $J$ (Hz)]	$\delta$ $^{13}\text{C}$ (ppm) HSQC H $\rightarrow$ C	C-n HMBC H $\rightarrow$ C
 Alanine	- $\delta$ 3.85 (m, o, 1H-2) $\delta$ 1.49 (d, 7.0, 3H-3)	178.7 (C-1) 52.6 (C-2) 18.9 (C-3)	C-1
 Valine	- $\delta$ 3.62 (m, o, 1H-2) $\delta$ 2.33 (m, o, 1H-3) $\delta$ 1.02 (m, o, 3H-4a) $\delta$ 1.06 (m, o, 3H-4b)	no (C-1) 63.0 (C-2) 32.6 (C-3) 19.4 (C-4) 19.9 (C-4)	
 Threonine	- $\delta$ 3.64 (m, o, 1H-2) $\delta$ 4.03 (m, o, 1H-3) $\delta$ 1.30 (m, o, 3H-4)	no (C-1) 63.4 (C-3) 69.1 (C-3) 23.0 (C-3)	
 Acetic acid	$\delta$ 1.94 (s, 3H-1) -	25.8 (C-2) 179.0 (C-1)	C-1
 Formate	$\delta$ 8.46 (s, 1H-4)	175.8	
 GABA	$\delta$ 2.95 (t,o, 7.4, 2H-2) $\delta$ 1.99 (q, 7.4, 2H-3) $\delta$ 2.30 (m, o, 2H-4) -	38.0 (C-2) 27.6 (C-3) 41.8 (C-4) 178.0 (C-1)	C-1 C-4
 Malic acid	- $\delta$ 4.79 (d, 7.4, 2H-2) $\delta$ 2.55, 2.77 (m, o, 7.4, 2H-3a, 3b) -	176.1 (C-1) 38.0 (C-2) 27.6 (C-3) no (C-4)	C-1
 Caffeine <sup>b</sup>	$\delta$ 3.36 (s, 1H-1') $\delta$ 3.55 (s, 1H-3') $\delta$ 3.97 (s, 1H-7') $\delta$ 7.89 (s, 1H-8)	30.4 (C-1') 157.4 (C-2') 32.4 (C-3') 153.7 (C-4') 110.9 (C-5') 156.9 (C-6') 36.2 (C-7') 145.8 (C-8)	C-6' C-2' C-8 C-5 C-4 C-7'

Tabela 6 – Table 6. Proton and carbon chemical shifts, coupling constant, multiplicity and long-range heteronuclear  $^1\text{H}$ - $^{13}\text{C}$  HMBC correlations gathered for the twelve identified guaraná metabolites. (continuation).

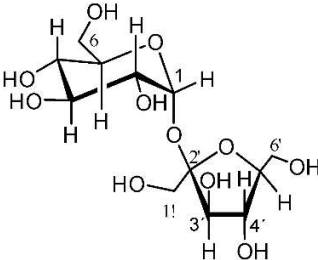
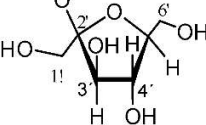
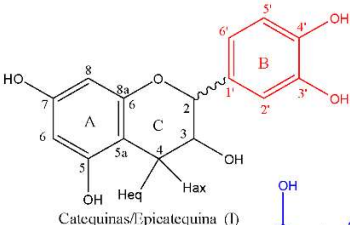
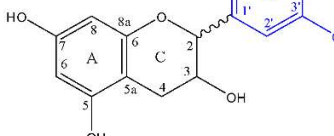
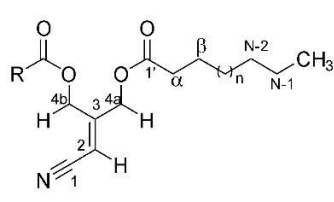
Numbered structure <sup>a</sup>	$\delta$ $^1\text{H}$ (ppm), [multiplicity, $J$ (Hz)]	$\delta$ $^{13}\text{C}$ (ppm) HSQC H→C	C-n HMBC H→C	
<b>Glucose moiety</b>				
	$\delta$ 5.40 (d, 3.5, 1H-1)	92.25 (C-1)	C-2 C-2'	
	$\delta$ 3.45 (dd, 3.5, 8.0, 1H-2)	72.8 (C-2)		
	$\delta$ 3.70 (m, o, 3.2, 1H-3)	76.2 (C-3)		
	$\delta$ 3.78 (m, o, 1H-4)	73.3 (C-4)		
	$\delta$ 3.99 (m, o, 1H-5)	76.0 (C-5)		
	$\delta$ 3.72 (dd, 3.2, 8.0, 2H-6)	63.7 (C-6)		
<b>Fructose moiety</b>				
	$\delta$ 3.62 (m, 2H-1')	64.0 (C-1')	C-2'	
	-	103.4 (C-2')		
	$\delta$ 4.07 (d, 8.0, 1H-3')	75.7 (C-3')		
	$\delta$ 4.02 (d, 8.0, 1H-4')	72.9 (C-4')		
	$\delta$ 3.80 (m, 1H-5')	102.4 (C-5')		
	$\delta$ 3.77 (m, 2H-6')	63.5 (C-6')		
<b>Ring A - Spin AB Coupling System</b>				
 <p>Catequina I / Epicatequina (I)</p>	$\delta$ 5.99/5.91 (d, 2.3, 1H-8 – I) <sup>c</sup>	95.7 (C-8)	C-7	
	$\delta$ 6.09/6.15 (d, 2.1, 1H-8 – II) <sup>c</sup>	95.7 (C-8)		
	-	158.7 (C-7)		
	$\delta$ 5.93 (d, 2.3, 1H-6 – I e II)	94.7 (C-6)	C-5 C-5a	
	-	157.0 (C-5)		
	-	157.5 (C-8a)		
	-	102.4 (C-5a)		
	<b>C ring - ABMX Spin Coupling System</b>			
	$\delta$ 4.65 (d, 8.0, 1H-2 – I e II)	81.2 (C-2')		
	$\delta$ 4.06 (m, o, 1H-3 – I e II)	69.5 (C-3')		
$\delta$ 2.89 (dd, 16.1, 8.0, 1H-4eq – I) <sup>c</sup>	27.8 (C-4')			
$\delta$ 2.88 (dd, 16.0, 8.0, 1H-4eq – II) <sup>c</sup>	27.8 (C-4')	C-3'		
$\delta$ 2.51 (dd, 16.0, 8.2, 1H-4ax – I e II)	26.9 (C-4')	C-5a'		
26.9 (C-4')				
<b>Ring B of I - ABC Spin Coupling System</b>				
 <p>Galocatequina / Gallocatequina (II)</p>	-	130.8 (C-1')		
	$\delta$ 6.84 (d, 1.5, 1H-2' – I)	115.0 (C-2')		
	-	145.8 (C-3')		
	-	144.9 (C-4')		
	$\delta$ 6.86 (d, 8.2, 1H-5' – I)	114.7 (C-5')		
	$\delta$ 6.76 (dd, 8.2, 1.5, 1H-6' – I)	118.7 (C-6')		
<b>Ring B of II</b>				
-	131.9 (C-1')			
$\delta$ 6.99 (s, 2H-2' e 6' – II)	114.3 (C-2' e 6'')	C-2 C-5'		
-	146.4 (C-3')	C-1'		
-	133.0 (C-4')			
-	146.9 (C-5')			

Tabela 6 – Table 6. Proton and carbon chemical shifts, coupling constant, multiplicity and long-range heteronuclear  $^1\text{H}$ - $^{13}\text{C}$  HMBC correlations gathered for the twelve identified guaraná metabolites. (conclusion).

Numbered structure <sup>a</sup>	$\delta$ $^1\text{H}$ (ppm), [multiplicity, $J$ (Hz)]	$\delta$ $^{13}\text{C}$ (ppm) HSQC H $\rightarrow$ C	C-n HMBC H $\rightarrow$ C
 <p>R = fatty acid chain* N = total number of carbons in the fatty acid chain n = number of <math>\text{CH}_2</math></p> <p>Cyanolipid<sup>d</sup></p>	-	118.4 (C-1)	
	$\delta$ 5.35 (s, 1H-2)	129.8 (C-2)	
	-	114.3 (C-3)	
	$\delta$ 4.85 (s, 4H-4a)	78.4 (C-4a, 4b)	C-2 C-3
	$\delta$ 2.28 (m, $\alpha\text{CH}_2$ )	$\alpha\text{CH}_2$ – (C-2')	
	$\delta$ 1.58 (m, $\beta\text{CH}_2$ )	$\beta\text{CH}_2$ – (C-3')	
	$\delta$ 1.27 (m, $\text{CH}_2$ )	31.3 [C-(N-2)]	
	$\delta$ 1.28 (m, $\text{CH}_2$ )	22.3 [C-(N-1)]	
	$\delta$ 0.87 (s, $\text{CH}_3$ )	13.2	C-(N-1) C-(N-2) <sup>d</sup>

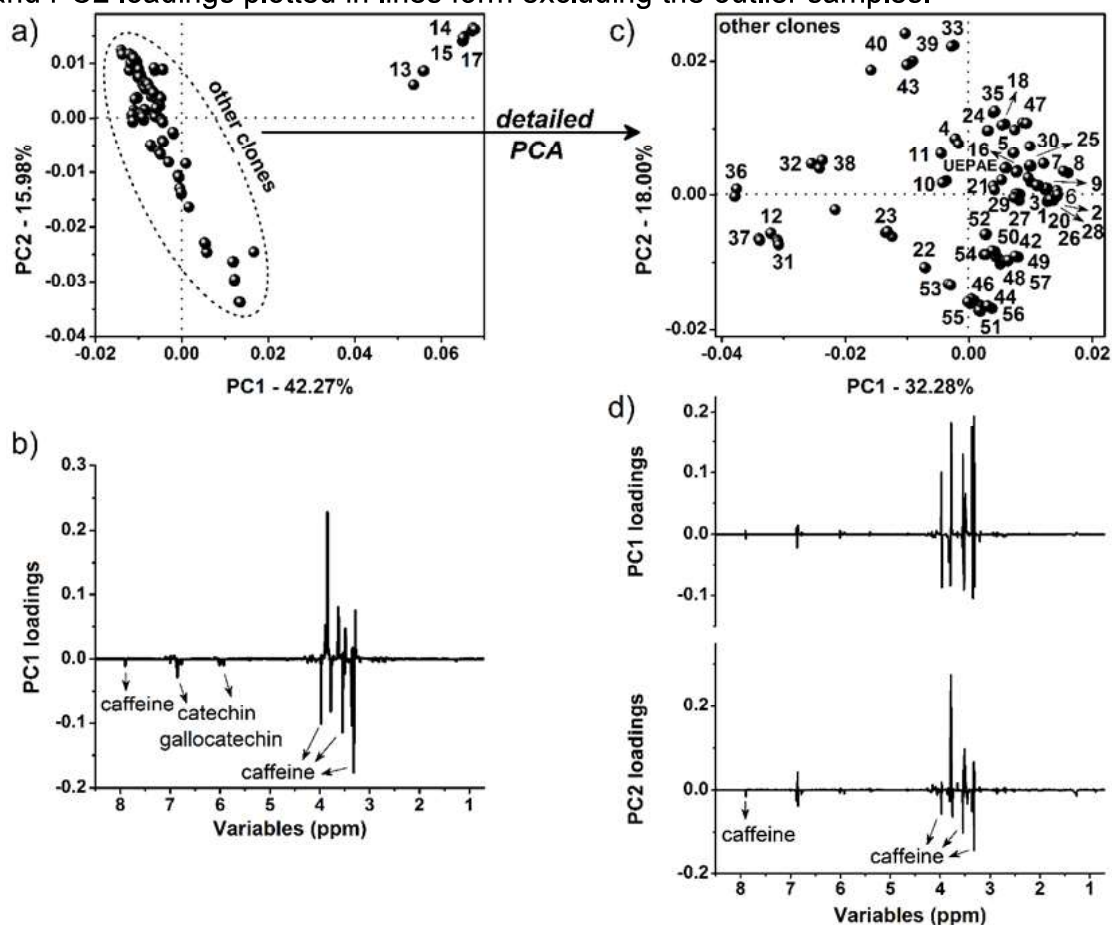
Source: Prepared by the Author.

Abbreviations - The resonance multiplicity of protons: singlet, doublet, quartet, multiplet for s, d, q and m, respectively; GABA: Gamma-aminobutyric acid; eq: equatorial, ax, axial; Absence of the expected sign (-); signal overlap; (O); unidentified: ni; C-n: numbering given to carbon in the structure of the molecule.

Remarks – <sup>a</sup> All spectroscopic data collected are in accordance with data from the Spectral Database for Organic Compounds, SDBS ([https://sdb.sdb.aist.go.jp/sdb/cgi-bin/cre\\_index.cgi](https://sdb.sdb.aist.go.jp/sdb/cgi-bin/cre_index.cgi)) and Human Metabolome Database, HMDB (<https://hmdb.ca/>); <sup>b</sup> The spectroscopic data of catechin, sucrose, and caffeine were compared with the analytical standard and the chemical shift used for the quantification step of the relative concentrations of the metabolites is highlighted in bold; <sup>c</sup> Most of the signals from the A and C rings of I and II are in a superposition, however, some differences in the proton and carbon assignments were possible through couplings via heteronuclear correlation map (qHSQC and qHMBC); <sup>d</sup> The size of the cyanolipid fatty acid ( $N^*$ ) chain is not fully understood.

Fig. 17(a) illustrates the PC1  $\times$  PC2 scores coordinate system that retained the main information regarding the study aim with 58.25% of the total data variance. The relevant loadings are plotted in line form in Fig. 17(b). Additionally, to improve the composition variability among the guaraná clones, a second PCA was developed excluding the samples numbered 13, 14, 15, and 17, since these guaraná seed powders were considered outliers based on the Hotelling's  $T^2 \times Q$  residuals and leverage  $\times$  studentized residuals plots, and therefore, these samples impaired the visualization of the data variability from the other samples. Fig. 17(c) illustrates the PC1  $\times$  PC2 scores coordinate system from this detailed PCA evaluation that retained 50.28% of the total data variance, with respective loadings plotted in lines form in Figure 5(d).

Figura 17 – Fig. 17. PCA results using the  $^1\text{H}$  NMR dataset ( $\delta$  0.7 and 8.7) of the guaraná seed powder: a) PC1  $\times$  PC2 scores coordinate system from the total samples; b) PC1 loadings plotted in lines form of the total samples; c) PC1  $\times$  PC2 scores coordinate system excluding the outliers' samples (13, 14, 15 and 17); d) PC1 and PC2 loadings plotted in lines form excluding the outlier samples.



Source: Prepared by the Author.

It was clear that caffeine was the main compound for discrimination of guaraná clones, considering the entire  $^1\text{H}$  NMR spectra followed by catechin and gallic acid. The clones numbered 13, 14, 15, and 17 at positive PC1 and PC2 scores provided samples with fewer amounts of these compounds mentioned above (caffeine, catechin, and gallic acid). The additional PCA developed excluding the clones 13, 14, 15, and 17 revealed that caffeine was also the main compound for samples discrimination according to the PC2 axis. In general, the samples from the clones 42, 44, 46, 48, 49, 50, 51, 52, 53, 54,

55, 56, and 57 presented high amounts of caffeine considering all the compounds detected by the  $^1\text{H}$  NMR signals ( $\delta$  0.7 and 8.7).

In order to complement and corroborate the composition variability of the samples according to the guaraná clone, quantification by  $^1\text{H}$  NMR ( $^1\text{H}$  qNMR) of each organic compound with a non-overlapped signal was developed. Figure 6 presents concentrations (%) of acetic acid (a), alanine (b), caffeine (c), gallic acid (d), malic acid (e), sucrose (f), and total catechin (g) in guaraná seed powder.

Caffeine variability (Fig. 6(c)) detected by PCA analyses was corroborated by quantitative analyses, revealing its high concentration (above average) in guaraná samples from clones 42, 44, 46, 48, 49, 50, 51, 52, 53, 54, 55, 56, and 57 (as described by the PCA), in addition to complementing the results with clones 11, 27, 31, 32, 36, and 37. Fascinating data obtained by employing PCA analysis and confirming by analysis of variance demonstrated that clones 13, 14, 15, and 17 have a low percentage of caffeine. This information is very relevant, as seeds with low caffeine content are of great interest to the food industry. Thus, these clones can be selected by genetic improvement systems to generate new guaraná cultivars that meet industry demands regarding the development of products with low caffeine content.

According to the literature, guaraná contains a significantly higher amount of caffeine than other foods, such as coffee, cocoa, and yerba tea, respectively about 4, 30, and 10 times higher. Excessive consumption of caffeine can lead to some adverse reactions such as increased heart rate, increased blood pressure, stress, anxiety, reduced fertility, and insomnia (Santana; Macedo, 2018). In addition, studies show that consumption of more than 200 mg of caffeine per day can cause problems such as tachycardia, ventricular arrhythmia, and seizures disease states including type 2 diabetes mellitus, Parkinson's disease, liver disease, stroke risk, Alzheimer's disease, some cancers and depression (Yousefi *et al.*, 2017). These are some of the possible effects. Therefore, guaraná with low levels of caffeine prevents these

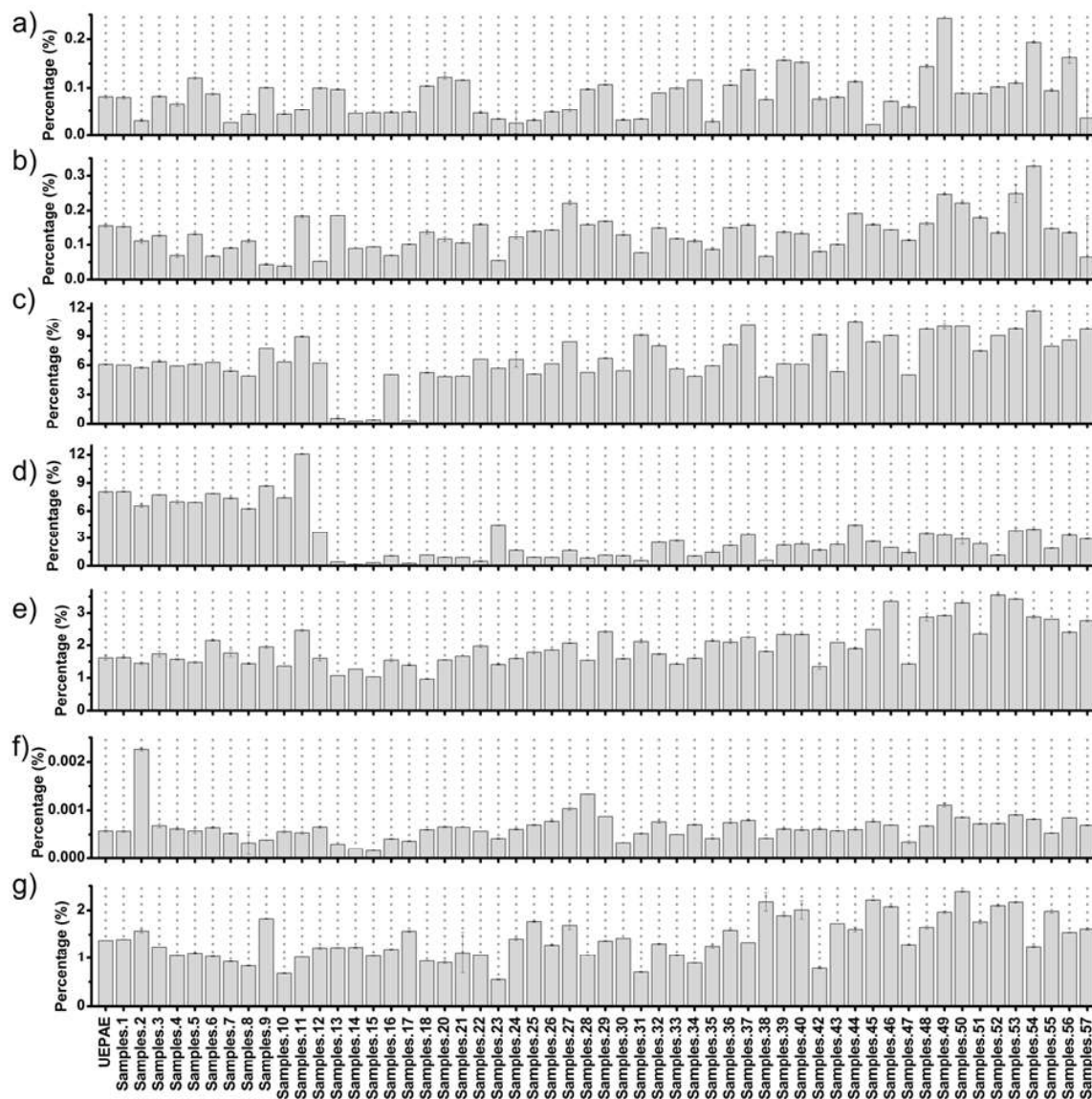


adverse reactions from being triggered by excessive consumption (Patrick *et al.*, 2019; Tfouni *et al.*, 2007).

Furthermore, considering the concentration means, guaraná samples from the clones 46, 48, 49, 50, 52, 53, 54, 55, 56, and 57 showed higher concentrations of gallic catechin and total catechin in samples from the clones UEPAE, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, and 11 (Fig. 18a). Moreover, catechins exert antioxidant activity, acting in the destruction of free radicals and neutralizing active ions of transition metals. In addition, because of the many phenolic hydroxyl groups present in their structures, they have also demonstrated the ability to prevent cardiovascular metabolic diseases and have positive effects on lipid metabolism (Portella *et al.*, 2013; Yonekura *et al.*, 2016). Additionally, studies have shown that ingesting catechin from green tea can increase fat oxidation during exercise. The magnitude of the increase is usually between 3% and 7% (about 250 – 600 kJ/day) (Bag *et al.*, 2022; Hodgson; Croft, 2010). Other beneficial health effects are attributed to catechins, such as antimicrobial, antiviral, anti-inflammatory, anti-allergic, and anticancer action. This bioactive compound can contribute to the treatment and prevention of various diseases, infectious or not. Furthermore, catechins have been widely studied for their ability to prevent premature aging and boost immunity (Bae *et al.*, 2020; Musial; Kuban-Jankowska; Gorska-Ponikowska, 2020).

In this study, we also verified that the clones 49, 54, and 56 presented higher concentrations of acetic acid and malic acid in samples from clones 27, 28, and 49 (Fig. 18(a)). In general, organic acids, such as acetic and malic, are widely used in the food industry as acidulants, flavoring, and/or food preservatives. They can also play an important role as natural antimicrobials, aiming to inhibit the growth and proliferation of microbial pathogens (Bevilacqua *et al.*, 2023; Bushell *et al.*, 2019).

Figura 18 – Fig. 18. Quantification using  $^1\text{H}$  NMR: acetic acid (a), alanine (b), caffeine (c), gallic acid (d), malic acid (e), sucrose (f), and total catechin (g) in guaraná seed powder from different clones.



Source: Prepared by the Author.

Finally, a higher concentration of amino acid alanine was observed in samples from the clones 49, 50, 53, and 54 as well as a higher concentration of sucrose in samples from the clones 9, 38, 39, 40, 45, 46, 49, 50, 52, 53, and 55.

## 4 Conclusions

Through the non-target metabolomics approach, the metabolic profile of fifty-six guaraná seed clones was established by UPLC-QTOF-MS<sup>E</sup>, so that a total of 19 metabolites were duly annotated, including caffeine and procyanidins. On the other hand, through the target metabolomics approach, metabolites in guarana seeds were identified and quantified by NMR.

The large volume of data obtained by the UPLC-QTOF-MS<sup>E</sup> and NMR analyses was submitted for chemometric analysis. Thus, the use of unsupervised multivariate analysis tools such as PCA allowed the evaluation of similarities and differences between clones. Furthermore, it was possible to infer some specific characteristics of some of the fifty-six guaraná clones studied, such as the presence, absence, and different concentration levels of certain metabolites in some samples, such as caffeine.

Finally, the set of analyzes methods carried out (UPLC-QTOF-MS<sup>E</sup>, NMR, and chemometric tools) allowed obtaining valuable information about the different chemical compositions of guaraná seeds. Thus, contributing to the selection of the best clones with different purposes, enabling an adequate use in the different types of industry, such as food industry and pharmaceutical industry, in addition to the possibility of being a potential source of bioactive compounds.

## Author Contributions

Tamyris de Aquino Gondim: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing; Jhonyson Arruda Carvalho Guedes: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing; Elenilson de Godoy Alves Filho: Conceptualization, Investigation, Software, Validation, Visualization, Writing – original draft; Gisele Silvestre da Silva:

Conceptualization, Data curation, Formal Analysis, Methodology, Software, Supervision, Writing – original draft; Gisele Simone Lopes: Conceptualization, Investigation, Methodology, Resources, Supervision, Validation, Writing – review & editing; Natasha Veruska dos Santos Nina: Funding acquisition, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing; Firmino José do Nascimento Filho: Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Visualization, Writing – review & editing; André Luiz Atroch: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing; Guilherme Julião Zocolo: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

### **Conflicts of interest**

There are no conflicts to declare.

### **Acknowledgements**

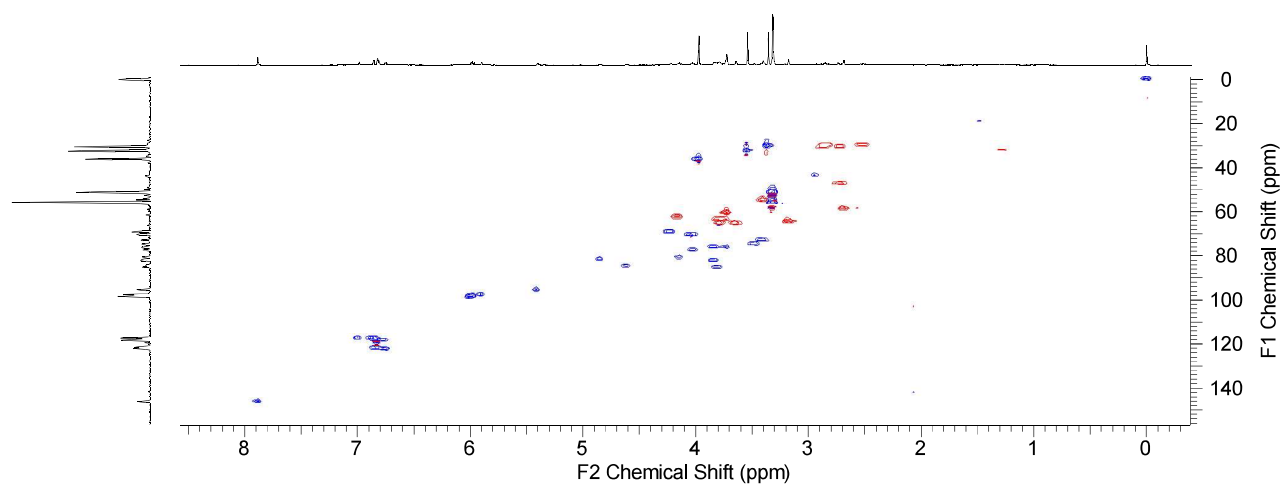
The authors gratefully acknowledge the financial support from the CNPq, National Council for Scientific and Technological Development (303791/2016-0), INCT BioNat, National Institute of Science and Technology (grant # 465637/2014-0). We would also like to thank Embrapa (SEG 03.14.01.012.00.00 and SEG 10.19.00.035.00.00). Part of the work was financed by FAPEAM – Fundação de Amparo à Pesquisa do Estado do Amazonas, through EDITAL N°. 004/2018 - AMAZONAS STRATEGIC, Project title: Conservation and use of the collection of guarana tree genotypes in Amazonas. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (PROEX 23038.000509/2020-82).

**Electronic supplementary information (ESI)****Metabolomic approaches to explore chemodiversity in seeds of guaraná****(*Paullinia cupana*) using UPLC-QTOF-MS<sup>E</sup> and NMR analysis**

Tamyris de Aquino Gondim, Jhonyson Arruda Carvalho Guedes, Elenilson de Godoy Alves Filho, Gisele Silvestre da Silva, Natasha Veruska dos Santos Nina, Firmino José do Nascimento Filho, André Luiz Atroch, Gilvan Ferreira Da Silva, Gisele Simone Lopes and Guilherme Julião Zocolo

Representative 1D and 2D NMR spectra of guaraná seeds, highlighting the resonance signals of catechin, gallic acid and caffeine

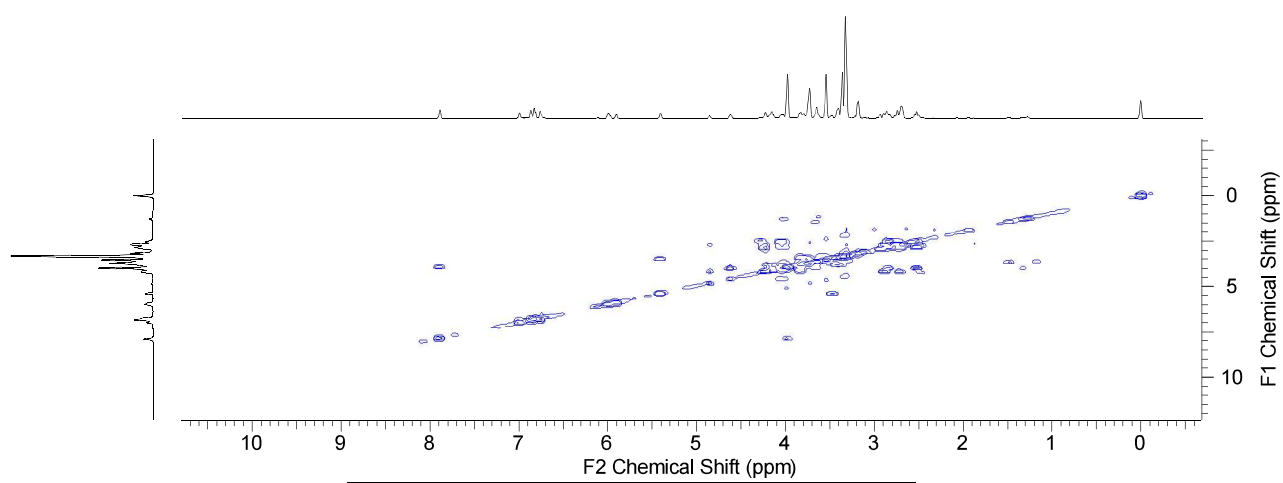
Figura 19 – Fig. S1.  $^1\text{H}$ - $^{13}\text{C}$  HSQC contour maps (600/150 MHz,  $\text{CD}_3\text{OD-}d_4 + \text{D}_2\text{O} + \text{EDTA}$ ) of guaraná seeds.



Parameter	Value
Acquisition time (sec)	(0.1500, 0.0066)
Frequency (MHz)	(599.56, 150.76)
Nucleus	( $^1\text{H}$ , $^{13}\text{C}$ )
Number of transients	32
Original points count	(1442, 200)
Points count	(2048, 1024)
Pulse sequence	gHSQCAD
Solvent	$\text{CD}_3\text{OD-}d_4 + \text{D}_2\text{O}$
Spectrum type	HSQC
Sweep width (Hz)	(9610.69, 30125.10)

Source: Prepared by the Author.

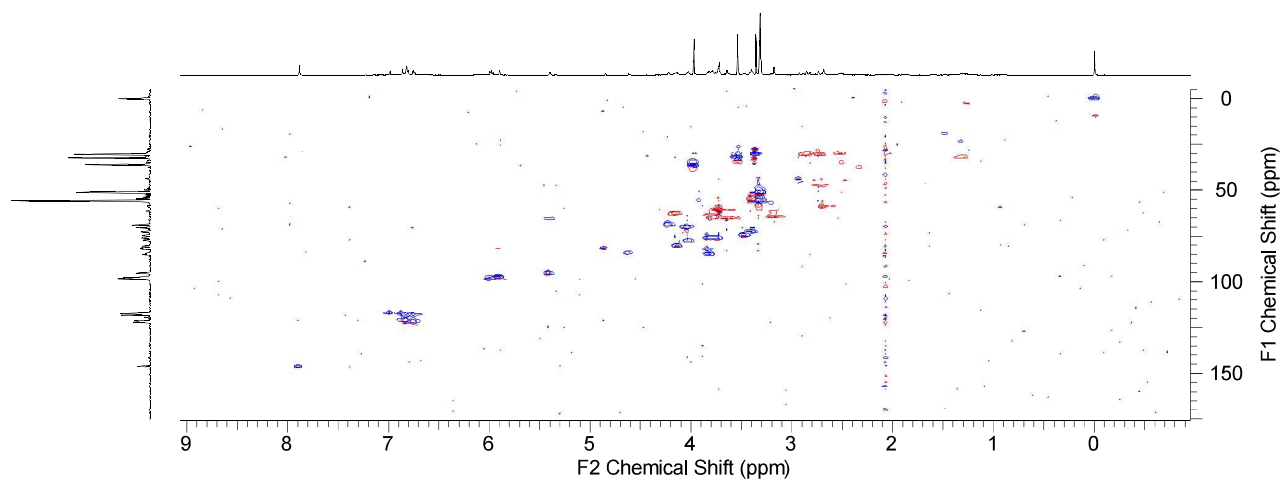
Figura 20 – Fig. S2.  $^1\text{H}$ - $^1\text{H}$  COSY contour maps ( $\text{CD}_3\text{OD-}d_4 + \text{D}_2\text{O} + \text{EDTA}$ ) of guaraná seeds.



Parameter	Value
Acquisition time (sec)	(0.1500, 0.0208)
Frequency (MHz)	(599.56, 599.56)
Nucleus	( $^1\text{H}$ , $^1\text{H}$ )
Number of transients	16
Original points count	(1442, 200)
Points count	(2048, 2048)
Pulse sequence	gCOSY
Solvent	$\text{CD}_3\text{OD-}d_4 + \text{D}_2\text{O}$
Spectrum type	COSY
Sweep width (Hz)	(9610.69, 9610.69)

Source: Prepared by the Author.

Figura 21 – Fig. S3.  $^1\text{H}$ - $^{13}\text{C}$  HMBC contour maps (600/150 MHz,  $\text{CD}_3\text{OD}-d_4 + \text{D}_2\text{O} + \text{EDTA}$ ) of guaraná seeds.

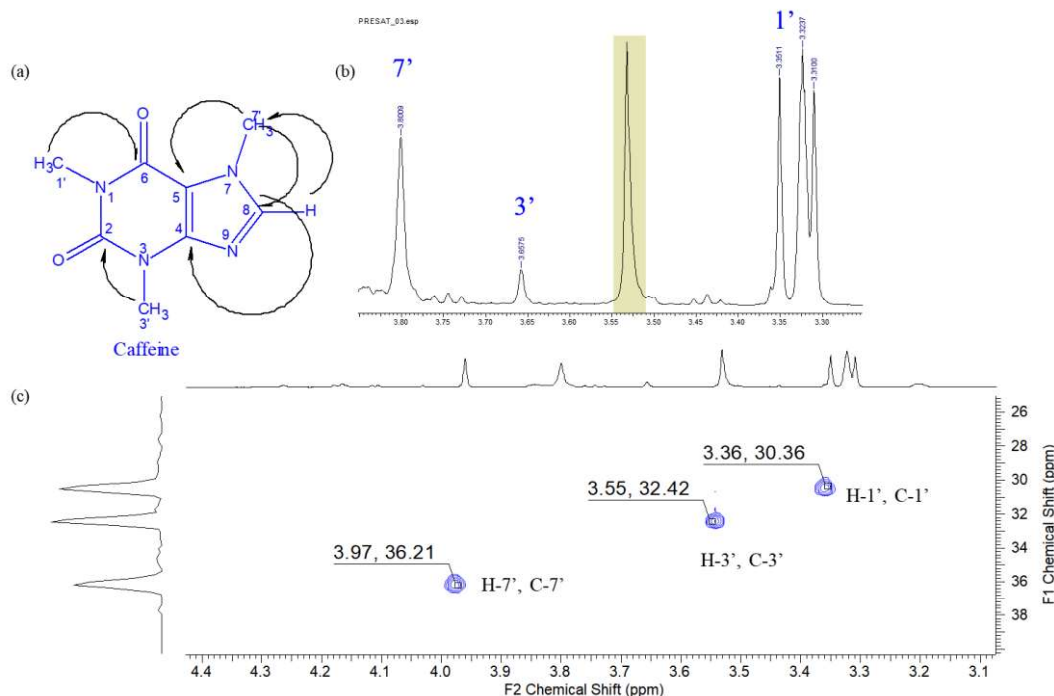


Parameter	Value
Acquisition time (sec)	(0.1500, 0.0066)
Frequency (MHz)	(599.56, 150.76)
Nucleus	( $^1\text{H}$ , $^{13}\text{C}$ )
Number of transients	32
Original points count	(1442, 200)
Points count	(2048, 1024)
Pulse sequence	gHSQCAD
Solvent	$\text{CD}_3\text{OD}-d_4 + \text{D}_2\text{O}$
Spectrum type	HSQC
Sweep width (Hz)	(9610.69, 30125.10)

Source: Prepared by the Author.

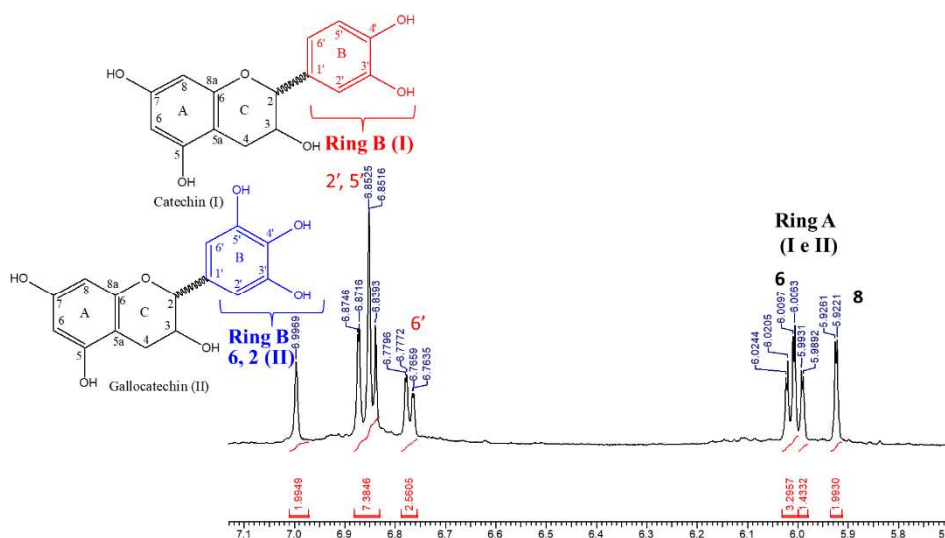


Figura 22 – Fig. S4. (a) Caffeine molecule - arrows indicate the Key HMBC correlations (H → C) observed in a representative sample of guaraná seeds, (b) expansion of PRESAT  $^1\text{H}$  NMR spectrum ( $\delta$  3.30 to 3.90 ppm) and (c) Expansion of  $^1\text{H}$ - $^{13}\text{C}$  HSQC contour map selected region ( $\delta$  3.1 to 4.4 ppm).



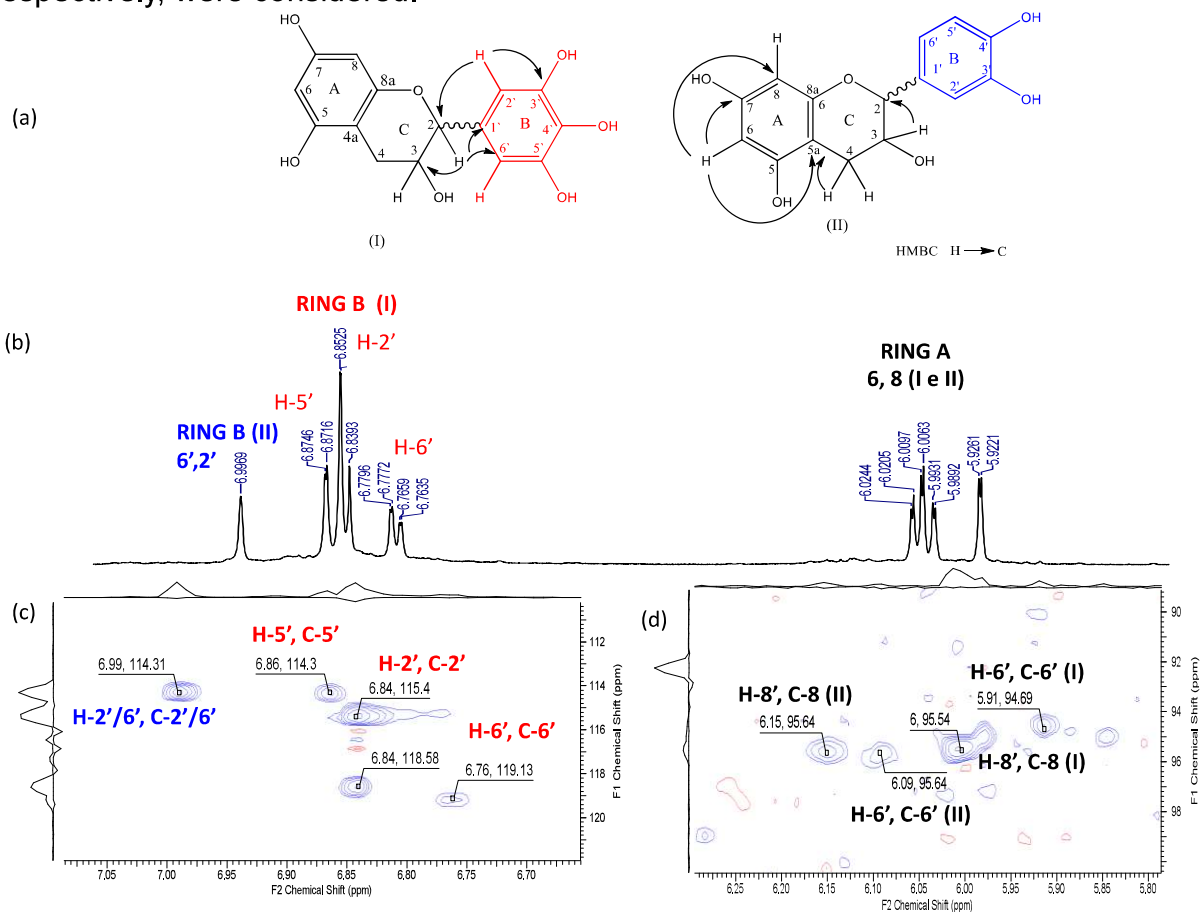
Source: Prepared by the Author.

Figura 23 – Fig. S5. Expansion of the representative  $^1\text{H}$  NMR spectrum ( $\delta$  5.70 to 7.10 ppm) highlighting catechin and gallocatechin A and B-ring signals.



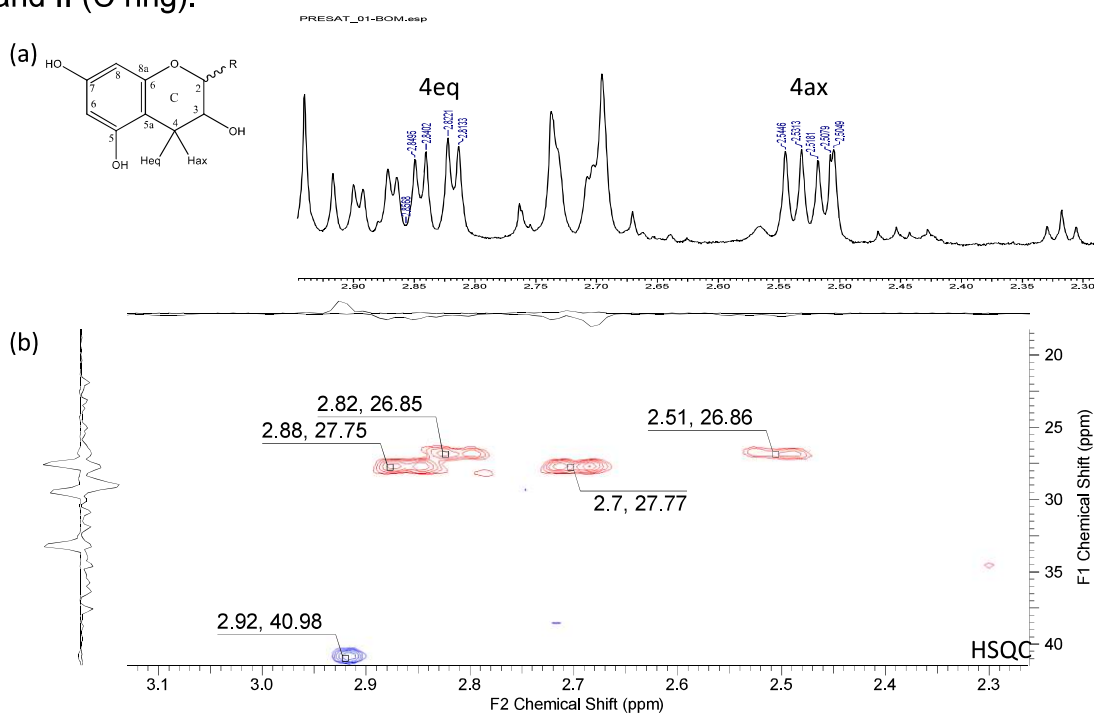
Source: Prepared by the Author.

Figura 24 – Fig. S6. (a) Numbered catechin and gallocatechin structure - arrows indicate the Key HMBC correlations observed in a representative sample, (b)  $^1\text{H}$  NMR spectrum expansion ( $\delta$  5.70 to 7.10 ppm) and  $^1\text{H}$ - $^{13}\text{C}$  HSQC heteronuclear correlation map in the region at  $\delta$  5.80 – 6.30 and 6.60 – 7.10 ppm regions (c and d, respectively). In the quantification step, the signals in ring B  $\delta$  6.86 and 6.99 ppm for I and II, respectively, were considered.



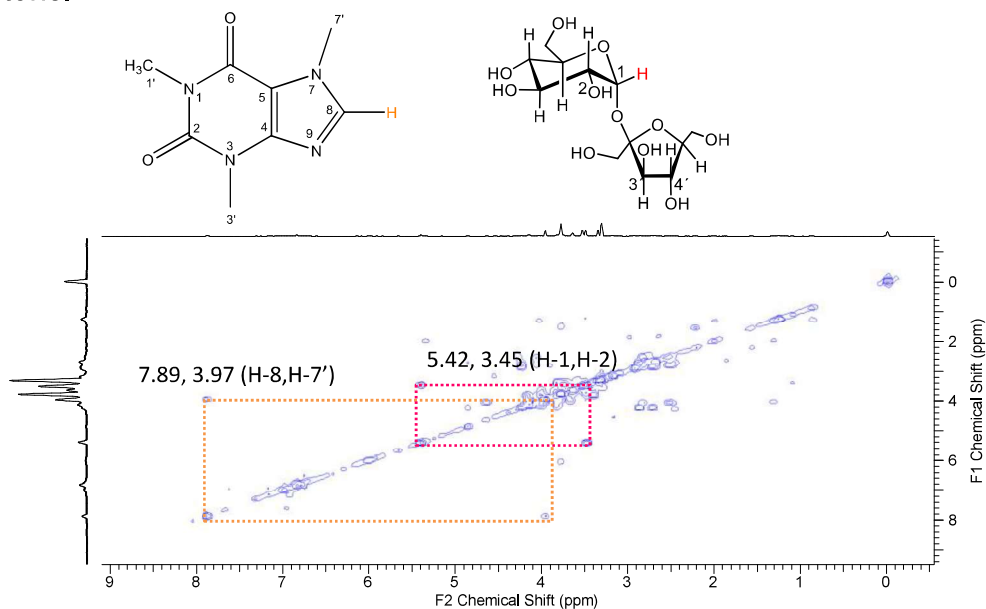
Source: Prepared by the Author.

Figura 25 – Fig. S7. Expansion of the representative  $^1\text{H}$ - $^{13}\text{C}$  HSQC contour map of guaraná seeds ( $\delta$  2.1 – 3.1 ppm) highlighting the equatorial (eq) and axial (ax) signals of I and II (C ring).



Source: Prepared by the Author.

Figura 26 – Fig. S8. A representative  $^1\text{H}$ - $^1\text{H}$  COSY contour map of guaraná seeds highlighting specific correlations between caffeine (H-8 and H-7') and sucrose (H-1 and H-2) protons.



Source: Prepared by the Author.

## 6 CONCLUSÃO

Metabolômica é um campo interdisciplinar que está intimamente relacionada a várias áreas de conhecimentos, como química, biologia e informática, sendo um elo importante no estudo das plantas. Sendo assim, por meio da abordagem metabolômica foi possível traçar o perfil metabólico de sementes de cinquenta e seis clones diferentes de guaraná e de folhas de abacaxizeiro de sete variedades comerciais.

Através da avaliação dos extratos hidroetanólicos das sete variedades comerciais de folhas de abacaxi analisadas por UPLC-QTOF-MS<sup>E</sup>, foi possível estabelecer o perfil metabólico das amostras, onde foram detectados trinta e quatro metabólitos, dos quais vinte e oito foram devidamente anotados. Dentre as classes de metabólitos podemos destacar aminoácidos, ácidos orgânicos e compostos fenólicos.

Além disso, a análise utilizando ICP-OES permitiu a determinação e quantificação dos oligoelementos (Zn, Cr, Cu, Cd, Mn, P e Fe), sendo importante ressaltar que Cr e Cd não foram observados nas amostras analisadas. Somado a esses fatos, temos também que os extratos apresentaram baixos níveis de citotoxicidade. Com base na composição mineral e metabólica das folhas das diferentes variedades de abacaxi avaliadas foi possível verificar as semelhanças e dissimilaridades entre as amostras através de ferramentas quimiométricas como PCA e HCA. Assim o perfil metabólico gerou informações importantes na caracterização de resíduos agroindustriais de abacaxi. Corroborando com a caracterização e consequente uma possível utilização de coprodutos agroindustriais de abacaxi em aplicações em diferentes tipos de indústrias, como farmacológica, cosmética e alimentar.

Além disso, a abordagem metabolômica não-alvo foi de suma importância para estabelecer os perfis metabólicos das sementes dos cinquenta e seis clones distintos de guaraná, os quais foram obtidos por UPLC-QTOF-MS<sup>E</sup>. Com isso, um total de 19 metabólitos foram devidamente anotados, incluindo cafeína e procianidinas. Por outro lado, através da abordagem da metabolômica alvo, os metabólitos nas sementes de guaraná foram identificados e quantificados por RMN.

O grande volume de dados provenientes das análises metabolômicas do UPLC-QTOF-MS<sup>E</sup> e do RMN, foram submetidos à análise quimiométrica. Assim, através da PCA se pode avaliar semelhanças e diferenças entre clones. Desse modo,

foi possível inferir algumas características específicas de alguns dos cinquenta e seis clones de guaraná estudados, como a presença, ausência e diferentes níveis de concentração de determinados metabólitos (por exemplo a cafeína) em algumas amostras. Contribuindo assim para a seleção dos melhores clones com diferentes finalidades, possibilitando uma utilização adequada nos diversos tipos de indústria. Como exemplo podemos citar a indústria alimentícia, a qual de posse dessas informações pode selecionar determinados clones para desenvolver produtos com maiores ou menores teores de cafeína, visando atender diferentes públicos, tais como crianças, adultos e idosos.

Diante disso, em geral, podemos concluir que tanto as folhas de abacaxizeiro como as sementes de guaraná podem se constituir, respectivamente como coproduto e um produto de interesse agroindustrial. Haja vista que ambos podem ser utilizados como uma fonte potencial de compostos bioativos e para o desenvolvimento de novos produtos, os quais podem encontrar aplicações em variados setores industriais, como farmacêutico, cosmético e alimentício.

## REFERÊNCIAS

- ABOUZEID, E. M. *et al.* Comprehensive metabolite profiling of *Phoenix rupicola* pulp and seeds using UPLC-ESI-MS/MS and evaluation of their estrogenic activity in ovariectomized rat model. **Food Research International**, [s. l.], v. 157, p. 111308, 2022.
- ABU-REIDAH, I. M. *et al.* HPLC–DAD–ESI-MS/MS screening of bioactive components from *Rhus coriaria* L. (Sumac) fruits. **Food Chemistry**, [s. l.], v. 166, p. 179–191, 2015. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0308814614008978>.
- AHMED, O. H. *et al.* Economic Viability of Pineapple Residues Recycling. **Journal of Sustainable Agriculture**, [s. l.], v. 21, n. 4, p. 129–137, 2003. Disponível em: [https://www.tandfonline.com/doi/full/10.1300/J064v21n04\\_07](https://www.tandfonline.com/doi/full/10.1300/J064v21n04_07).
- AKWU, N. A.; NAIDOO, Y.; SINGH, M. A comparative study of the proximate, FTIR analysis and mineral elements of the leaves and stem bark of *Grewia lasiocarpa* E.Mey. ex Harv.: An indigenous southern African plant. **South African Journal of Botany**, [s. l.], v. 123, p. 9–19, 2019.
- ALONSO, A.; MARSAL, S.; JULIÀ, A. Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. **Frontiers in Bioengineering and Biotechnology**, [s. l.], v. 3, n. MAR, 2015. Disponível em: [http://www.frontiersin.org/Bioinformatics\\_and\\_Computational\\_Biology/10.3389/fbioe.2015.00023/abstract](http://www.frontiersin.org/Bioinformatics_and_Computational_Biology/10.3389/fbioe.2015.00023/abstract).
- ALSEEKH, S. *et al.* Mass spectrometry-based metabolomics: a guide for annotation, quantification and best reporting practices. **Nature Methods**, [s. l.], v. 18, n. 7, p. 747–756, 2021.
- ALSEEKH, S.; FERNIE, A. R. Metabolomics 20 years on: what have we learned and what hurdles remain?. **The Plant Journal**, [s. l.], v. 94, n. 6, p. 933–942, 2018.
- ALVES FILHO, E. G. *et al.* <sup>1</sup>H NMR spectra dataset and solid-state NMR data of cowpea (*Vigna unguiculata*). **Data in Brief**, [s. l.], v. 11, p. 136–146, 2017. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S2352340917300136>.
- ANSES. **ANSES-CIQUAL: French food composition table version 2020**. [S. l.], 2020. Disponível em: <https://ciqual.anses.fr/>. Acesso em: 26 dez. 2022.
- ATANASOV, A. G. *et al.* Natural products in drug discovery: advances and opportunities. **Nature Reviews Drug Discovery**, [s. l.], v. 20, n. 3, p. 200–216, 2021.
- AZI, F. *et al.* Metabolite dynamics and phytochemistry of a soy whey-based beverage bio-transformed by water kefir consortium. **Food Chemistry**, [s. l.], p. 128225, 2020.
- BAE, J. *et al.* Activity of catechins and their applications. **Biomedical Dermatology**, [s. l.], v. 4, n. 1, p. 8, 2020.
- BAG, S. *et al.* Tea and its phytochemicals: Hidden health benefits & modulation of

signaling cascade by phytochemicals. **Food Chemistry**, [s. l.], v. 371, p. 131098, 2022. Disponível em:  
<https://linkinghub.elsevier.com/retrieve/pii/S030881462102104X>.

BALDISSEROTTO, A. *et al.* Moringa oleifera Leaf Extracts as Multifunctional Ingredients for “Natural and Organic” Sunscreens and Photoprotective Preparations. **Molecules**, [s. l.], v. 23, n. 3, p. 664, 2018.

BANERJEE, S. *et al.* Valorisation of pineapple wastes for food and therapeutic applications. **Trends in Food Science & Technology**, [s. l.], v. 82, p. 60–70, 2018.

BARROS, S. de S. *et al.* Pineapple (*Ananas comosus*) leaves ash as a solid base catalyst for biodiesel synthesis. **Bioresource Technology**, [s. l.], v. 312, p. 123569, 2020. Disponível em:  
<https://linkinghub.elsevier.com/retrieve/pii/S0960852420308415>.

BASKARAN, R.; PULLENCHERI, D.; SOMASUNDARAM, R. Characterization of free, esterified and bound phenolics in custard apple (*Annona squamosa* L) fruit pulp by UPLC-ESI-MS/MS. **Food Research International**, [s. l.], v. 82, p. 121–127, 2016.

BEEBE, K. R.; PELL, R. J.; SEASHOLTZ, M. B. **Chemometrics: A Practical Guide**. 1. ed. [S. l.]: Wiley-Interscience, 1998.

BELEW, Z. M. *et al.* Transport engineering in microbial cell factories producing plant-specialized metabolites. **Current Opinion in Green and Sustainable Chemistry**, [s. l.], v. 33, p. 100576, 2022.

BELINATO, J. *et al.* Metabolômica microbiana: inovações e aplicações. **Química Nova**, [s. l.], 2019. Disponível em:  
[http://quimicanova.s bq.org.br/audiencia\\_pdf.asp?aid2=6899&nomeArquivo=RV20180487.pdf](http://quimicanova.s bq.org.br/audiencia_pdf.asp?aid2=6899&nomeArquivo=RV20180487.pdf).

BEVILACQUA, A. *et al.* Using regression and Multifactorial Analysis of Variance to assess the effect of ascorbic, citric, and malic acids on spores and activated spores of *Alicyclobacillus acidoterrestris*. **Food Microbiology**, [s. l.], v. 110, p. 104158, 2023.

BHAMBHANI, S.; KONDHARE, K. R.; GIRI, A. P. Diversity in Chemical Structures and Biological Properties of Plant Alkaloids. **Molecules**, [s. l.], v. 26, n. 11, p. 3374, 2021.

BONADIMAN, B. da S. R. *et al.* Guarana (*Paullinia cupana*): Cytoprotective effects on age-related eye dysfunction. **Journal of Functional Foods**, [s. l.], v. 36, p. 375–386, 2017. Disponível em:  
<https://linkinghub.elsevier.com/retrieve/pii/S1756464617304152>.

BOUBAKRI, H. *et al.* Phenolic composition as measured by liquid chromatography/mass spectrometry and biological properties of Tunisian barley. **International Journal of Food Properties**, [s. l.], p. 1–15, 2017.

BOUFRIDI, A.; QUINN, R. J. Turning Metabolomics into Drug Discovery. **Journal of the Brazilian Chemical Society**, [s. l.], v. 27, n. 8, p. 1334–1338, 2016.

BRITO, T. B. N. *et al.* Antimicrobial, Antioxidant, Volatile And Phenolic Profiles Of Cabbage-Stalk And Pineapple-Crown Flour Revealed By GC-MS And UPLC-MS<sup>E</sup>. **Food Chemistry**, [s. l.], p. 127882, 2020.

BUSHELL, F. M. L. *et al.* Synergistic Impacts of Organic Acids and pH on Growth of *Pseudomonas aeruginosa*: A Comparison of Parametric and Bayesian Non-parametric Methods to Model Growth. **Frontiers in Microbiology**, [s. l.], v. 9, 2019. Disponível em: <https://www.frontiersin.org/article/10.3389/fmicb.2018.03196/full>.

CABRAL, J. R. S.; JUNGHANS, D. T. Variedades de Abacaxi. **Embrapa Mandioca e Fruticultura**, [s. l.], p. 4, 2003. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/81569/1/Circular-Tecnica-63-Variedade-Abacaxi-Renato-Cabral-2003.pdf>.

CANUTO, G. A. B. *et al.* Metabolômica: definições, estado-da-arte e aplicações representativas. **Química Nova**, [s. l.], v. 41, n. 1, p. 75–91, 2018. Disponível em: <http://www.scielo.br/pdf/qn/v41n1/0100-4042-qn-41-01-0075.pdf>.

CARLA CADONÁ, F. *et al.* Guarana a Caffeine-Rich Food Increases Oxaliplatin Sensitivity of Colorectal HT-29 Cells by Apoptosis Pathway Modulation. **Anti-Cancer Agents in Medicinal Chemistry**, [s. l.], v. 16, n. 8, p. 1055–1065, 2016. Disponível em: <http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1871-5206&volume=16&issue=8&page=1055>.

CARNEIRO, A. M. *et al.* Soya agricultural waste as a rich source of isoflavones. **Food Research International**, [s. l.], v. 130, n. April 2019, p. 108949, 2020. Disponível em: <https://doi.org/10.1016/j.foodres.2019.108949>.

CAXITO, M. L. C. *et al.* *In Vitro* Antileukemic Activity of *Xanthosoma sagittifolium* (Taioba) Leaf Extract. **Evidence-Based Complementary and Alternative Medicine**, [s. l.], v. 2015, p. 1–10, 2015. Disponível em: <http://www.hindawi.com/journals/ecam/2015/384267/>.

CHAGAS-PAULA, D. A. *et al.* A Metabolomic Approach to Target Compounds from the Asteraceae Family for Dual COX and LOX Inhibition. **Metabolites**, [s. l.], v. 5, n. 3, p. 404–430, 2015. Disponível em: <http://www.mdpi.com/2218-1989/5/3/404/>.

CHAKRABORTY, A. J. *et al.* Bromelain a Potential Bioactive Compound: A Comprehensive Overview from a Pharmacological Perspective. **Life**, [s. l.], v. 11, n. 4, p. 317, 2021. Disponível em: <https://www.mdpi.com/2075-1729/11/4/317>.

CHEN, A. *et al.* Production of renewable fuel and value-added bioproducts using pineapple leaves in Costa Rica. **Biomass and Bioenergy**, [s. l.], v. 141, p. 105675, 2020.

CHIN, Y.-W. *et al.* Drug discovery from natural sources. **The AAPS Journal**, [s. l.], v. 8, n. 2, p. E239–E253, 2006. Disponível em: <http://www.springerlink.com/index/10.1007/BF02854894>.

CHOI, J. S. *et al.* Effects of C-glycosylation on anti-diabetic, anti-Alzheimer's disease and anti-inflammatory potential of apigenin. **Food and Chemical Toxicology**, [s. l.], v. 64, p. 27–33, 2014.



- CONTINI, C. *et al.* Does attitude moderate the effect of labelling information when choosing functional foods?. **Food Quality and Preference**, [s. l.], v. 106, p. 104795, 2023.
- CORBO, M. R. *et al.* Functional Beverages: The Emerging Side of Functional Foods. **Comprehensive Reviews in Food Science and Food Safety**, [s. l.], v. 13, n. 6, p. 1192–1206, 2014.
- CORREIA, P. R. M.; FERREIRA, M. M. C. Reconhecimento de padrões por métodos não supervisionados: explorando procedimentos quimiométricos para tratamento de dados analíticos. **Química Nova**, [s. l.], v. 30, n. 2, p. 481–487, 2007. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-40422007000200042&lng=pt&nrm=iso&tlng=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422007000200042&lng=pt&nrm=iso&tlng=pt).
- CREN-OLIVÉ, C. *et al.* Catechin and epicatechin deprotonation followed by <sup>13</sup>C NMR. **Tetrahedron Letters**, [s. l.], v. 43, n. 25, p. 4545–4549, 2002. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0040403902007451>.
- DA SILVA, G. *et al.* <sup>1</sup>H quantitative nuclear magnetic resonance and principal component analysis as tool for discrimination of guarana seeds from different geographic regions of Brazil. *In*: PROCEEDINGS OF THE XIII INTERNATIONAL CONFERENCE ON THE APPLICATIONS OF MAGNETIC RESONANCE IN FOOD SCIENCE. [S. l.]: IM Publications, 2016. p. 21. Disponível em: [https://www.impopen.com/mrfs-abstract/M16\\_0021](https://www.impopen.com/mrfs-abstract/M16_0021).
- DA SILVA, G. S. *et al.* Chemical profiling of guarana seeds (*Paullinia cupana*) from different geographical origins using UPLC-QTOF-MS combined with chemometrics. **Food Research International**, [s. l.], v. 102, n. June, p. 700–709, 2017. Disponível em: <https://doi.org/10.1016/j.foodres.2017.09.055>.
- DAGLIA, M. *et al.* Untargeted and targeted methodologies in the study of tea (*Camellia sinensis* L.). **Food Research International**, [s. l.], v. 63, p. 275–289, 2014.
- DE LEÓN-SOLIS, C.; CASASOLA, V.; MONTERROSO, T. Metabolomics as a tool for geographic origin assessment of roasted and green coffee beans. **Heliyon**, [s. l.], v. 9, n. 11, p. e21402, 2023.
- DEL GIGLIO, A. B. *et al.* Purified Dry Extract of *Paullinia cupana* (Guaraná) (PC-18) for Chemotherapy-Related Fatigue in Patients with Solid Tumors: An Early Discontinuation Study. **Journal of Dietary Supplements**, [s. l.], v. 10, n. 4, p. 325–334, 2013. Disponível em: <http://www.tandfonline.com/doi/full/10.3109/19390211.2013.830676>.
- DENG, B. *et al.* Effects of nitrogen availability on mineral nutrient balance and flavonoid accumulation in *Cyclocarya paliurus*. **Plant Physiology and Biochemistry**, [s. l.], v. 135, p. 111–118, 2019.
- DETTMER, K.; ARONOV, P. A.; HAMMOCK, B. D. Mass spectrometry-based metabolomics. **Mass Spectrometry Reviews**, [s. l.], v. 26, n. 1, p. 51–78, 2007.
- DIFONZO, G. *et al.* Characterisation and classification of pineapple (*Ananas comosus* [L.] Merr.) juice from pulp and peel. **Food Control**, [s. l.], v. 96, p. 260–270,

2019.

DJOUMBOU FEUNANG, Y. *et al.* ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. **Journal of Cheminformatics**, [s. l.], v. 8, n. 1, p. 61, 2016. Disponível em: <https://jcheminf.biomedcentral.com/articles/10.1186/s13321-016-0174-y>.

DO, N. H. N. *et al.* Advanced fabrication and application of pineapple aerogels from agricultural waste. **Materials Technology**, [s. l.], v. 35, n. 11–12, p. 807–814, 2020. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/10667857.2019.1688537>.

DORNELES, I. M. P. *et al.* Guarana (*Paullinia cupana*) presents a safe and effective anti-fatigue profile in patients with chronic kidney disease: A randomized, double-blind, three-arm, controlled clinical trial. **Journal of Functional Foods**, [s. l.], v. 51, p. 1–7, 2018. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1756464618305231>.

DUDZIK, D. *et al.* Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. **Journal of Pharmaceutical and Biomedical Analysis**, [s. l.], 2018.

DUEÑAS, M. *et al.* Effects of different industrial processes on the phenolic composition of white and brown teff (*Eragrostis tef* (Zucc.) Trotter). **Food Chemistry**, [s. l.], v. 335, p. 127331, 2021. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0308814620311936>.

ELLIS, D. I. *et al.* Metabolic fingerprinting as a diagnostic tool. **Pharmacogenomics**, [s. l.], v. 8, n. 9, p. 1243–1266, 2007.

EMBRAPA. **Abacaxi BRS Ajubá**. [S. l.], 2009. Disponível em: <https://www.embrapa.br/en/busca-de-solucoes-tecnologicas/-/produto-servico/1424/abacaxi-brs-ajuba>. Acesso em: 4 nov. 2023.

EMBRAPA. **Abacaxi BRS Imperial**. [S. l.], 2004. Disponível em: <https://www.embrapa.br/en/busca-de-solucoes-tecnologicas/-/produto-servico/7346/abacaxi-brs-imperial>. Acesso em: 4 nov. 2023.

EMBRAPA. **Abacaxi BRS Vitória**. [S. l.], 2006. Disponível em: <https://www.embrapa.br/en/busca-de-solucoes-tecnologicas/-/produto-servico/1425/abacaxi-brs-vitoria>. Acesso em: 4 nov. 2023.

EMBRAPA. **Abacaxi Pérola**. [S. l.], 2015. Disponível em: <https://www.embrapa.br/en/busca-de-imagens/-/midia/2276037/abacaxi-perola>. Acesso em: 4 nov. 2023.

EMBRAPA. **Abacaxi Smooth Cayenne (havaiano)**. [S. l.], 2021. Disponível em: <https://www.embrapa.br/en/busca-de-imagens/-/midia/5455023/abacaxi-smooth-cayenne-havaiano>. Acesso em: 4 nov. 2023.

EMBRAPA. **Cultivar de abacaxi MD-2 ou Gold**. [S. l.], 2021. Disponível em: <https://www.embrapa.br/en/busca-de-imagens/-/midia/5856014/cultivar-de-abacaxi->

md-2-ou-gold. Acesso em: 4 nov. 2023.

EMBRAPA. **Fusariose do Abacaxizeiro**. [S. l.], 2021. Disponível em: [EZURUIKE, U. F.; PRIETO, J. M. The use of plants in the traditional management of diabetes in Nigeria: Pharmacological and toxicological considerations. \*\*Journal of Ethnopharmacology\*\*, \[s. l.\], v. 155, n. 2, p. 857–924, 2014.](https://www.embrapa.br/en/busca-de-imagens/-/midia/5858008/fusariose-do-abacaxizeiro#:~:text=Fusariose do abacaxi%2C causada pelo,forma de controle da doena. Acesso em: 4 nov. 2023.</a></p></div><div data-bbox=)

FAN, C. *et al.* Multi-ingredients determination and fingerprint analysis of leaves from *Ilex latifolia* using ultra-performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry. **Journal of Pharmaceutical and Biomedical Analysis**, [s. l.], v. 84, p. 20–29, 2013.

FANG, C.; FERNIE, A. R.; LUO, J. Exploring the Diversity of Plant Metabolism. **Trends in Plant Science**, [s. l.], v. 24, n. 1, p. 83–98, 2019.

FAOSTAT. **Crops and livestock products**. [S. l.], 2020. Disponível em: <https://www.fao.org/faostat/en/#data/QCL/visualize>. Acesso em: 26 dez. 2022.

FAOSTAT. **Crops and livestock products**. [S. l.], 2021. Disponível em: <https://www.fao.org/faostat/en/#data/QCL/visualize>. Acesso em: 5 nov. 2023.

FARAG, M. A. *et al.* Metabolite profiling of three *Opuntia ficus-indica* fruit cultivars using UPLC-QTOF-MS in relation to their antioxidant potential. **Food Bioscience**, [s. l.], v. 36, p. 100673, 2020. Disponível em: <https://doi.org/10.1016/j.fbio.2020.100673>.

FARAG, M. A. *et al.* Metabolomic fingerprints of 21 date palm fruit varieties from Egypt using UPLC/PDA/ESI-qTOF-MS and GC-MS analyzed by chemometrics. **Food Research International**, [s. l.], v. 64, p. 218–226, 2014. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0963996914004165>.

FEIZI, N. *et al.* Recent trends in application of chemometric methods for GC-MS and GC×GC-MS-based metabolomic studies. **TrAC Trends in Analytical Chemistry**, [s. l.], v. 138, p. 116239, 2021.

FIEHN, O. Combining Genomics, Metabolome Analysis, and Biochemical Modelling to Understand Metabolic Networks. **Comparative and Functional Genomics**, [s. l.], v. 2, n. 3, p. 155–168, 2001. Disponível em: <http://www.hindawi.com/journals/ijg/2001/914970/abs/>.

FIEHN, O. Metabolomics - The link between genotypes and phenotypes. **Plant Molecular Biology**, [s. l.], v. 48, n. 1–2, p. 155–171, 2002.

FORCISI, S. *et al.* Liquid chromatography – mass spectrometry in metabolomics research: Mass analyzers in ultra high pressure liquid chromatography coupling. **Journal of Chromatography A**, [s. l.], v. 1292, p. 51–65, 2013. Disponível em: <http://dx.doi.org/10.1016/j.chroma.2013.04.017>.

FU, Y.-J. *et al.* Enzyme assisted extraction of luteolin and apigenin from pigeonpea

[*Cajanuscajan* (L.) Millsp.] leaves. **Food Chemistry**, [s. l.], v. 111, n. 2, p. 508–512, 2008.

FU, C. *et al.* Phytochemical analysis and geographic assessment of flavonoids, coumarins and sesquiterpenes in *Artemisia annua* L. based on HPLC-DAD quantification and LC-ESI-QTOF-MS/MS confirmation. **Food Chemistry**, [s. l.], v. 312, p. 1–10, 2020. Disponível em: <https://doi.org/10.1016/j.foodchem.2019.126070>.

FUNARI, C. S. *et al.* Metabolômica, uma abordagem otimizada para exploração da biodiversidade brasileira: estado da arte, perspectivas e desafios. **Química Nova**, [s. l.], v. 36, n. 10, p. 1605–1609, 2013. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-40422013001000019&lng=pt&nrm=iso&tlng=en](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422013001000019&lng=pt&nrm=iso&tlng=en).

FUNASAKI, M. *et al.* Amazon Rainforest Cosmetics: Chemical Approach for Quality Control. **Química Nova**, [s. l.], 2016.

GOBBO-NETO, L.; LOPES, N. P. Plantas medicinais: Fatores de influência no conteúdo de metabólitos secundários. **Química Nova**, [s. l.], v. 30, n. 2, p. 374–381, 2007.

GONZÁLEZ-RIANO, C. *et al.* Recent developments along the analytical process for metabolomics workflows. **Analytical Chemistry**, [s. l.], v. 92, n. 1, p. 203–226, 2020.

GREMBECKA, M.; SZEFER, P. Comparative assessment of essential and heavy metals in fruits from different geographical origins. **Environmental Monitoring and Assessment**, [s. l.], v. 185, n. 11, p. 9139–9160, 2013.

GUEDES, J. A. C. *et al.* Comparative analyses of metabolic fingerprint integrated with cytotoxic activity and *in silico* approaches of the leaves extract of *Spondias mombin* L. and *Spondias tuberosa* Arr. Cam. from Northeast, Brazil. **Phytochemistry Letters**, [s. l.], v. 40, n. September, p. 26–36, 2020. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S187439002030642X>.

GUEDES, J. A. C. **Estudo do perfil metabolômico de folhas de cajazeira, umbuzeiro e abacaxizeiro e sua correlação com potencial atividade anticâncer por meio de análise multivariada**. 2018. 154 f. - Universidade Federal do Ceará, [s. l.], 2018. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/7556065><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC394507><http://dx.doi.org/10.1016/j.humphath.2017.05.005><https://doi.org/10.1007/s00401-018-1825-z><http://www.ncbi.nlm.nih.gov/pubmed/27157931>.

GUEDES, J. A. C. *et al.* Metabolic profile and cytotoxicity of non-polar extracts of pineapple leaves and chemometric analysis of different pineapple cultivars. **Industrial Crops and Products**, [s. l.], v. 124, n. August, p. 466–474, 2018. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0926669018307258>.

HASSAN, M. A. *et al.* Health benefits and phenolic compounds of *Moringa oleifera* leaves: A comprehensive review. **Phytomedicine**, [s. l.], v. 93, p. 153771, 2021.

HODGSON, J. M.; CROFT, K. D. Tea flavonoids and cardiovascular health.

**Molecular Aspects of Medicine**, [s. l.], v. 31, n. 6, p. 495–502, 2010.

HÖGNER, C. *et al.* Development and validation of a rapid ultra-high performance liquid chromatography diode array detector method for *Vitex agnus-castus*. **Journal of Chromatography B**, [s. l.], v. 927, n. 0, p. 181–190, 2013. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1570023213001402>.

HU, C.; XU, G. Mass-spectrometry-based metabolomics analysis for foodomics. **TrAC - Trends in Analytical Chemistry**, [s. l.], v. 52, p. 36–46, 2013. Disponível em: <http://dx.doi.org/10.1016/j.trac.2013.09.005>.

HUANG, M. *et al.* Identification and quantification of phenolic compounds in *Vitex negundo* L. var. *cannabifolia* (Siebold et Zucc.) Hand.-Mazz. using liquid chromatography combined with quadrupole time-of-flight and triple quadrupole mass spectrometers. **Journal of Pharmaceutical and Biomedical Analysis**, [s. l.], v. 108, p. 11–20, 2015.

IBÁÑEZ, C. *et al.* The role of direct high-resolution mass spectrometry in foodomics. **Analytical and Bioanalytical Chemistry**, [s. l.], v. 407, n. 21, p. 6275–6287, 2015. Disponível em: <http://link.springer.com/10.1007/s00216-015-8812-1>.

ISAH, T. Stress and defense responses in plant secondary metabolites production. **Biological Research**, [s. l.], v. 52, n. 1, p. 39, 2019.

JANDRIĆ, Z. *et al.* Assessment of fruit juice authenticity using UPLC–QToF MS: A metabolomics approach. **Food Chemistry**, [s. l.], v. 148, p. 7–17, 2014. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0308814613014465>.

JOHNSON, C. H.; IVANISEVIC, J.; SIUZDAK, G. Metabolomics: beyond biomarkers and towards mechanisms. **Nature Reviews Molecular Cell Biology**, [s. l.], v. 17, n. 7, p. 451–459, 2016.

JOO, Y. H. *et al.* UPLC-QTOF-MS/MS screening and identification of bioactive compounds in fresh, aged, and browned *Magnolia denudata* flower extracts. **Food Research International**, [s. l.], v. 133, p. 109192, 2020.

JU, L. *et al.* Chemical profiling of *Houttuynia cordata* Thunb. by UPLC-Q-TOF-MS and analysis of its antioxidant activity in C2C12 cells. **Journal of Pharmaceutical and Biomedical Analysis**, [s. l.], v. 204, p. 114271, 2021. Disponível em: <https://doi.org/10.1016/j.jpba.2021.114271>.

KAMEL, R.; MOSTAFA, D. M. Rutin nanostructured lipid cosmeceutical preparation with sun protective potential. **Journal of Photochemistry and Photobiology B: Biology**, [s. l.], v. 153, p. 59–66, 2015. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1011134415002845>.

KAVUTHODI, B.; SEBASTIAN, D. Biocatalysis and Agricultural Biotechnology Biotechnological valorization of pineapple stem for pectinase production by *Bacillus subtilis* BKDS1: Media formulation and statistical optimization for submerged fermentation. **Biocatalysis and Agricultural Biotechnology**, [s. l.], v. 16, n. May, p. 715–722, 2018. Disponível em: <https://doi.org/10.1016/j.bcab.2018.05.003>.

- KESSLER, A.; KALSKE, A. Plant Secondary Metabolite Diversity and Species Interactions. **Annual Review of Ecology, Evolution, and Systematics**, [s. l.], v. 49, n. 1, p. 115–138, 2018.
- KIM, H. W. *et al.* NPClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products. **Journal of Natural Products**, [s. l.], v. 84, n. 11, p. 2795–2807, 2021.
- KITE, G. C. *et al.* Acyl Spermidines in Inflorescence Extracts of Elder (*Sambucus nigra* L., Adoxaceae) and Elderflower Drinks. **Journal of Agricultural and Food Chemistry**, [s. l.], v. 61, n. 14, p. 3501–3508, 2013.
- KUEGER, S. *et al.* High-resolution plant metabolomics: from mass spectral features to metabolites and from whole-cell analysis to subcellular metabolite distributions. **The Plant Journal**, [s. l.], v. 70, n. 1, p. 39–50, 2012.
- KUMAR, S. *et al.* Rapid qualitative and quantitative analysis of bioactive compounds from *Phyllanthus amarus* using LC/MS/MS techniques. **Industrial Crops and Products**, [s. l.], v. 69, p. 143–152, 2015. Disponível em: <http://dx.doi.org/10.1016/j.indcrop.2015.02.012>.
- LACCHINI, E.; VENEGAS-MOLINA, J.; GOOSSENS, A. Structural and functional diversity in plant specialized metabolism signals and products: The case of oxylipins and triterpenes. **Current Opinion in Plant Biology**, [s. l.], v. 74, p. 102371, 2023.
- LAI, Z. *et al.* Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics. **Nature Methods**, [s. l.], v. 15, n. 1, p. 53–56, 2018. Disponível em: <http://www.nature.com/articles/nmeth.4512>.
- LEE, Jung-Ha *et al.* The potential use of bromelain as a natural oral medicine having anticarcinogenic activities. **Food Science & Nutrition**, [s. l.], v. 7, n. 5, p. 1656–1667, 2019. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/fsn3.999>.
- LEI, Z. *et al.* Construction of an Ultrahigh Pressure Liquid Chromatography-Tandem Mass Spectral Library of Plant Natural Products and Comparative Spectral Analyses. **Analytical Chemistry**, [s. l.], v. 87, n. 14, p. 7373–7381, 2015.
- LI, Z. *et al.* Functional Endophytes Regulating Plant Secondary Metabolism: Current Status, Prospects and Applications. **International Journal of Molecular Sciences**, [s. l.], v. 24, n. 2, p. 1153, 2023.
- LI, S. *et al.* Identification of A-series oligomeric procyanidins from pericarp of Litchi chinensis by FT-ICR-MS and LC-MS. **Food Chemistry**, [s. l.], v. 135, n. 1, p. 31–38, 2012.
- LI, M. *et al.* The occurrence of tricin and its derivatives in plants. **Green Chemistry**, [s. l.], v. 18, n. 6, p. 1439–1454, 2016.
- LIANG, L. *et al.* Metabolomics applications for plant-based foods origin tracing, cultivars identification and processing: Feasibility and future aspects. **Food Chemistry**, [s. l.], v. 449, n. April, p. 139227, 2024. Disponível em: <https://doi.org/10.1016/j.foodchem.2024.139227>.

LIM, P. K.; JULCA, I.; MUTWIL, M. Redesigning plant specialized metabolism with supervised machine learning using publicly available reactome data. **Computational and Structural Biotechnology Journal**, [s. l.], v. 21, p. 1639–1650, 2023.

LIU, S. *et al.* Molecular networks of secondary metabolism accumulation in plants: Current understanding and future challenges. **Industrial Crops and Products**, [s. l.], v. 201, p. 116901, 2023.

LUQUE DE CASTRO, M. D.; DELGADO-POVEDANO, M. M. Ultrasound: A subexploited tool for sample preparation in metabolomics. **Analytica Chimica Acta**, [s. l.], v. 806, p. 74–84, 2014.

LV, H.-P. *et al.* Bioactive compounds from Pu-erh tea with therapy for hyperlipidaemia. **Journal of Functional Foods**, [s. l.], v. 19, p. 194–203, 2015.

MA, C. *et al.* Characterization of active phenolic components in the ethanolic extract of *Ananas comosus* L. leaves using high-performance liquid chromatography with diode array detection and tandem mass spectrometry. **Journal of Chromatography A**, [s. l.], v. 1165, p. 39–44, 2007.

MA, C. *et al.* Rapid screening of potential  $\alpha$ -amylase inhibitors from *Rhodiola rosea* by UPLC-DAD-TOF-MS/MS-based metabolomic method. **Journal of Functional Foods**, [s. l.], v. 36, p. 144–149, 2017. Disponível em: <http://dx.doi.org/10.1016/j.jff.2017.06.060>.

MAEDA, H. A. Evolutionary Diversification of Primary Metabolism and Its Contribution to Plant Chemical Diversity. **Frontiers in Plant Science**, [s. l.], v. 10, 2019.

MAI, L. H. *et al.* Antivascular and anti-parasite activities of natural and hemisynthetic flavonoids from New Caledonian *Gardenia* species (Rubiaceae). **European Journal of Medicinal Chemistry**, [s. l.], v. 93, p. 93–100, 2015.

MAIA, I. da C. *et al.* Effect of solid-state fermentation over the release of phenolic compounds from brewer's spent grain revealed by UPLC-MS<sup>E</sup>. **LWT**, [s. l.], v. 133, p. 110136, 2020. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0023643820311257>.

MANETTI, L. M.; DELAPORTE, R. H.; LAVERDE JR., A. Metabólitos secundários da família bromeliaceae. **Química Nova**, [s. l.], v. 32, n. 7, p. 1885–1897, 2009. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-40422009000700035&lng=pt&nrm=iso&tling=pt](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422009000700035&lng=pt&nrm=iso&tling=pt).

MARONE, D. *et al.* Specialized metabolites: Physiological and biochemical role in stress resistance, strategies to improve their accumulation, and new applications in crop breeding and management. **Plant Physiology and Biochemistry**, [s. l.], v. 172, p. 48–55, 2022.

MARQUES, L. L. M. *et al.* Guaraná (*Paullinia cupana*) seeds: Selective supercritical extraction of phenolic compounds. **Food Chemistry**, [s. l.], v. 212, p. 703–711, 2016. Disponível em: <http://dx.doi.org/10.1016/j.foodchem.2016.06.028>.

MARQUES, L. L. M. *et al.* *Paullinia cupana*: a multipurpose plant – a review. **Revista**

**Brasileira de Farmacognosia**, [s. l.], v. 29, n. 1, p. 77–110, 2019.

MARQUES, C. F.; JUSTINO, G. C. An Optimised MS-Based Versatile Untargeted Metabolomics Protocol. **Separations**, [s. l.], v. 10, n. 5, p. 314, 2023.

MARSHALL, D. D.; POWERS, R. Beyond the paradigm: Combining mass spectrometry and nuclear magnetic resonance for metabolomics. **Progress in Nuclear Magnetic Resonance Spectroscopy**, [s. l.], v. 100, p. 1–16, 2017.

MATOS, T. K. *et al.* Integrated UPLC-HRMS, Chemometric Tools, and Metabolomic Analysis of Forage Palm (*Opuntia* spp. and *Nopalea* spp.) to Define Biomarkers Associated with Non-Susceptibility to Carmine Cochineal (*Dactylopius opuntiae*). **Journal of the Brazilian Chemical Society**, [s. l.], v. 00, n. 00, p. 1–11, 2021.

MAULIDIANI, M. *et al.* Detection of bioactive compounds in persimmon (*Diospyros kaki*) using UPLC-ESI-Orbitrap-MS/MS and fluorescence analyses. **Microchemical Journal**, [s. l.], v. 149, p. 103978, 2019.

MIKYŠKA, A. *et al.* Chemotaxonomic characterization of hop genotypes based on profiling of proanthocyanidins using liquid chromatography coupled with high-resolution accurate mass spectrometry. **Journal of Food Composition and Analysis**, [s. l.], v. 112, p. 104702, 2022.

MIZUNO, T. *et al.* Identification of anthocyanin and other flavonoids from the green–blue petals of *Puya alpestris* (Bromeliaceae) and a clarification of their coloration mechanism. **Phytochemistry**, [s. l.], v. 181, p. 112581, 2021.

MOCO, S. *et al.* Metabolomics technologies and metabolite identification. **Trends in Analytical Chemistry**, [s. l.], v. 26, n. 9, p. 855–866, 2007.

MOHD ALI, M. *et al.* Pineapple (*Ananas comosus*): A comprehensive review of nutritional values, volatile compounds, health benefits, and potential food products. **Food Research International**, [s. l.], v. 137, p. 109675, 2020. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0963996920307006>.

MOHEB, A. *et al.* Tricin biosynthesis during growth of wheat under different abiotic stresses. **Plant Science**, [s. l.], v. 201–202, p. 115–120, 2013.

MORAES, S. L. de; TOMAZ, J. C.; LOPES, N. P. Liquid chromatography–tandem mass spectrometric method for determination of the anti-inflammatory compound vicenin-2 in the leaves of *L. ericoides* Mart. **Biomedical Chromatography**, [s. l.], v. 21, n. 9, p. 925–930, 2007. Disponível em: <http://doi.wiley.com/10.1002/bmc.828>.

MOSMANN, T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. **Journal of Immunological Methods**, [s. l.], v. 65, n. 1–2, p. 55–63, 1983.

MUSHTAQ, M. Y. *et al.* Extraction for Metabolomics: Access to The Metabolome. **Phytochemical Analysis**, [s. l.], v. 25, n. 4, p. 291–306, 2014.

MUSIAL, C.; KUBAN-JANKOWSKA, A.; GORSKA-PONIKOWSKA, M. Beneficial Properties of Green Tea Catechins. **International Journal of Molecular Sciences**,



[s. l.], v. 21, n. 5, p. 1744, 2020.

NEHME, C. J. *et al.* Intraspecific variability of flavonoid glycosides and styrylpyrones from leaves of *Cryptocarya mandioccana* Meisner (Lauraceae). **Biochemical Systematics and Ecology**, [s. l.], v. 36, n. 8, p. 602–611, 2008.

NOVAKI, L. P. *et al.* Analysis of Consumer Products: Demonstrating the Power of LC-MS/MS for the Simultaneous Analysis of Caffeine and Other Methylxanthines in Guaraná Fruit Powder Extract. **Journal of Chemical Education**, [s. l.], v. 98, n. 6, p. 2083–2089, 2021.

OLDONI, T. L. C. *et al.* Bioguided extraction of phenolic compounds and UHPLC-ESI-Q-TOF-MS/MS characterization of extracts of *Moringa oleifera* leaves collected in Brazil. **Food Research International**, [s. l.], v. 125, p. 108647, 2019.

ONONAMADU, C. J. *et al.* Identification of potential antioxidant and hypoglycemic compounds in aqueous- methanol fraction of methanolic extract of *Ocimum canum* leaves. **Analytical and Bioanalytical Chemistry Research**, [s. l.], v. 6, n. 2, p. 431–439, 2019.

OUYANG, M. *et al.* Application of sparse linear discriminant analysis for metabolomics data. **Anal. Methods**, [s. l.], v. 6, n. 22, p. 9037–9044, 2014. Disponível em: <http://xlink.rsc.org/?DOI=C4AY01715C>.

OZAROWSKI, M. *et al.* Comparison of bioactive compounds content in leaf extracts of *Passiflora incarnata*, *P. caerulea* and *P. alata* and *in vitro* cytotoxic potential on leukemia cell lines. **Revista Brasileira de Farmacognosia**, [s. l.], v. 28, n. 2, p. 179–191, 2018.

PACHECO, M. T. *et al.* Determination by HPLC-DAD-ESI/MS<sup>n</sup> of phenolic compounds in Andean tubers grown in Ecuador. **Journal of Food Composition and Analysis**, [s. l.], v. 84, p. 103258, 2019.

PAPETTI, A. *et al.* HPLC–DAD–ESI/MS<sup>n</sup> characterization of environmentally friendly polyphenolic extract from *Raphanus sativus* L. var. “Cherry Belle” skin and stability of its red components. **Food Research International**, [s. l.], v. 65, n. PB, p. 238–246, 2014. Disponível em: <http://dx.doi.org/10.1016/j.foodres.2014.04.046>.

PATRICK, M. *et al.* Safety of Guarana Seed as a Dietary Ingredient: A Review. **Journal of Agricultural and Food Chemistry**, [s. l.], v. 67, n. 41, p. 11281–11287, 2019.

PAUL, A.; DE BOVES HARRINGTON, P. Chemometric applications in metabolomic studies using chromatography-mass spectrometry. **TrAC Trends in Analytical Chemistry**, [s. l.], v. 135, p. 116165, 2021.

PEER, W. A.; MURPHY, A. S. Flavonoids and auxin transport: modulators or regulators?. **Trends in Plant Science**, [s. l.], v. 12, n. 12, p. 556–563, 2007.

PEREIRA, J. B.; DANTAS, K. G. F. Evaluation of inorganic elements in cat’s claw teas using ICP OES and GF AAS. **Food Chemistry**, [s. l.], v. 196, p. 331–337, 2016.

PICHERSKY, E.; LEWINSOHN, E. Convergent Evolution in Plant Specialized Metabolism. **Annual Review of Plant Biology**, [s. l.], v. 62, n. 1, p. 549–566, 2011.

PILON, A. *et al.* Metabolômica de plantas: métodos e desafios. **Química Nova**, [s. l.], v. X, n. 00, p. 1–26, 2020. Disponível em: [http://quimicanova.s bq.org.br/audiencia\\_pdf.asp?aid2=8055&nomeArquivo=RV2019-0431.pdf](http://quimicanova.s bq.org.br/audiencia_pdf.asp?aid2=8055&nomeArquivo=RV2019-0431.pdf).

PONTES, J. G. M. *et al.* NMR-based metabolomics strategies: plants, animals and humans. **Analytical Methods**, [s. l.], v. 9, n. 7, p. 1078–1096, 2017. Disponível em: <http://xlink.rsc.org/?DOI=C6AY03102A>.

PORTELLA, R. de L. *et al.* Guaraná (*Paullinia cupana* Kunth) effects on LDL oxidation in elderly people: an *in vitro* and *in vivo* study. **Lipids in Health and Disease**, [s. l.], v. 12, n. 1, p. 12, 2013.

POTT, D. M.; OSORIO, S.; VALLARINO, J. G. From Central to Specialized Metabolism: An Overview of Some Secondary Compounds Derived From the Primary Metabolism for Their Role in Conferring Nutritional and Organoleptic Characteristics to Fruit. **Frontiers in Plant Science**, [s. l.], v. 10, 2019.

PYTLAKOWSKA, K. *et al.* Multi-element analysis of mineral and trace elements in medicinal herbs and their infusions. **Food Chemistry**, [s. l.], v. 135, n. 2, p. 494–501, 2012.

QIANG, L. *et al.* Identification of proanthocyanidins from litchi (*Litchi chinensis* Sonn.) pulp by LC-ESI-Q-TOF-MS and their antioxidant activity. **PLoS ONE**, [s. l.], v. 10, n. 3, p. 1–17, 2015.

QUIFER-RADA, P. *et al.* A comprehensive characterisation of beer polyphenols by high resolution mass spectrometry (LC-ESI-LTQ-Orbitrap-MS). **Food Chemistry**, [s. l.], v. 169, p. 336–343, 2015.

RAI, A.; SAITO, K.; YAMAZAKI, M. Integrated omics analysis of specialized metabolism in medicinal plants. **The Plant Journal**, [s. l.], v. 90, n. 4, p. 764–787, 2017.

RAMOS, P. A. B. *et al.* Chemical Characterisation, Antioxidant and Antibacterial Activities of *Pinus pinaster* Ait. and *Pinus pinea* L. Bark Polar Extracts: Prospecting Forestry By-Products as Renewable Sources of Bioactive Compounds. **Applied Sciences**, [s. l.], v. 12, n. 2, p. 784, 2022.

RAUF, A. *et al.* Proanthocyanidins: A comprehensive review. **Biomedicine & Pharmacotherapy**, [s. l.], v. 116, p. 108999, 2019.

REINHARDT, D. H. R. C. *et al.* Advances in pineapple plant propagation. **Revista Brasileira de Fruticultura**, [s. l.], v. 40, n. 6, 2018.

REN, M. *et al.* Qualitative and quantitative analysis of phenolic compounds by UPLC-MS/MS and biological activities of *Pholidota chinensis* Lindl. **Journal of Pharmaceutical and Biomedical Analysis**, [s. l.], v. 187, p. 113350, 2020.

RICO, X. *et al.* Recovery of high value-added compounds from pineapple, melon, watermelon and pumpkin processing by-products: An overview. **Food Research International**, [s. l.], v. 132, n. December 2019, p. 109086, 2020.

RODRIGUES, C. I. *et al.* Assessment of *in vitro* anthelmintic activity and bio-guided chemical analysis of BRS Boyrá pineapple leaf extracts. **Veterinary Parasitology**, [s. l.], v. 285, p. 109219, 2020.

RODRÍGUEZ-PÉREZ, C. *et al.* Comprehensive metabolite profiling of *Solanum tuberosum* L. (potato) leaves by HPLC-ESI-QTOF-MS. **Food Research International**, [s. l.], v. 112, p. 390–399, 2018.

RODRÍGUEZ-PÉREZ, C. *et al.* Optimization of extraction method to obtain a phenolic compounds-rich extract from *Moringa oleifera* Lam leaves. **Industrial Crops and Products**, [s. l.], v. 66, p. 246–254, 2015.

RODRÍGUEZ, D. *et al.* Polymorphic microsatellite markers in pineapple (*Ananas comosus* (L.) Merrill). **Scientia Horticulturae**, [s. l.], v. 156, p. 127–130, 2013.

RUE, E. A.; RUSH, M. D.; VAN BREEMEN, R. B. Procyanidins: a comprehensive review encompassing structure elucidation via mass spectrometry. **Phytochemistry Reviews**, [s. l.], v. 17, n. 1, p. 1–16, 2018. Disponível em: <http://link.springer.com/10.1007/s11101-017-9507-3>.

RUI, W. *et al.* Rapid Analysis of the Main Components of the Total Glycosides of *Ranunculus japonicus* by UPLC/Q-TOF-MS. **Natural Product Communications**, [s. l.], v. 5, n. 5, p. 1934578X1000500, 2010.

RUTZ, A. *et al.* The LOTUS initiative for open knowledge management in natural products research. **eLife**, [s. l.], v. 11, p. 1–41, 2022. Disponível em: <https://elifesciences.org/articles/70780>.

SAID, R. Ben *et al.* Tentative Characterization of Polyphenolic Compounds in the Male Flowers of *Phoenix dactylifera* by Liquid Chromatography Coupled with Mass Spectrometry and DFT. **International Journal of Molecular Sciences**, [s. l.], v. 18, n. 3, p. 512, 2017.

SALEM, M. A. *et al.* Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis. **Metabolites**, [s. l.], v. 10, n. 1, p. 37, 2020.

SALLES, R. C. de O. *et al.* Geographical origin of guarana seeds from untargeted UHPLC-MS and chemometrics analysis. **Food Chemistry**, [s. l.], v. 371, p. 131068, 2022.

SANG, J. *et al.* Combination of a deep eutectic solvent and macroporous resin for green recovery of anthocyanins from *Nitraria tangutorun* Bobr. fruit. **Separation Science and Technology**, [s. l.], v. 54, n. 18, p. 3082–3090, 2019. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/01496395.2018.1559190>.

SANTANA, Á. L.; MACEDO, G. A. Effects of hydroalcoholic and enzyme-assisted extraction processes on the recovery of catechins and methylxanthines from crude

and waste seeds of guarana (*Paullinia cupana*). **Food Chemistry**, [s. l.], v. 281, p. 222–230, 2019.

SANTANA, Á. L.; MACEDO, G. A. Health and technological aspects of methylxanthines and polyphenols from guarana: A review. **Journal of Functional Foods**, [s. l.], v. 47, p. 457–468, 2018.

SCHIMPL, F. C. *et al.* Guarana: Revisiting a highly caffeinated plant from the Amazon. **Journal of Ethnopharmacology**, [s. l.], v. 150, n. 1, p. 14–31, 2013. Disponível em: <http://dx.doi.org/10.1016/j.jep.2013.08.023>.

SENA NETO, A. R. *et al.* Comparative study of 12 pineapple leaf fiber varieties for use as mechanical reinforcement in polymer composites. **Industrial Crops and Products**, [s. l.], v. 64, p. 68–78, 2015. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0926669014006608>.

SENA NETO, A. R. *et al.* Poly(lactic acid) composites reinforced with leaf fibers from ornamental variety of hybrid pineapple (Potyra). **Polymer Composites**, [s. l.], v. 39, n. 11, p. 4050–4057, 2017. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/pc.24464>.

SEPÚLVEDA, L. *et al.* Valorization of pineapple waste for the extraction of bioactive compounds and glycosides using autohydrolysis. **Innovative Food Science & Emerging Technologies**, [s. l.], v. 47, p. 38–45, 2018. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1466856417301753>.

SHEN, S. *et al.* Metabolomics-centered mining of plant metabolic diversity and function: Past decade and future perspectives. **Molecular Plant**, [s. l.], v. 16, n. 1, p. 43–63, 2023.

SIBALY, S.; JEETAH, P. Production of paper from pineapple leaves. **Journal of Environmental Chemical Engineering**, [s. l.], v. 5, n. 6, p. 5978–5986, 2017.

SILVA, F. de A. *et al.* Diversity of cultivable fungal endophytes in *Paullinia cupana* (Mart.) Ducke and bioactivity of their secondary metabolites. **PLOS ONE**, [s. l.], v. 13, n. 4, p. e0195874, 2018.

SILVA, M. de F. G. da *et al.* Evaluation of nutritional and chemical composition of yacon syrup using  $^1\text{H}$  NMR and UPLC-ESI-Q-TOF-MS<sup>E</sup>. **Food Chemistry**, [s. l.], v. 245, p. 1239–1247, 2018. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S030881461731912X>.

SIMÕES, C. M. O. *et al.* **Farmacognosia: do produto natural ao medicamento**. Porto Alegre: Artmed, 2017.

SINGH, G.; AGRAWAL, H.; BEDNAREK, P. Specialized metabolites as versatile tools in shaping plant–microbe associations. **Molecular Plant**, [s. l.], v. 16, n. 1, p. 122–144, 2023.

SOGA, T. Advances in capillary electrophoresis mass spectrometry for metabolomics. **TrAC Trends in Analytical Chemistry**, [s. l.], v. 158, p. 116883, 2023.

SONG, Q. *et al.* Optimized flash extraction and UPLC-MS analysis on antioxidant compositions of *Nitraria sibirica* fruit. **Journal of Pharmaceutical and Biomedical Analysis**, [s. l.], v. 172, p. 379–387, 2019. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0731708519300408>.

STEINGASS, C. B. *et al.* Studies into the phenolic patterns of different tissues of pineapple (*Ananas comosus* [L.] Merr.) infructescence by HPLC-DAD-ESI-MS<sup>n</sup> and GC-MS analysis. **Analytical and Bioanalytical Chemistry**, [s. l.], v. 407, n. 21, p. 6463–6479, 2015.

SUI, Y. *et al.* Characterization and preparation of oligomeric procyanidins from *Litchi chinensis* pericarp. **Fitoterapia**, [s. l.], v. 112, p. 168–174, 2016.

SUMNER, L. W. *et al.* Proposed minimum reporting standards for chemical analysis. **Metabolomics**, [s. l.], v. 3, n. 3, p. 211–221, 2007. Disponível em: <http://link.springer.com/10.1007/s11306-007-0082-2>.

SUN, G. M. *et al.* **Nutritional Composition of Pineapple (*Ananas comosus* (L.) Merr.)**. [S. l.]: Elsevier Inc., 2016.

TEBANI, A.; AFONSO, C.; BEKRI, S. Advances in metabolome information retrieval: turning chemistry into biology. Part I: analytical chemistry of the metabolome. **Journal of Inherited Metabolic Disease**, [s. l.], v. 41, n. 3, p. 379–391, 2018. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1007/s10545-017-0074-y>.

TECHNICAL COMMITTEE. **ISO 10993-5: Biological evaluation of medical devices: part 5: tests for in vitro cytotoxicity**. [S. l.], 2009. Disponível em: <https://www.iso.org/standard/36406.html>. Acesso em: 28 set. 2022.

TFOUNI, S. A. V. *et al.* Contribuição do guaraná em pó (*Paullinia cupana*) como fonte de cafeína na dieta. **Revista de Nutrição**, [s. l.], v. 20, n. 1, p. 63–68, 2007.

THOMAS, M.; HUGHES, R. E. A relationship between ascorbic acid and threonic acid in guinea-pigs. **Food and Chemical Toxicology**, [s. l.], v. 21, n. 4, p. 449–452, 1983. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/0278691583901011>.

TOMASI, G.; VAN DEN BERG, F.; ANDERSSON, C. Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. **Journal of Chemometrics**, [s. l.], v. 18, n. 5, p. 231–241, 2004. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/cem.859>. Acesso em: 27 jan. 2023.

TSOLMON, B. *et al.* Structural identification and UPLC-ESI-QTOF-MS2 analysis of flavonoids in the aquatic plant *Landoltia punctata* and their *in vitro* and *in vivo* antioxidant activities. **Food Chemistry**, [s. l.], p. 128392, 2020.

TSUGAWA, H. *et al.* A cheminformatics approach to characterize metabolomes in stable-isotope-labeled organisms. **Nature Methods**, [s. l.], v. 16, n. 4, p. 295–298, 2019. Disponível em: <http://dx.doi.org/10.1038/s41592-019-0358-2>.

TSUGAWA, H. *et al.* Hydrogen Rearrangement Rules: Computational MS/MS Fragmentation and Structure Elucidation Using MS-FINDER Software. **Analytical Chemistry**, [s. l.], v. 88, n. 16, p. 7946–7958, 2016.

TSUGAWA, H. *et al.* MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. **Nature Methods**, [s. l.], v. 12, n. 6, p. 523–526, 2015. Disponível em: <http://www.nature.com/articles/nmeth.3393>.

VAZ JÚNIOR, S. **Aproveitamento de resíduos agroindustriais: Uma abordagem sustentável**. 1. ed. Brasília: Embrapa Agroenergia, 2020.

VERPOORTE, R.; CHOI, Y. H.; KIM, H. K. NMR-based metabolomics at work in phytochemistry. **Phytochemistry Reviews**, [s. l.], v. 6, n. 1, p. 3–14, 2007.

VIANA, E. D. S. *et al.* Caracterização físico-química de novos híbridos de abacaxi resistentes à fusariose. **Ciência Rural**, [s. l.], v. 43, n. 7, p. 1155–1161, 2013.

VILLATE, A. *et al.* Review: Metabolomics as a prediction tool for plants performance under environmental stress. **Plant Science**, [s. l.], v. 303, p. 110789, 2021.

WANG, K. *et al.* Metabolomics: A promising technique for uncovering quality-attribute of fresh and processed fruits and vegetables. **Trends in Food Science & Technology**, [s. l.], v. 142, p. 104213, 2023.

WANG, S. *et al.* Natural variance at the interface of plant primary and specialized metabolism. **Current Opinion in Plant Biology**, [s. l.], v. 67, p. 102201, 2022.

WATANABE, M. *et al.* Profiling contents of water-soluble metabolites and mineral nutrients to evaluate the effects of pesticides and organic and chemical fertilizers on tomato fruit quality. **Food Chemistry**, [s. l.], v. 169, p. 387–395, 2015.

WENG, J. The evolutionary paths towards complexity: a metabolic perspective. **New Phytologist**, [s. l.], v. 201, n. 4, p. 1141–1149, 2014.

WISHART, D. S. *et al.* HMDB 3.0—The Human Metabolome Database in 2013. **Nucleic Acids Research**, [s. l.], v. 41, n. D1, p. D801–D807, 2012. Disponível em: <http://academic.oup.com/nar/article/41/D1/D801/1055560/HMDB-30The-Human-Metabolome-Database-in-2013>.

YABOR, L. *et al.* Mineral composition of a transgenic pineapple clone grown in the field for 8 yr. **In Vitro Cellular & Developmental Biology - Plant**, [s. l.], v. 53, n. 5, p. 489–493, 2017.

YAN, M. *et al.* Separation and analysis of flavonoid chemical constituents in flowers of *Juglans regia* L. by ultra-high-performance liquid chromatography-hybrid quadrupole time-of-flight mass spectrometry. **Journal of Pharmaceutical and Biomedical Analysis**, [s. l.], v. 164, p. 734–741, 2019. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0731708518324786>.

YANG, C. *et al.* Application of metabolomics profiling in the analysis of metabolites and taste quality in different subtypes of white tea. **Food Research International**, [s. l.], v. 106, p. 909–919, 2018.

YONEKURA, L. *et al.* Bioavailability of catechins from guaraná (*Paullinia cupana*) and its effect on antioxidant enzymes and other oxidative stress markers in healthy human subjects. **Food & Function**, [s. l.], v. 7, n. 7, p. 2970–2978, 2016.

YOUSEFI, S. *et al.* A simple, effective and highly sensitive analytical method used for the determination of caffeine in tea and energy drink samples, and method optimization using a central composite design. **Analytical Methods**, [s. l.], v. 9, n. 10, p. 1665–1671, 2017. Disponível em: <http://xlink.rsc.org/?DOI=C6AY03490J>.

YUAN, L.; GROTEWOLD, E. Plant specialized metabolism. **Plant Science**, [s. l.], v. 298, p. 110579, 2020.

YULIANA, N. D. *et al.* Metabolomics for the rapid dereplication of bioactive compounds from natural sources. **Phytochemistry Reviews**, [s. l.], v. 12, n. 2, p. 293–304, 2013.

ZAYNAB, M. *et al.* Role of secondary metabolites in plant defense against pathogens. **Microbial Pathogenesis**, [s. l.], v. 124, p. 198–202, 2018.

ZHANG, L. *et al.* Antihyperglycemic, antioxidant activities of two *Acer palmatum* cultivars, and identification of phenolics profile by UPLC-QTOF-MS/MS: New natural sources of functional constituents. **Industrial Crops and Products**, [s. l.], v. 89, p. 522–532, 2016.

ZHANG, X. *et al.* HPLC/QTOF-MS/MS application to investigate phenolic constituents from *Ficus pandurata* H. aerial roots. **Biomedical Chromatography**, [s. l.], v. 29, n. 6, p. 860–868, 2015.

ZHANG, W. *et al.* Metallomics and NMR-based metabolomics of *Chlorella* sp. reveal the synergistic role of copper and cadmium in multi-metal toxicity and oxidative stress. **Metallomics**, [s. l.], v. 7, n. 3, p. 426–438, 2015. Disponível em: <https://academic.oup.com/metallomics/article/7/3/426-438/6015141>.

ZHANG, K. *et al.* Optimization of the sample preparation method for adherent cell metabolomics based on ultra-performance liquid chromatography coupled to mass spectrometry. **Analytical Methods**, [s. l.], v. 11, n. 29, p. 3678–3686, 2019. Disponível em: <http://xlink.rsc.org/?DOI=C9AY00326F>.

ZHANG, W.; TAN, N. G. J.; LI, S. F. Y. NMR-based metabolomics and LC-MS/MS quantification reveal metal-specific tolerance and redox homeostasis in *Chlorella vulgaris*. **Mol. BioSyst.**, [s. l.], v. 10, n. 1, p. 149–160, 2014.

ZHONG, M. *et al.* Santin inhibits influenza A virus replication through regulating MAPKs and NF- $\kappa$ B pathways. **Journal of Asian Natural Products Research**, [s. l.], v. 21, n. 12, p. 1205–1214, 2019.