



LUIZ HENRIQUE GALLI VARGAS

**PROSPECÇÃO DE GENES TECIDO
ESPECÍFICO E METABÓLITOS EM *Elaeis* spp.**

LAVRAS - MG

2014

LUIZ HENRIQUE GALLI VARGAS

**PROSPECÇÃO DE GENES TECIDO ESPECÍFICO E METABÓLITOS
EM *Elaeis* spp.**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para a obtenção do título de Mestre.

Orientador

Dr. Manoel Teixeira Souza Júnior

Coorientadores

Dr. Eduardo Fernandes Formighieri

Dra. Patrícia Verardi Abdelnur

LAVRAS – MG

2014

**Ficha Catalográfica Elaborada pela Coordenadoria de Produtos e
Serviços da Biblioteca Universitária da UFLA**

Vargas, Luiz Henrique Galli.

Prospecção de genes tecido específico e metabólitos em *Elaeis
spp.* / Luiz Henrique Galli Vargas. – Lavras : UFLA, 2014.

138 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2014.

Orientador: Manoel Teixeira Souza Júnior.

Bibliografia.

1. Genoma. 2. Palma de óleo. 3. Caiapé. 4. UHPLC-MS. 5.
Expressão relativa. I. Universidade Federal de Lavras. II. Título.

CDD – 633.851233

LUIZ HENRIQUE GALLI VARGAS

**PROSPECÇÃO DE GENES TECIDO ESPECÍFICO E METABÓLITOS
EM *Elaeis* spp.**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Biotecnologia Vegetal, área de concentração em Biotecnologia Vegetal, para a obtenção do título de Mestre.

APROVADO em 15 de agosto de 2014.

Dr. Eduardo Fernandes Formighieri	EMBRAPA - Agroenergia
Dra. Patrícia Verardi Abdelnur	EMBRAPA - Agroenergia
Dr. João Ricardo Moreira de Almeida	EMBRAPA - Agroenergia
Dr. Alexandre Alonso Alves	EMBRAPA - Agroenergia

Dr. Manoel Teixeira Souza Júnior
Orientador

**LAVRAS – MG
2014**

Aos meus pais, irmãos e toda família. Aos meus avós (*in memoriam*).

DEDICO

AGRADECIMENTOS

Aos meus pais, Luiz Sérgio Oliveira Vargas e Neli Galli, por permitirem a realização deste sonho, à minha “mãe” de coração, Magna Viana pelos incentivos durante os anos.

Meus irmãos, principalmente, ao Roger William Galli Vargas, que esteve ao meu lado, suportando meu humor variável, companheiro de conversa e amigo.

Meu padrasto Alexandre Ferreira e madrastra Janete Aguilar.

Ao professor Dr. Clevison Luiz Giacobbo e minha “dinda” Dra. Julcéia Camillo, por confiarem e aconselharem sempre que necessário.

A meu orientador, Dr. Manoel Teixeira Souza Júnior, por me receber, confiar e sempre estar disposto a resolver todos os problemas, meus coorientadores, Dr. Eduardo Fernandes Formighieri e Dra. Patrícia Verardi Abdelnur, importantíssimos pela disposição e sempre prestativos nas várias dúvidas referentes aos trabalhos.

A meu amigo, que aprendi a respeitar profissional e pessoalmente hoje em quem me espelho; André Pereira Leão, sem você não seria possível, pode ter certeza!

Os sinceros agradecimentos ao analista da Embrapa Agroenergia, José Antônio Ribeiro, o qual se multiplica diariamente no laboratório da Central de Análises Químicas.

Os analistas e colaboradores do LGB e da CAQ.

Aos amigos Daniel Sifuentes, João Ricardo, Marcelo Picanço, Renato Molina e Rodrigo Furtado.

A turma do “LGBBeer” que tornou esta caminhada sempre divertida, aos amigos de Lavras - MG.

Agradeço aos amigos de graduação em Engenharia Agrônômica, principalmente, àqueles que hoje considero irmãos.

À Universidade Federal de Lavras e ao Programa de Pós Graduação em Biotecnologia Vegetal pela oportunidade e a Capes pela concessão da bolsa.

A todos, muito obrigado, esta dissertação é dedicada a vocês.

*“Toda vitória tem seu
início na decisão pessoal
em conquistá-la”*

(autor desconhecido)

RESUMO GERAL

Objetivou-se neste trabalho realizar a prospecção e validação de genes que possuam expressão em tecido específico (folha) presentes no genoma da espécie *Elaeis oleifera* (Kunth) Cortés (caiaué). Também, a elaboração de um protocolo eficiente que possibilite a identificação de metabólitos secundários presentes em folhas da espécie *Elaeis guineensis* Jacq (palma de óleo) com o uso da espectrometria de massas acoplada à cromatografia líquida de Ultra Alta Eficiência (UHPLC-MS). Para a etapa de prospecção, foram realizadas buscas em bancos de dados genômicos por genes que indicavam expressão específica em folha. Após serem realizadas as etapas de alinhamento das sequências dos genes candidatos com o *draft* do genoma do caiaué, os resultados foram validados com o uso da técnica de reação em cadeia da polimerase em tempo real (qPCR). Foram testados dois tecidos (folha e raiz) para comparação da diferença de expressão. Como resultado, obteve-se indicação de maior expressão relativa do gene MyBHV5 no tecido folha quando comparado ao tecido raiz e confirmação da presença deste gene no genoma do caiaué. Enquanto com o uso da técnica de UHPLC-MS foi possível elaborar um protocolo de extração de metabólitos secundários, presentes em folhas de palma de óleo, o qual se mostrou eficiente, simples e com boa reprodutibilidade. Foi possível a identificação de possíveis compostos os quais deverão ter sua elucidação estrutural realizada em estudos futuros.

Palavras-chave: Genoma. Palma de óleo. Caiaué. UHPLC-MS. Expressão relativa.

GENERAL ABSTRACT

The aim of this study was the exploration and validation of genes that have expression in specific tissue (leaf) present in the species *Elaeis oleifera* (Kunth) Cortés (caiaué) genome. Also, the development of an efficient protocol that enables the identification of secondary metabolites in leaves of the species *Elaeis guineensis* (oil palm) with the use of mass spectrometry coupled to liquid chromatography ultra high efficiency (UHPLC-MS). To step prospecting searches in genomic databases for genes that showed specific expression in leaf data were performed. After aligning the sequences of the candidate genes in the genome caiaué be performed, the results were validated using the technique of polymerase chain reaction in real time (qPCR). Two tissues (root and leaf) to compare differences in expression were tested. As a result was obtained indicating highest relative MyBHV5 gene expression in leaf tissue compared to root tissue and confirmation of the presence of this gene in the genome caiaué. While using the UHPLC-MS technique was possible to develop a protocol for extraction of secondary metabolites in leaves of palm oil which is efficient, simple and with good repeatability. It was possible to identify secondary metabolites which must be performed structural elucidation in future studies.

Keywords: Genome. Palm oil. Caiaué. UHPLC-MS. Relative expression.

SUMÁRIO

PRIMEIRA PARTE		
1	INTRODUÇÃO	12
2	REFERENCIAL TEÓRICO	14
2.1	O gênero <i>Elaeis</i>	14
2.2	Aspectos morfológicos	14
2.3	<i>Elaeis guineensis</i> Jacq – palma de óleo	16
2.4	<i>Elaeis oleifera</i> (H. B. K) Cortés – caiaué	16
2.5	Uso de caiaué no melhoramento genético da palma de óleo	17
2.6	Bioinformática na prospecção de genes	18
2.7	<i>Data mining</i>	19
2.8	Sequência FASTA	21
2.9	BLAST (Basic Local Alignment Search Tools)	21
2.10	Blastn (<i>Nucleotide Blast</i>)	22
2.11	tBlastn	23
2.12	Blastp (<i>Protein Blast</i>)	23
2.13	Blastx	23
2.14	tBlastx	24
2.15	Extração de ácidos nucleicos	24
2.16	PCR (Reação em Cadeia da Polimerase)	26
2.17	Desenho de <i>primers</i> específicos	27
2.18	PCR quantitativo (qPCR)	28
2.19	Síntese do DNA complementar (cDNA)	29
2.20	Fluoróforos	30
2.21	Histórico espectrometria de massas	31
2.22	Espectrometria de massas (MS - <i>Mass Spectrometry</i>): definições	32
2.23	Componentes do espectrômetro de massas	33
2.24	Fontes de ionização	34
2.25	Ionização por eletrospray (ESI – <i>Electrospray Ionization</i>)	34
2.26	Analisadores de massas (<i>Mass Analyzers</i>)	36
2.27	Quadrupolo (Q – <i>Quadrupole</i>)	36
2.28	Tempo de Voo (TOF – <i>Time of Flight</i>)	37
2.29	Detectores	38
2.30	Multiplicadora de elétrons (EM – <i>Electron Multiplier</i>)	38
2.31	Espectrometria de massas sequencial – MS/MS (<i>Tandem Mass Spectrometry</i> – MS/MS)	39
2.32	Métodos cromatográficos acoplados à espectrometria de massas	40
2.33	Cromatografia Líquida de Ultra Alta Eficiência (<i>Ultra High Performance Liquid Chromatography</i> - UHPLC)	41

2.34	Cromatografia líquida acoplada à espectrometria de massas (<i>Liquid Chromatography/Mass Spectrometry</i> – LC-MS)	42
2.35	Aplicações da espectrometria de massas na determinação estrutural de biomoléculas	43
2.36	Metabolômica.....	44
2.37	Espectrometria de massas na metabolômica.....	45
2.38	Metabólitos secundários.....	47
2.39	Terpenos	48
2.40	Compostos fenólicos	49
2.41	Flavonoides.....	49
2.42	Alcaloides.....	50
	REFERÊNCIAS	51
	SEGUNDA PARTE - ARTIGOS	64
	ARTIGO 1 Prospecção e caracterização de genes tecido específico (folha) de <i>Elaeis oleifera</i> (H. B. K.) Cortés.....	64
1	INTRODUÇÃO	66
2	MATERIAL E MÉTODOS	68
3	RESULTADOS E DISCUSSÃO.....	78
4	CONCLUSÃO.....	92
	REFERÊNCIAS	94
	ARTIGO 2 Extração e análise de metabólitos secundários de folhas de <i>Elaeis guineensis</i> Jacq por cromatografia líquida de Ultra Alta Eficiência acoplada à espectrometria de massas	99
1	INTRODUÇÃO	101
2	MATERIAL E MÉTODOS	103
2.1	Obtenção do material biológico.....	103
2.2	Extrações	103
2.3	Recuperação dos extratos	105
2.4	Métodos analíticos	106
2.5	Análise de Dados.....	108
3	RESULTADOS E DISCUSSÃO.....	109
4	CONCLUSÃO.....	132
	REFERÊNCIAS	134
	CONSIDERAÇÕES FINAIS.....	137

PRIMEIRA PARTE

1 INTRODUÇÃO

A espécie americana *Elaeis oleifera* (caiaué) é importante fonte de recursos genéticos nos programas de melhoramento da palma de óleo no Brasil. O Centro Nacional de Pesquisa de Agroenergia – CNPAE (Embrapa Agroenergia) tem contribuído de forma crescente com a ampliação e divulgação do conhecimento acerca desta espécie por meio de diversas pesquisas. Dentre os resultados destas pesquisas, encontra-se o banco de dados genômicos do caiaué nos servidores do Laboratório de Bioinformática da Embrapa Agroenergia (dados não publicados), viabilizando a prospecção de genes de interesse no genoma da espécie. A validação de genes expressos, especificamente, em folhas de caiaué é uma área pouco explorada atualmente, acarretando em baixas quantidades de genes específicos de folhas descritos para esta espécie (NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION- NCBI, 2013).

Além da bioinformática, a genômica funcional contempla outras áreas, tais como a transcriptômica, proteômica e a metabolômica (OLIVER et al., 1998). Estas ciências possuem funções, para tentar compreender mudanças no funcionamento do genoma, em diferentes estágios do desenvolvimento e sob diferentes condições ambientais. No entanto, ao contrário de outras abordagens de genômica funcional, a identificação e quantificação de metabolitos de plantas, via metabolômica, têm sido em grande parte negligenciadas (FIEHN, 2002).

Assim, especificamente, o estudo do metaboloma possui um enorme potencial a ser explorado, haja vista que a espécie *Elaeis guineensis* (palma de óleo) não possui seus metabolitos secundários identificados em sua totalidade, havendo poucos estudos realizados na área da espectrometria de massas com

esta espécie. O Centro Nacional de Pesquisa de Agroenergia – CNPAE (Embrapa Agroenergia) possui estudos de caracterização de recursos genéticos com ambas as culturas citadas (*E. guineensis* e *E. oleifera*), onde diversos projetos nas áreas de genômica, genética molecular e metabolômica encontram-se em andamento com os mais variados objetivos.

A metabolômica é considerada uma área nova e de grande valia, para que haja o entendimento dos mais variados sistemas biológicos (ABDELNUR, 2011) e a elucidação do metaboloma é, particularmente, difícil em razão da natureza química de diversas moléculas pequenas. Para obtenção do sucesso na técnica, é necessário realizar um manuseio eficiente da amostra, assim como possuir boas técnicas de separação e, ainda, existe grande possibilidade de ocorrer perdas de compostos nessa etapa (HOFFMANN; STROOBANT, 2007). As técnicas analíticas, baseadas em cromatografia líquida e espectrometria de massas, vêm sendo amplamente utilizadas em metabolômica, em virtude da capacidade de separar e identificar os metabólitos polares e apolares presentes mesmo em pequenas quantidades.

Assim, é possível a caracterização de recursos genéticos sem *E. guineensis* e *E. oleifera* com a união de ferramentas de bioinformática, biotecnológicas (PCR (*Polymerase Chain Reaction*), RT-PCR (*Reverse Transcription-Polymerase Chain Reaction*) e qPCR (*Real Time Quantitative PCR*)). e a metabolômica. Estes estudos possuem grande valia no auxílio ao programa de melhoramento dessas espécies que possuem estudos na Embrapa Agroenergia.

2 REFERENCIAL TEÓRICO

2.1 O gênero *Elaeis*

O gênero *Elaeis*, pertence à classe *Liliopsida* (Monocotiledônea), ordem *Arecales* (Palmales), família *Arecaceae* (Palmaceae), subfamília *Arecoideae*, tribo *Cocoseae* (Cocoinaea), subtribo *Elaeidinae* (DRANSFIELD et al., 2005). Esse gênero é composto por duas espécies: *Elaeis guineensis* (palma de óleo), *Elaeis oleifera* (caiaué) (CUNHA et al., 2009).

2.2 Aspectos morfológicos

A palma de óleo e ocaiaú é compartilhadas muitas características semelhantes (CONCEIÇÃO; MULLER, 2000). Ambas são espécies monocotiledônias, perenes e monoicas, produzem inflorescências masculinas e femininas em uma mesma planta, em ciclos alternados, com sistema de reprodução alógamo. O caiaué apresenta uma baixa taxa de crescimento do tronco, sendo caracterizado por seu porte baixo, quando comparado com a palma de óleo (SCHWENDIMAN; PALLARES; AMBLARD, 1982; TAN, 1976).

O sistema radicular dessas espécies é fasciculado, possuindo raízes adventícias, com raízes primárias, geralmente, cerca de 6-10 mm de diâmetro, proveniente da base do tronco e raízes secundárias que se espalham em diversos ângulos no solo (CORLEY; BARETT; JONES, 1976). As raízes são formadas com base no bulbo radicular, que é um órgão de, aproximadamente, 80 cm de diâmetro, localizado na base do estipe. A densidade de raízes diminui do estipe para a periferia e a extensão máxima da parte superficial baseado no estipe depende da idade da planta (FERWERDA, 1975).

Nas folhas do caiaué, os folíolos são dispostos em um mesmo plano sobre a raque, conferindo aspecto crespo às folhas, enquanto na palma de óleo esta característica não é observada (CORLEY; TINKER, 2003). No caiaué, assim como na palma de óleo, a folha é composta por um eixo central, subdividido em parte basal, que se liga ao estipe, denominada pecíolo, cujo comprimento é bastante variável. Há centenas de espinhos na parte distal, denominada ráquis, que suporta os folíolos. Repartidos dos dois lados do ráquis encontram-se de 250 a 350 folíolos, com variados ângulos de inserção (SURRE; ZILLER, 1969). O comprimento das folhas é muito variável, uma planta adulta pode conter até 42 folhas funcionais e o comprimento varia de 4, 2 a 7, 4 m (CORLEY; TINKER, 2003; CUNHA et al., 2009).

Em relação ao aspecto reprodutivo, as flores masculinas e femininas são produzidas em diferentes ciclos sexuais (CUNHA et al., 2009). É considerada uma planta alógama com frequente autofecundação em condições naturais (cerca de 5%) (RAMOS, 2006). As inflorescências dessas espécies são protegidas pela espata, formadas na axila de cada folha. A inflorescência feminina é envolta por duas espatas, protegendo a raque floral, que apresenta comprimento de 15 a 30 cm, suportando as ráquias que medem de 4 a 9 cm de comprimento, terminadas em ápice não-pontiagudo (CUNHA et al., 2009).

Os frutos são do tipo drupa ou baga, geralmente, apresentam três camadas: exocarpo, mesocarpo e endocarpo (MIRANDA et al., 2001). Possuem uma característica esférica, possuindo de 2 a 5 cm de comprimento e 3 a 30 g, sendo o peso médio de 12, 2 g (GONÇALVES, 2001). A palma de óleo possui uma característica marcante, a maioria das plantas comerciais são híbridos F1 entre seleções de sementes do tipo dura e pisifera, dando origem ao híbrido tenera, a diferença entre elas se dá pela espessura da amêndoa (BILLOTTE et al., 1999).

As sementes possuem um endocarpo duro, uma consistência lenhosa. A superfície do endocarpo é estriada, opaca e com coloração escura. O endosperma é constituído por tegumento fino, tecido de reserva (amêndoa) e embrião, apresenta consistência rígida, oleosa e coloração branca (SURRE; ZILLER, 1969).

2.3 *Elaeis guineensis* Jacq – palma de óleo

A palma de óleo tem origem no oeste africano, concentrando-se em Guiné, Serra Leoa, Libéria, Costa do Marfim, Gana, Togo, Benim, Nigéria, Camarões, Gabão, República Democrática do Congo e Angola (RAJANAIDU, 1986). A planta, também, é conhecida como dendê, palma-de-guiné, dendém, palmeira dendém e coqueiro-dedendezeiro (LORENZI et al., 1996).

2.4 *Elaeis oleifera* (H. B. K) Cortés – caiaué

As populações naturais de caiaué podem ser encontradas desde a América Central até o norte da América do Sul (GHESQUIÈRE et al., 1997; MEUNIER; BOUTIN, 1975). Esta região é considerada o provável centro de origem do *Elaeis oleifera*, em razão da ocorrência de maior diversidade morfológica (OOI et al., 1994) e a bacia amazônica seria considerada um centro secundário de diversificação do caiaué (GHESQUIÈRE et al., 1997; MEUNIER; BOUTIN, 1975).

O caiaué tem se mostrado tolerante a pragas e doenças da palma de óleo comercial do Brasil. No entanto, seu plantio comercial não é de interesse econômico em função da baixa produtividade de óleo quando comparado à quantidade produzida pela palma de óleo (BARCELOS, 1986). Quanto às exigências agroclimáticas, o cultivo da espécie é favorecido pela ocorrência de

temperaturas entre 24 e 28°C, número de horas de insolação ao ano próximo de 2.000 e nunca inferior a 1.500 horas, umidade relativa na faixa de 75 a 90%, e a pluviosidade favorável é de 2.000 mm/ano, regularmente distribuída (BARCELOS et al., 1995; BASTOS, 2001).

2.5 Uso de caiaué no melhoramento genético da palma de óleo

Apesar do isolamento geográfico entre as duas espécies, o cruzamento da palma de óleo com o caiaué é viável, produzindo híbridos interespecíficos férteis com potencial comercial. Assim, o caiaué é utilizado como fonte de variabilidade genética nos programas de melhoramento da palma de óleo (RIOS et al., 2012). No Brasil, extensas coletas de subamostras foram realizadas na região Amazônica desde o início dos anos 80 (BASTOS, 2001).

O interesse em germoplasma de caiaué é em razão de algumas características interessantes para a palma de óleo, tais como: baixo porte o qual viabiliza um ciclo da cultura mais longo, aumentando, assim, o tempo de produção da cultura. Possui uma melhor qualidade do óleo e tolerância a doenças e pragas, incluindo o amarelecimento fatal que é o principal problema encontrado nos plantios comerciais da palma de óleo na América (AMORIM et al., 1997).

Enquanto o óleo da palma de óleo possui média de 700 ppm de teor de caroteno, o óleo do caiaué possui teor de cerca de 3.000 ppm, o caroteno é importante por ter atividade antioxidante em humanos, além de possuir uma alta concentração de vitamina E (importante na prevenção de doenças cardiovasculares), esteróis, e de tocotrienóis, os quais ajudam a reduzir os níveis de colesterol (JALANI et al., 1997; LUBRANO; ROBIN, 1997). Todavia, o caiaué, ainda, é pouco explorado, demandando intenso trabalho de seleção, o

que representa algumas décadas de processo seletivo, por se tratar de uma espécie perene.

Apesar do cruzamento interespecífico, as cultivares híbridas têm produtividade inferior à das cultivares tradicionais de palma óleo (CUNHA et al., 2012). No Brasil, iniciativas têm sido tomadas visando aproveitar melhor a variabilidade genética do caiaué em benefício da palma de óleo.

2.6 Bioinformática na prospecção de genes

Com o sequenciamento do genoma humano surgiu a “era pós-genômica”. Neste contexto, outras técnicas foram avançando como a transcriptômica, proteômica e metabolômica, objetivando o isolamento e caracterização do RNA, proteínas e os metabólitos, respectivamente. As ferramentas de bioinformática foram desenvolvidas concomitantemente. O termo “ômica” refere-se à análise global dos sistemas biológicos (PRIMROSE; TWYMAN, 2003).

Em razão do conjunto de dados produzidos, tornou-se fundamental a organização dessas informações em bancos de dados que permitam acesso *on-line*. Esses bancos de dados são classificados como primários e secundários. Os primários são formados pela deposição direta de sequências de nucleotídeos, aminoácidos ou estruturas proteicas. Os bancos de dados de secundários são aqueles que derivam dos primários, após serem submetidos a um processo de “cura”, ou seja, foram formados com as informações depositadas nos bancos de dados primários (CLOTE; BACKOFEN, 2000).

2.7 *Data mining*

A procura de sequências de nucleotídeos e de proteínas nos bancos de dados públicos é uma maneira eficiente para a obtenção de informações sobre essas moléculas. Os bancos de dados mais utilizados em bioinformática estão citados na tabela 1 (PROSDOCIMI, 2002).

Tabela 1 Lista de bancos de dados públicos (páginas da *Web*) mais conhecidos em bioinformática

Nome	Endereço eletrônico	Sigla	Função
<i>National Center for Biotechnology Information</i>	http://www.ncbi.nlm.nih.gov/	NCBI	Banco genômico de dados norte americano.
<i>European Bioinformatics Institute</i>	http://www.ebi.ac.uk/	EBI	Banco de dados europeu de sequências de DNA.
<i>DNA Data Bank of Japan</i>	http://www.ddbj.nig.ac.jp	DDBJ	Banco de dados japonês de sequências de DNA.
<i>Protein Data Bank</i>	http://www.rcsb.org/pdb	PDB	Armazena estruturas tridimensionais resolvidas de proteínas.
<i>Protein Information Resource</i>	http://www-brf.georgetown.edu/	PIR	Banco de proteínas anotadas.
<i>Protein knowledgebase</i>	http://www.expasy.ch/spro/	SWISS-PROT	Armazena sequências de proteínas e suas respectivas características moleculares, anotado manualmente por uma equipe de especialistas.
<i>Kyoto Encyclopedia of Genes and Genomes</i>	http://www.genome.ad.jp/kegg/	KEGG	Banco com dados de sequências de genomas de vários organismos diferentes e informações relacionadas às suas vias metabólicas.
<i>Protein sequence analysis & classification</i>	http://www.ebi.ac.uk/interpro/	INTERPRO	Banco de dados de famílias, domínios e assinaturas de proteínas.

Fonte: Adaptado de Prosdocimi (2002).

As informações desejadas, após serem identificadas, a próxima etapa geralmente é a busca por similaridade. Os *softwares*, geralmente, usam heurísticas computacionais no processo de busca, podendo-se citar como uns dos mais conhecidos o BLAST (*Basic Local Alignment Search Tools*) (ALTSCHUL et al., 1997) e o FASTA (LIPMAN; PEARSON, 1985).

2.8 Sequência FASTA

O formato FASTA é o tipo de sequência de nucleotídeo e de aminoácidos mais utilizado, este é baseado em texto que representa sequências de nucleotídeos ou sequências de aminoácidos, utilizando os códigos de uma letra. Tem início com uma descrição de uma única linha e, após uma quebra de linha, existem linhas em sequência. A linha de descrição é distinguida com base em dados de sequências de um símbolo de maior (">") (LIPMAN; PEARSON, 1985).

2.9 BLAST (Basic Local Alignment Search Tools)

A ferramenta BLAST (*Basic Local Alignment Search Tools*) é utilizada para comparação de sequências de nucleotídeos e aminoácidos. A pesquisa inicial é feita por sequências de entrada e é comparada com a consulta, utilizando uma matriz de substituição. Acertos de palavras são, então, estendidos em qualquer direção, na tentativa de gerar um alinhamento com uma pontuação a ser avaliada, chamada de *E-value*. O valor de *E-value* corresponde à possibilidade da sequência ser alinhada ao acaso ao fazer a busca (ALTSCHUL et al., 1997).

Análises de similaridade de sequências são os tipos mais fundamentais e frequentemente realizadas sobre os dados no principal banco de dados de

sequências. O NCBI (*National Center for Biotechnology Information*) oferece a “família” BLAST de programas (blast. ncbi. nlm. nih. gov) para detectar similaridades entre uma sequência de consulta e as sequências do banco de dados (ALTSCHUL et al., 1997; ZHANG et al., 1998). No NCBI são encontrados os cinco tipos clássicos de BLAST (nucleotide blast, protein blast, blastx, tblastn e tblasx.) ou, também, por meio de um conjunto de programas autônomos distribuídos por FTP (File Transfer Protocol ou Protocolo de Transferência de Arquivos) (SAYERS et al., 2009).

2.10 Blastn (*Nucleotide Blast*)

Compara uma sequência de nucleotídeos contra um banco de dados de nucleotídeos. O blastn é muito utilizado objetivando a procura de sequências de nucleotídeos que são muito conservados. O blastn encontra sequências similares por meio da quebra das sequências maiores em sequências menores. O programa identifica as correspondências exatas encontradas para as sequências. O *nucleotide blast* pode ser definido como um programa que “estende” as sequências que ele mesmo quebrou, para poder, assim, achar a similaridade com uma sequência alvo, passo a passo e, assim, gerar um alinhamento final com ambas as sequências completas.

Um dos parâmetros importantes que regem a sensibilidade de pesquisas da ferramenta blastn é o comprimento das palavras iniciais, ou tamanho da palavra, como é chamado. Por causa disso, é melhor em encontrar alinhamentos de sequências nucleotídicas de outros organismos relacionados. O tamanho de palavra é ajustável e pode ser reduzido baseado no valor padrão para um mínimo de 7 para aumentar a sensibilidade da pesquisa (NCBI, 2013).

2.11 tBlastn

Utiliza uma sequência de aminoácidos contra um banco de nucleotídeos. A pesquisa tblastn permite comparar uma sequência de proteína para as traduções de seis quadros de um banco de dados de nucleotídeos.

Pode ser uma maneira muito produtiva de encontrar regiões homólogas codificantes de proteínas em sequências de nucleotídeos não anotadas como etiquetas de sequências expressas (ESTs). Assim como registros de projetos genoma (HTG - *high-throughput genomic sequences*), localizados nos bancos de dados EST e HTGs, respectivamente (NCBI, 2013).

2.12 Blastp (*Protein Blast*)

O blastp utiliza como sequência de entrada uma sequência de aminoácidos contra um banco de dados de proteínas. Esse tipo de BLAST é muito utilizado, quando se tem uma proteína e deseja-se saber se existem, em outros organismos, proteínas similares. É utilizado tanto para a identificação de uma sequência de aminoácidos da consulta e para descobrir sequências semelhantes em bases de dados de proteínas.

Como outros programas BLAST, blastp é projetado para encontrar regiões locais de semelhança. Quando a semelhança de sequência abrange toda a sequência, o blastp, também, irá reportar um alinhamento global, que é o resultado de preferência para fins de identificação de proteínas (NCBI, 2013).

2.13 Blastx

Compara uma sequência de nucleotídeos contra um banco de dados de proteínas. O blastx traduz a sequência de consulta em todos os seis quadros de

leitura possíveis e fornece significância estatística combinada para acessos a quadros diferentes, é, particularmente, útil quando o quadro de leitura da sequência de consulta é desconhecido ou contém erros que podem levar a mudanças de quadros ou outros erros de codificação.

Assim, o blastx é, muitas vezes, a primeira análise realizada com uma sequência de nucleótidos recentemente determinada. Esta procura é muito mais sensível do que o blast de nucleotídeos, uma vez que a comparação é feita ao nível da proteína (NCBI, 2013).

2.14 tBlastx

Utiliza como sequência de entrada uma sequência de nucleotídeos contra um banco de nucleotídeos. Compara os seis frames tradicionais de uma sequência de nucleotídeos com os seis frames tradicionais de outra sequência de nucleotídeos do banco de dados.

O tblastx diminui a possibilidade de ambiguidades que podem impedir que determinados quadros de leitura aberta venham a ser detectados. Isto é muito útil na identificação de potenciais proteínas codificadas pelos ESTs em um único quadro de leitura, sendo muito útil para a identificação de novos genes (NCBI, 2013).

2.15 Extração de ácidos nucleicos

A extração de DNA de plantas e de material vegetal é uma etapa importante na análise da estrutura e organização do genoma de plantas. O DNA vegetal, muitas vezes, tem a necessidade de ser preparado, para que se possam produzir amostras puras e com qualidade suficiente para estarem aptas a serem

utilizadas nas mais variadas técnicas de biologia molecular, como, por exemplo, a reação em cadeia da polimerase - PCR (BRASILEIRO; CARNEIRO, 1998).

Para o isolamento de DNA, é fundamental manter a integridade das moléculas, que devem permanecer inalteradas, durante o processo de extração, pois as informações contidas no DNA dependem da sequência intacta. A técnica de extração de DNA genômico de tecidos vegetais com o uso de brometo de cetil-trimetil amônio (CTAB) é a mais amplamente utilizada (DOYLE; DOYLE, 1990; FERREIRA; GRATTAPAGLIA, 1998).

Já, obtenção de um RNA de boa qualidade e em quantidades adequadas, faz-se necessária para o estudo de genes diferencialmente expressos, incluindo transcrição reversa, PCR quantitativo em tempo real (qPCR) e construção de bibliotecas de cDNA (CARDILLO; GIULIETTI; MARCONI, 2006).

Diferentes tecidos demonstram características distintas, na etapa de extração de RNA, pois são encontrados diferentes compostos em cada parte da planta e cada um com uma característica distinta (GEUNA; HARTINGS; SCIENZA, 1998). Adaptações nas metodologias se fazem necessárias, para obtenção de RNA, para cada tipo de tecido utilizado, muitas dessas modificações são encontradas na literatura (GEUNA; HARTINGS; SCIENZA, 1998; SALZMAN et al., 1999).

A utilização de polivinilpirrolidona (PVP) (SALZMAN et al., 1999) e *kits* comerciais como o TRIzol® Reagent (Invitrogen™) e o Concert™ Plant RNA Reagent (Invitrogen™), são comumente utilizados na extração de RNAs totais de tecidos vegetais (INVITROGEN, 2001). A presença de polissacarídeos, polifenóis e uma gama de metabólitos secundários tornam difícil a padronização de protocolo entre os diferentes tecidos para obtenção de RNA (CAMPOS et al., 2010). Outra etapa essencial, para a obtenção de boas amostras de RNA, é a eliminação total de DNA, após a extração, pois apenas uma única cópia de DNA é capaz de gerar resultados não confiáveis (BIRTIC; KRANNER, 2006).

Apesar de existirem muitos protocolos para diferentes espécies e tecidos, tais metodologias para isolamento de ácidos nucleicos sempre se baseiam na lise e na desnaturação das células permitindo a liberação dos ácidos nucleicos totais (AUSUBEL et al., 1998).

2.16 PCR (Reação em Cadeia da Polimerase)

A técnica de PCR (Reação em Cadeia da Polimerase ou *Polymerase Chain Reaction*) é um método desenvolvido por Saiki et al. (1988).

A reação parte de um DNA molde e a técnica consiste em três passos (desnaturação, anelamento e extensão): a fita dupla de DNA é desnaturada sob alta temperatura; em seguida, cada oligonucleotídeo anela com uma das fitas de DNA; os oligonucleotídeos anelados são, então, estendidos pela ação de uma DNA polimerase (ERLICH; GEFLAND; SNINSKY, 1991). Nesse processo, o DNA é desnaturado a 95-97°C, os iniciadores são hibridizados a 30-60°C e, posteriormente, a síntese de DNA é realizada pela adição de desoxirribonucleotídeos trifosfatos (dNTPs) pela DNA-polimerase a 72°C. A repetição dessas etapas, por cerca de 20 a 30 ciclos, permite a amplificação de um segmento de DNA em bilhões de vezes (PASSAGLIA; ZAHA, 2001).

A DNA polimerase é uma enzima termoestável, que suporta altas temperaturas e é utilizada como componente essencial na PCR (MULLIS, 1990). Existem no mercado *kits* comerciais que contêm todos os componentes necessários para a reação, (solução tampão, mistura dos desoxirribonucleotídeos dATP, dCTP, dGTP e dTTP) que devem ser adicionados à reação juntamente com os iniciadores (*primers*) e o DNA molde (alvo) para que a reação se dê de forma completa (VIEIRA, 2013).

2.17 Desenho de *primers* específicos

Primers podem ser definidos como oligonucleotídeos que hibridizam com uma cadeia de DNA molde complementar (*template*) e fornecem uma extremidade 3'-hidroxila para a iniciação da síntese e amplificação de uma sequência de DNA (BORÉM; VIEIRA, 2005).

Algumas características devem ser observadas no momento de desenho dos iniciadores, como o conteúdo de GC superior a 50%, comprimento do produto amplificado (*amplicon*), temperatura de *melting* (T_m), temperatura de pareamento dos *primers* (t_a), *primer dimer* (autocomplementariedade), estabilidade da região 3' e comprimento dos *primers* (LINHART; SHAMIR, 2002).

Esses parâmetros devem seguir algumas recomendações gerais como: tamanho do *amplicon* que deve ser, preferencialmente, entre 80 pb e 150 pb, conteúdo de GC ser entre 45 e 55 %, temperatura de desnaturação entre 95 e 97°C, temperatura de *melting* (T_m) na faixa dos 50°C, temperatura de pareamento é selecionada de acordo com a T_m consensual sendo-lhe diminuída 2 a 4°C, exigência de se ter um par de nucleotídeos do tipo GC na extremidade 3' dos iniciadores, comprimento dos *primers* variando entre 18 e 30 bases (ABD-ELSALAM, 2003; CHEN; JANES, 2002; LI et al., 1997; PELT-VERKUIL; BELKUM; HAYS, 2008; PROMEGA, 2013; SUGGS et al., 1981).

Existem programas disponíveis, para auxiliar o desenho de *primers*, como a CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR *primer design* (ROSE et al., 1998). Algumas páginas *web* gratuitas possibilitam trabalhar *on line*, sem a necessidade de instalação de *softwares* no computador como, por exemplo, os domínios primer 3 plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/>) e Integrated dna Technologies – IDT (<http://www.idtdna.com/site>). Estas ferramentas são

extremamente úteis na verificação dos parâmetros de qualidade necessários no desenho dos *primers*.

2.18 PCR quantitativo (qPCR)

A PCR em tempo real foi descrita pela primeira vez no início da década de 1990 por Higuchi e seus colaboradores, que desenvolveram um sistema ao qual uma câmera de vídeo foi acoplada, de modo a monitorar a PCR durante todos os ciclos. Este mecanismo permitiu-lhes observar o aumento da fluorescência, durante a reação, em virtude da ligação do brometo de etídio às moléculas de DNA de dupla fita recém-sintetizadas (HIGUCHI et al., 1993).

O monitoramento da PCR, em tempo real, revolucionou o processo de quantificação de fragmentos de DNA e RNA. A, então, denominada, PCR em tempo real, permite a quantificação destes ácidos nucleicos de maneira precisa e com maior reprodutibilidade, uma vez que determina valores de quantificação durante a fase exponencial da reação. O acúmulo de produtos da PCR é detectado e monitorado diretamente pela leitura do aumento da fluorescência do fluóforo utilizado, enquanto na PCR tradicional os produtos precisam ser, posteriormente, analisados por eletroforese.

O principal fundamento na utilização deste método é o chamado limiar do ciclo, ou *threshold cycle* (Ct). O Ct é definido como o ciclo da reação de PCR no qual o sinal fluorescente do corante sinalizador atravessa uma linha arbitrária denominada limiar. Para apresentar os dados na forma de Ct, deve-se certificar que o limiar está delimitado na fase exponencial de amplificação, acima do ruído de fundo e abaixo da fase de platô da reação de PCR. O valor numérico do Ct é inversamente proporcional à quantidade inicial do transcrito de interesse na reação, ou seja, quanto menor o valor de Ct, maior a quantidade inicial do transcrito na amostra (SCHMITTGEN; LIVAK, 2008).

A PCR em tempo real possui componentes essenciais para seu funcionamento, entre eles um termociclador com sistema ótico para a excitação da fluorescência e um computador com um *software* para aquisição de dados e análise final da reação, (AMBION, 2013). Duas maneiras distintas são aplicáveis, para que se possa realizar a análise de resultados obtidos em experimentos utilizando a PCR em tempo real, a quantificação absoluta e relativa. Quantificação absoluta representa a quantidade de cópias do transcrito de interesse, enquanto a quantificação relativa demonstra alteração na expressão do gene alvo relativa a algum grupo referência como um controle não tratado ou uma amostra tempo zero. Quantificação absoluta é utilizada quando se precisa conhecer a quantidade de *amplicons*, por exemplo, na quantificação em cálculos de carga viral (SUGDEN; WINTER, 2008).

2.19 Síntese do DNA complementar (cDNA)

Um dos motivos para a não utilização da molécula de RNA é por causa de sua instabilidade a altas temperaturas, existindo, assim, a necessidade da síntese de um cDNA (DNA complementar, ou recombinante). Esse cDNA é utilizado na reação de qPCR, o cDNA é sintetizado numa reação catalisada pela enzima transcriptase reversa e DNA polimerase. Essa técnica foi desenvolvida na década de 1970 em trabalho realizado por Kacian e colaboradores (KANCIAN et al., 1971).

As transcriptases reversas, geralmente, são provenientes de vírus, que usam amostras de RNA mensageiro ou RNA total como molde. Para que a reação de síntese ocorra, são necessários dNTPs (desoxirribonucleotídeos fosfatados), uma solução de $MgCl_2$ (cloreto de magnésio), enzima transcriptase reversa e o RNA mensageiro total (SAMBROOK; RUSSEL, 2002).

2.20 Fluoróforos

Os fluoróforos são moléculas que possuem a capacidade de absorção e emissão de luz em um determinado comprimento de onda. Os sistemas de detecção da PCR em Tempo Real utilizam estas moléculas que proporcionam o acompanhamento da reação ao longo dos ciclos realizados (MARRAS, 2008). Dois tipos distintos de fluoróforos são os mais utilizados na técnica de PCR em tempo real, o SYBR® Green e a TaqMan® (*Applied Biosystems*).

O SYBR® Green, liga-se de maneira inespecífica ao DNA, a detecção é realizada no momento em que o DNA dupla fita é gerado, em razão do aumento da quantidade do fluoróforo entre as cadeias de DNA geradas (MACKAY et al., 2007). Um sinal de fluorescência fraco é produzido, quando não ocorre a ligação das moléculas SYBR® Green ao DNA, não sendo analisado pelo computador. Assim, a reação é monitorada continuamente e um aumento da fluorescência é observado em tempo real. No ciclo seguinte, na etapa de desnaturação do DNA, as moléculas do SYBR® Green são liberadas e há queda no sinal da fluorescência, ocorrendo o monitoramento da quantidade crescente de DNA amplificado em cada ciclo realizado. A utilização deste fluoróforo possui vantagens de boa sensibilidade e fácil manuseio, além de ter baixo custo.

Possui a desvantagem de que a ligação do fluoróforo pode ocorrer em todo DNA fita dupla que surge durante a reação, tais como os dímeros dos iniciadores e sequências indesejadas fazendo com que a concentração final do fragmento seja superestimada (VITZTHUM et al., 1999).

A outra fluorescência muito utilizada é chamada de Taq® DNA Polimerase (*Applied Biosystems*) e tem a característica de possuir um marcador fluorescente comumente chamado de repórter e um capturador de fluorescência denominado “*quencher*”. Enquanto estas duas moléculas estiverem ligadas à sequência linear de nucleotídeos, a fluorescência do repórter será capturada pelo

quencher e não haverá liberação de fluorescência pela sonda e detecção de sinal pelo aparelho (SPANAKOS et al., 2002). Quando a enzima Taq DNA polimerase encontra a sonda, ela a remove pela degradação de sua extremidade 5', liberando o "repórter" para a solução. Após a separação do "repórter" e do "quencher", o sinal fluorescente será, então, detectado pelo aparelho (MARRAS, 2008).

Possui a vantagem de detectar ampliações mais específicas, quando comparada ao SYBR® Green, porém possui a desvantagem de ter elevado custo, quando comparado ao mesmo.

2.21 Histórico espectrometria de massas

Os primórdios da espectrometria de massas remetem a J. J Thomson, que realizou experimentos em 1913, com transmissão de eletricidade por meio de gases e do desvio de raios catódicos por campos eletromagnéticos (DAHL, 1997).

No entanto, em 1918, Arthur Jeffrey Dempster desenvolveu o primeiro espectrômetro de massas capaz de aferir as massas de átomos carregados e, em 1919, Francis W. Aston melhorou a técnica, com uma maior sensibilidade, o que lhe permitiu descobrir 212 isótopos naturais, recebendo o Prêmio Nobel de Química no ano de 1922 pela pesquisa (BUDZIKIEWICZ; GRIGSBY, 2006; DAHL, 1997). Até hoje, esses estudos servem como base para as inovações tecnológicas dos espectrômetros de massas modernos.

A espectrometria de massas apresenta uma evolução constante e tem sido bastante utilizada, principalmente, nos campos da metabolômica e proteômica sendo frequente o desenvolvimento de seus equipamentos, cada vez mais robustos, sensíveis e com maior resolução (HOFFMANN; STROOBANT, 2007).

A espectrometria de massas (MS) é, atualmente, a melhor técnica de detecção para se acoplar à cromatografia, pois o espectrômetro de massas é sensível a pequenas quantidades de analito, fornece informações qualitativas e quantitativas sobre os compostos eluídos e pode distinguir substâncias diferentes com o mesmo tempo de retenção (HOFFMANN; STROOBANT, 2007).

2.22 Espectrometria de massas (MS - *Mass Spectrometry*): definições

A espectrometria de massas (MS) tem como princípio a conversão dos compostos químicos presentes em uma amostra em íons gasosos que sofrem separação com base na razão entre a massa (m) e a carga elétrica (z) (ALBERICI et al., 2010).

É uma técnica que identifica a composição química de diferentes compostos, isolados ou não, pelo estudo das massas de átomos, moléculas ou fragmentos de moléculas, desde que estes estejam eletricamente carregados. Um espectrômetro de massas determina a massa de uma molécula ou átomo medindo a razão massa/carga (m/z) dos íons gerados. A formação dos íons se dá por meio da perda ou do ganho de carga por uma espécie eletricamente neutra. Esses íons são direcionados para dentro do analisador de massas, separados de acordo com suas razões massa carga e detectados. A soma desses processos resulta no espectro de massas, que mostra a intensidade relativa em função da razão m/z para cada íon (HARRIS, 2008).

O estudo da matéria por meio da formação de íons, em fase gasosa e sua caracterização pela massa, carga, estrutura ou propriedades físico-químicas, emprega espectrômetros de massas (MURRAY et al., 2013).

2.23 Componentes do espectrômetro de massas

Um espectrômetro de massas é, basicamente, descrito por um sistema que possui um local para inserção de amostra, e possui componentes principais chamados de fonte de ionização, analisador de massas, detector e um sistema para processar os dados, (figura 1) (KINTER; SHERMAN, 2000; SKOOG; HOLLER; NIEMAN, 2002).

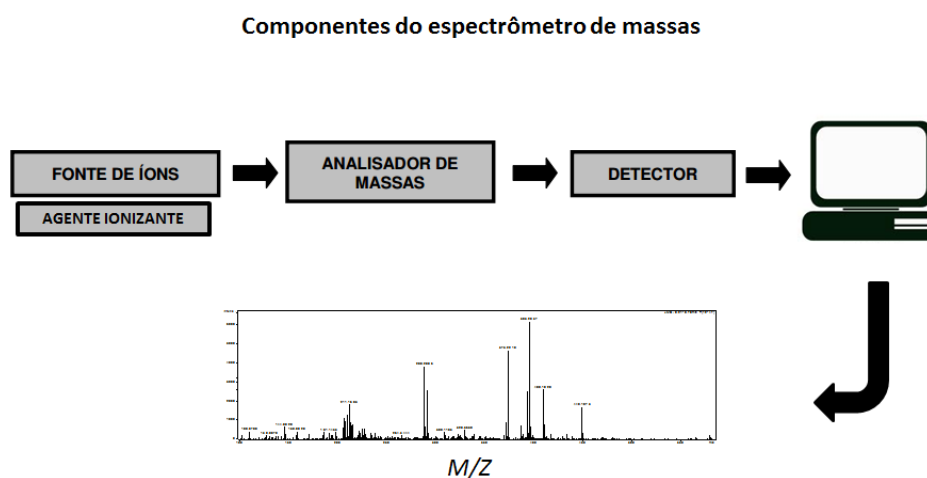


Figura 1 Ilustração de um de um espectrômetro de massas e seu funcionamento

Fonte: Adaptado de Skoog, Holler e Nieman (2002).

Com a ação de um agente ionizante, as amostras são convertidas em íons na fonte, esses (íons positivos ou negativos) são, então, acelerados por meio de um campo elétrico em direção ao analisador de massas.

A função do último componente principal é receber os íons que foram separados pelo analisador, transformando a corrente de íons em sinais elétricos

que são processados, armazenados na memória de um computador e mostrados em uma tela (MORAES; LAGO, 2003).

2.24 Fontes de ionização

A fonte de ionização pode ser definida como um componente do espectrômetro de massas capaz de gerar os íons considerando os analitos, e realiza sua transferência para a fase gasosa. Em um ambiente a vácuo, os íons são levados para o interior do espectrômetro e conduzidos até o detector, a quantidade de energia necessária para que este processo ocorra é chamada de potencial de ionização (WATSON; SPARKMAN, 2007).

2.25 Ionização por eletrospray (ESI – *Electrospray Ionization*)

Os primeiros relatos sobre a nebulização por eletrospray datam de 1917, descrito por Zeleny (NIESSEN, 1998). No entanto, apenas na década de 80 foi demonstrada a aplicabilidade da fonte de eletrospray (*Electrospray Ionization* – ESI) como um método de ionização branda.

A técnica pode ser definida como um processo no qual espécies ionizadas na fase gasosa são produzidas baseadas em uma solução pela formação e dessolvatação de minúsculas gotas altamente carregadas, que são resultantes da aplicação de uma diferença de potencial da ordem de kV entre um capilar e um contraeletrodo, sob pressão atmosférica (YAMASHITA; FENN, 1984). Então, uma fonte de alta tensão (1000 a 7000 V) é aplicada à solução contendo eletrólitos. A vazão é geralmente inferior a $10 \mu\text{L min}^{-1}$. A voltagem aplicada à solução pode ser positiva ou negativa, dependendo das características da amostra (MORAES; LAGO, 2003).

Em geral, a ionização por ESI é um processo que possui três fases: nebulização da solução de amostra em gotículas produzidas pela aplicação direta de voltagem no capilar, liberação dos íons por gotículas e transporte dos íons da região de pressão atmosférica da fonte de ionização para a região de alto vácuo do analisador de massas (COLE, 1997).

Embora os mecanismos que estão envolvidos nos processos de ionização por ESI não tenham sido completamente elucidados, existem dois principais mecanismos propostos. Um mecanismo de ionização é baseado no modelo do resíduo carregado (*charged-residue model* - CRM). Um modelo teórico para a produção de macromoléculas altamente carregadas na ionização por eletrospray, onde o excesso de carga gerado nas gotículas é transferido para estas moléculas e retido por elas mesmas após a evaporação do solvente.

O outro mecanismo é baseado no modelo de evaporação de íons (*ion evaporation model* - IEM), no qual os íons são diretamente ejetados da superfície da gotícula. Acredita-se que este mecanismo seja predominante para moléculas de baixa massa molar.

Nestes dois mecanismos, a ionização tem início pela formação do aerosol, produzido por um processo que envolve a formação do cone de Taylor, com base na ponta do capilar (MURRAY et al., 2013; YAMASHITA; FENN, 1984).

A ionização por ESI possui uma alta sensibilidade, permitindo que os mais variados tipos de compostos possam ser analisados, desde que sejam não-voláteis, podendo ser de média ou alta polaridade e seu peso molecular inferior a 200.000 Da. Atualmente, é uma das técnicas de ionização mais utilizadas em razão da ampla faixa de compostos detectados.

Sua utilização, geralmente, busca identificar substâncias usadas no combate e prevenção de pragas agrícolas, presentes na água e alimentos, identificação de produtos de degradação de relevância ambiental, mapeamento

proteômico, *screening* de drogas e adulteração em bebidas (MENDHAM et al., 2002), dentre outras diversas aplicações.

2.26 Analisadores de massas (*Mass Analyzers*)

O processo que ocorre, após a etapa de formação dos íons, é a separação dos mesmos e esta etapa é realizada pelo analisador de massas, que separa esses íons de acordo com a razão m/z (VESSECCHI et al., 2011). Dois dos principais tipos de analisadores estão descritos a seguir.

2.27 Quadrupolo (Q – *Quadrupole*)

Em função de seu baixo custo, o espectrômetro de massas com quadrupolo é um dos mais utilizados. É formado por quatro hastes metálicas paralelas às quais recebem um potencial elétrico constante e um potencial oscilante de radiofrequência (MENDHAM et al., 2002).

O quadrupolo é constituído por quatro barras cilíndricas paralelas duas a duas. Cada par de barras é conectado eletricamente estabelecendo, assim, um campo quadrupolar bidimensional no plano x-y. Os íons viajam na direção z. Enquanto viajam nesta direção, eles, também, oscilam no plano x-y, em virtude do potencial aplicado às barras. Portanto, em condições elétricas adequadas, os íons com uma única razão m/z ou que estão dentro de uma faixa de m/z selecionada, terão uma trajetória estável durante todo o caminho do quadrupolo (KINTER; SHERMAN, 2000; SKOOG; HOLLER; NIEMAN, 2002).

Campos elétricos oscilantes, formados por quatro barras metálicas, são utilizados para estabilizar ou desestabilizar seletivamente os íons, durante sua passagem pelo centro do quadrupolo, de acordo com seus valores de m/z . O quadrupolo funciona como um filtro. Isso permite que os íons de diferentes

valores de m/z cheguem com tempos diferentes ao detector e, desta forma, possam ser diferenciados (SCHALLEY, 2000).

Existem variações com mais de um quadrupolo, chamadas de espectrômetro de massas sequencial, composto por dois quadrupolos de transmissão montados em série, entre os quais há um quadrupolo Rf-exclusivo que funciona como célula de colisão (VESSECCHI et al., 2011). Nesta configuração, é possível realizar experimentos de MS/MS, os quais são utilizados para elucidação estrutural de compostos químicos.

2.28 Tempo de Voo (TOF – *Time of Flight*)

Os íons são separados em função do tempo de voo que cada um (com determinada razão m/z) leva para percorrer sob vácuo e a distância entre a amostra e o detector. Em geral, esse percurso é constituído por uma ou duas regiões de aceleração curta e outra região, longa, livre de campo elétrico. Este tipo de analisador considera o fato de que íons de massas diferentes, ao serem ionizados simultaneamente e, então, submetidos ao mesmo campo elétrico acelerador, levam tempos diferentes para chegar ao detector (COLLADO et al., 2004).

O instrumento separa íons por m/z , em uma região livre de campo, após aceleração por uma voltagem de aceleração fixa. Íons com a mesma energia translacional inicial e diferente m/z levam tempos diferentes para atravessar uma dada distância em uma região livre de campo (VESSECCHI et al., 2011). Silverstein, Webster e Kiemle (2005) descrevem o detector tipo TOF de maneira que os íons são acelerados por um potencial (V) e passam por um tubo até um detector. Se todos os íons que entram no tubo tiverem a mesma energia, então, íons de massas diferentes terão velocidades diferentes. No entanto, esse equipamento necessita produzir os íons em uma determinada localização e

momento conhecido, limitando os espectrômetros de massas por tempo de voo às técnicas de ionização pulsada (HARRIS, 2008; MENDHAM et al., 2002).

Os analisadores tipo TOF têm sido amplamente utilizados, pois são espectrômetros de massas de alta resolução que apresentam menor custo.

2.29 Detectores

Os detectores consistem na parte final dos espectrômetros de massas, possuem a função de detectar e amplificar o sinal da corrente de íons que são oriundos do analisador e transferir o sinal para o sistema de processamento de dados. Os detectores medem as correntes dos feixes de íons anteriormente separados pelo analisador (SKOOG; HOLLER; NIEMAN, 2002). Existem vários tipos de detectores atualmente e a escolha depende da concepção do instrumento e das aplicações analíticas que serão executadas (HOFFMANN; STROOBANT, 2007). O multiplicador de elétrons é um dos detectores mais simples, completos e que apresenta grande sensibilidade, utilizado até hoje, uma breve descrição deste detector é feita a seguir.

2.30 Multiplicadora de elétrons (EM – *Electron Multiplier*)

Um dos tipos de detectores mais utilizados atualmente em espectrometria de massas é a multiplicadora de elétrons (*Electron Multiplier* - EM). Neste detector, os íons são acelerados em altas velocidades, a fim de aumentar a eficiência de detecção. Isto é alcançado, segurando um eletrodo chamado de dínodo que realiza um alto potencial de conversão que varia de ± 3 a ± 30 kV. Um íon positivo ou negativo, ao chocar-se com dínodo de conversão provoca a emissão de várias partículas secundárias. Estas partículas secundárias podem incluir íons positivos, íons negativos, elétrons e cargas neutras. Estas

partículas secundárias são convertidas em elétrons no primeiro dínodo e, então, são amplificadas por efeito de cascata na multiplicadora de elétrons produzindo, assim, uma corrente.

Existem dois tipos de multiplicadoras de elétrons, do tipo discreto (*Discrete-Dynode*), constituído por 16 a 20 dínodos, geralmente, constituídos de cobre e berílio e do tipo contínuo (*Continuous-Dynode*), no qual a superfície atua como um eletrodo contínuo. Em ambos os detectores, os íons, provenientes do analisador, chocam-se com a superfície interna do dínodo e são refletidos, em um processo que se repete várias vezes (FINNIGAN, 1994).

2.31 Espectrometria de massas sequencial – MS/MS (*Tandem Mass Spectrometry* – MS/MS)

Espectrometria de massas sequencial, ou MS/MS, é um método, que, em geral, envolve pelo menos duas fases de análise da massa, ou em conjunto, com um processo de dissociação ou de uma reação química que provoca uma alteração na massa ou na carga de um íon (HOFFMANN; STROOBANT, 2007).

Na espectrometria de massas em tandem, geralmente, o primeiro analisador é usado para isolar um íon precursor, que, então, passa, espontaneamente ou por alguma ativação de fragmentação, para produzir íons produtos e fragmentos neutros, por um segundo analisador, o qual detecta os produtos dos íons (MARSHALL; HENDRICKSON; SHI, 2002).

Basicamente, um espectrômetro de massas em tandem pode ser produzido de duas maneiras: 1) pelo acoplamento de dois instrumentos fisicamente distintos, ou 2) executando uma sequência apropriada de acontecimentos num único dispositivo de armazenamento de íons. Assim, existem duas categorias principais de instrumentos que permitem experimentos

de espectrometria de massas em tandem: espectrômetros de massa em tandem no espaço ou no tempo.

Instrumentos comuns têm dois analisadores de massas, permitindo a realização de experimentos MS/MS. Um instrumento usado frequentemente neste tipo usa quadrupolos como analisadores. A configuração QqQ indica um instrumento com três quadrupolos em que o segundo, indicado por uma letra minúscula q, é a região de reação e da fragmentação do íon. Instrumentos TOF com um *reflecton*, ou uma combinação de um quadrupolo e TOF em um instrumento QTOF, também, são utilizados.

Para obter maior ordem de fragmentação são realizados experimentos MSⁿ, os quais requerem analisadores tipo armadilhas de íons (*Ion Trap*). Estes equipamentos são os únicos capazes de realizarem fragmentação sucessiva dos íons produtos, não havendo um limite no número de fragmentação, sendo de grande importância na elucidação estrutural de compostos químicos.

Analisadores em sucessão podem ser dispostos em qualquer número, porém o máximo que pode ser utilizado na prática são três a cinco analisadores, no caso dos instrumentos de feixe. Podem ser utilizados analisadores do tipo quadrupolo, TOF, *Ion Trap*, *Orbitrap* e FTICR (*Fourier Transform Ion Cyclotron Resonance*).

2.32 Métodos cromatográficos acoplados à espectrometria de massas

Entre os métodos de análise, a cromatografia ocupa um lugar de destaque em virtude da facilidade com que efetua a separação, identificação e quantificação das espécies químicas, por si mesmas, ou em conjunto com outras técnicas instrumentais de análise como a espectrofotometria ou a espectrometria de massas.

A fim de analisar misturas complexas, métodos cromatográficos podem ser acoplados em um espectrômetro de massas, tais como cromatografia gasosa (*Gas Chromatography*- GC), cromatografia líquida (*Liquid Chromatography* - LC) ou eletroforese capilar (*Capillary Electrophoresis* - CE). Os produtos separados são introduzidos no espectrômetro, tanto no estado gasoso por GC-MS, quanto em solução por LC-MS e CE-MS. Isto pode ocorrer de dois modos: o composto por eluição é coletado e analisados *off-line*, ou o cromatógrafo é ligado diretamente ao espectrômetro de massas e os espectros são adquiridos enquanto os compostos da mistura são eluídos (KOSTIAINEN et al., 2003).

O detector ideal para ser acoplado aos métodos de separação dos compostos químicos deve possuir algumas características como: nenhuma alteração da resolução cromatográfica, ter a maior sensibilidade possível, ser universal, fornecer o máximo de informação estrutural, ser seletivo, emitir um sinal proporcional à concentração da amostra, boa relação custo/desempenho e não ser prejudicial para o produto (HOFFMANN; STROOBANT, 2007). O espectrômetro de massas possui estas características e é, portanto, utilizado acoplado aos métodos cromatográficos com grande êxito em diferentes aplicações.

2.33 Cromatografia Líquida de Ultra Alta Eficiência (*Ultra High Performance Liquid Chromatography* - UHPLC)

A Cromatografia líquida de ultra eficiência (*Ultra Performance Liquid Chromatography* - UHPLC) é definida como uma técnica em que uma mistura de compostos é separada em componentes individuais, por meio de um líquido (fase móvel) em alta pressão que passa por uma fase estacionária (VESSECCI et al., 2011). É, também, chamada de cromatografia de alta velocidade, alta pressão, alto desempenho, alta resolução e alta eficiência, porém, a

nomenclatura mais aceita é CLAE (Cromatografia Líquida de Alta Eficiência) em português e em inglês HPLC (COLLINS; BRAGA; BONATO, 2006).

Geralmente os instrumentos são automatizados possuindo a capacidade de realizar separações e análises quantitativas de uma grande variedade de amostras em pouco tempo e com alta precisão (ARDREY, 2003).

O HPLC possui certas vantagens quando comparada à cromatografia gasosa, principalmente, no caso de separações de compostos polares e termolábeis, os quais, muitas vezes, são obtidos mais facilmente na cromatografia líquida que por cromatografia gasosa, pois existem duas fases cromatográficas na LC, para a interação seletiva com moléculas da amostra, contra apenas uma na GC. Outra vantagem é a maior variedade de colunas disponíveis (fases estacionárias) para HPLC e a possibilidade de utilização de temperaturas de separação mais (SNYDER; KIRKLAND, 1979).

Dois instrumentos são os mais utilizados, o HPLC (*High Performance Liquid Chromatography*) e o UPLC (*Ultra Performance Liquid Chromatography*) sendo o UPLC uma variante do HPLC. O UPLC pode trabalhar com partículas de colunas menores e pressões mais elevadas quando comparado com o HPLC, fazendo com que a sensibilidade seja maior se acoplado a um espectrômetro de massas (SWARTZ, 2005), além de promover análises mais rápidas e com melhor resolução na separação cromatográfica.

2.34 Cromatografia líquida acoplada à espectrometria de massas (*Liquid Chromatography/Mass Spectrometry – LC-MS*)

A cromatografia líquida (LC) normalmente é usada para a separação de compostos polares e termolábeis e, portanto, não são adequados para a separação por cromatografia gasosa (ARDREY, 2003; NIESSEN, 1998).

Em LC-MS, é necessário o acoplamento de dois instrumentos (LC e MS). A fase líquida, proveniente do LC, é inserida no MS, utilizando, normalmente, as fontes de ionização por ESI, ionização química a pressão atmosférica (APCI - *atmospheric pressure chemical ionization*) e fotoionização a pressão atmosférica (APPI - *atmospheric pressure photo ionization*), as quais podem tolerar as taxas de fluxo de cerca de 1 mL. min⁻¹ (HUANG et al., 1990).

A definição mais recente da técnica diz que LC-MS permite que uma mistura de compostos pode ser separada em componentes individuais e menos complexos, sendo o efluente da separação ionizado na fonte de íons do espectrômetro de massas (VESSECCHI et al., 2011).

2.35 Aplicações da espectrometria de massas na determinação estrutural de biomoléculas

A determinação do peso molecular é uma das primeiras medições utilizadas para caracterizar biopolímeros. Até o final da década de 1970, as únicas técnicas que forneciam essas informações foram eletroforéticas, métodos cromatográficos ou ultracentrifugação. Os resultados não eram muito precisos (10-100% de erro relativo, em média), em razão da dependência de algumas características como a conformação, o raio de Stokes e a hidrofobicidade de cada molécula analisada (HOFFMANN; STROOBANT, 2007).

O desenvolvimento dos métodos de ionização de dessorção, com base na emissão de íons pré-existentes, considerando um líquido ou uma superfície sólida, tal como a dessorção de plasma (*Plasma Desorption* - PD) (HARKANSSON et al., 1982), de bombardeamento atômico rápido (*Fast Atom Bombardment* - FAB) (BARBER et al., 1981) ou dessorção por laser (*Laser Desorption* - LD) (COTTER, 1984), permitiu um primeiro avanço para a utilização da espectrometria de massas no campo da análise de biomoléculas.

Desde então, o problema deixou de ser a produção de íons, mas, sim, de analisar tais massas isoladamente carregadas de íons, as quais são tecnicamente difíceis de serem detectadas com boa sensibilidade assim como de serem analisadas com uma boa resolução (HOFFMANN; STROOBANT, 2007).

No início da década de 1990, dois novos métodos de ionização, ESI (FENN et al., 1989) e MALDI, acoplados a um analisador TOF (KARAS; HILLENKAMP, 1988) foram desenvolvidos e continuaram a revolucionar o papel da espectrometria de massas na pesquisa biológica. Estes dois métodos de ionização são capazes de ionizar e transferir, para dentro do espectrômetro de massas, moléculas polares e macromoléculas, o que não era possível até então pelas técnicas de ionização utilizadas, como a ionização por elétrons (*electron ionization* – EI) e ionização química (chemical ionization – CI). Com o desenvolvimento destes métodos, foi possível a análise com alta precisão de biomoléculas de peso molecular muito elevado.

A análise por espectrometria de massas de diferentes classes de biomoléculas abrange os mais diversos compostos como peptídeos, proteínas, ácidos nucleicos, oligossacarídeos e lipídios, produtos do metabolismo secundário, entre outras aplicações (PENG; GYGI, 2001). Com isso, atualmente, a espectrometria de massas tornou-se uma das técnicas analíticas mais utilizadas nas ciências da vida (BURLINGAME, 2005).

2.36 Metabolômica

Os primórdios dos estudos relacionados a metabólitos foram iniciados há mais de 30 anos (HORNING; HORNING, 1971), no entanto, o termo metabolômica foi introduzido em 2002 por FIEHN. A expressão metabolômica é utilizada em paralelo a outros termos já existentes tais como: genômica, transcriptômica e proteômica, termo o qual é utilizado para definir uma técnica

que permite identificar e quantificar os mais variados tipos de metabólitos existentes em diversos organismos (FIEHN, 2002).

O metaboloma é uma das peças do sistema biológico o qual abrange todos os compostos produzidos em um sistema vivo. A composição desse sistema inclui o genoma, transcriptoma, proteoma e metaboloma. O estudo do metaboloma pode ser definido como o conjunto e a busca pelo entendimento completo de pequenas moléculas que estão envolvidas em reações metabólicas gerais e que são biossintetizados por uma célula, tecido ou organismo (BEECHER, 2003).

Dentro do campo da metabolômica existem dois tipos de análises tradicionais de estudos, a “*Targeted*” e “*Untargeted*”, ou “*Non Targeted*”. Quando se refere ao tipo *Targeted*, a análise é conduzida para moléculas alvos em que já se possui um conhecimento prévio, ao passo que na *Untargeted* é realizado um perfil (“*profiling*”) de todos os metabólitos presentes no sistema. Existindo, ainda, uma classificação quantitativa e qualitativa na *targeted* e *untargeted*, respectivamente (LIN et al., 2011).

2.37 Espectrometria de massas na metabolômica

As determinações do genoma e do transcriptoma não dependem da espectrometria de massas, pois outras técnicas permitem que as sequências de genes sejam alcançadas de uma forma mais eficiente. Já na proteômica, a espectrometria de massas possui um papel fundamental, pois nenhuma outra técnica, até o momento, é mais eficiente, para determinar a sequência de uma proteína tão rapidamente, utilizando-se quantidades muito baixas de amostra (HOFFMANN; STROOBANT, 2007).

Mas o conhecimento do genoma, transcriptoma e proteoma não revela o fenótipo de um sistema vivo, pois é difícil ou até mesmo impossível estabelecer

uma ligação direta entre a proteína e sua atividade enzimática realizada, e, assim, os metabólitos produzidos por ela. Muitas vezes, apesar de se possuir as sequências de determinadas proteínas, suas atividades conhecidas são apenas parciais. Para compreender a célula, é essencial saber a expressão ocasionada e os metabólitos produzidos pelas proteínas (HOFFMANN; STROOBANT, 2007).

No entanto, a elucidação do metaboloma é, particularmente, difícil em decorrência da natureza química de diversas moléculas pequenas. Para se obter sucesso na técnica, é necessário ter um manuseio eficiente da amostra e boas técnicas analíticas existe grande possibilidade de ocorrerem perdas de compostos nessa etapa. A espectrometria de massas e ressonância magnética nuclear (*Nuclear Magnetic Resonance* - NMR) são as técnicas mais amplamente utilizadas para a análise do metaboloma. A NMR pode medir, simultaneamente, todos os tipos de moléculas pequenas do metaboloma, além disso, a amostra pode ser recuperada para outras análises. No entanto, a NMR, também, possui limitações significativas de sensibilidade, por isso, apenas é aplicada a compostos de altas concentrações e não pode ser usada para analisar os compostos que estão em concentrações mais baixas e que geralmente, são os mais interessantes (HOFFMANN; STROOBANT, 2007).

Em comparação com a NMR, a espectrometria de massas é mais sensível e, portanto, pode ser usada para compostos em baixa concentração. Embora seja facilmente possível medir picomoles de compostos, os limites de detecção nos níveis atômico podem ser alcançados. A MS, também, permite identificar compostos, por meio da elucidação da sua estrutura química, por MS/MS e determinar suas massas exatas. A MS é uma técnica importante para a análise do metaboloma tendo como suporte outra ferramenta chave: os *softwares* específicos para processar e analisar os dados complexos obtidos (HOFFMANN; STROOBANT, 2007).

Esta técnica pode ser usada de maneira independente, neste caso, a amostra pode ser inserida por infusão direta no espectrômetro de massas (DIMS – *Direct Infusion Mass Spectrometry*) ou com separação prévia. A espectrometria de massas permite analisar cerca de 10.000 compostos em uma amostra (HUGHEY; RODGERS; MARSHALL, 2002). Geralmente é utilizada para identificar e quantificar compostos do metaboloma, após uma separação prévia, por meio de cromatografia sendo muitas vezes utilizado GC-MS ou LC-MS.

GC-MS tem sido amplamente usado, pois é um método poderoso que apresenta resolução cromatográfica muito elevada. No entanto, alguns compostos polares e com elevada massa molecular não podem ser analisados por esta técnica, visto que este método de separação é limitado a moléculas que sejam voláteis ou que possam se tornar voláteis por derivatização e termolábeis. Em consequência das limitações da técnica GC-MS, a LC-MS tornou-se uma técnica mais popular para análise metabolômica. Uma gama muito mais ampla de analitos pode ser analisada por LC-MS. Os métodos de ionização mais utilizados neste campo são ionizações por pressão atmosférica, principalmente, ESI e APCI (HOFFMANN; STROOBANT, 2007).

2.38 Metabólitos secundários

O Metabolismo pode ser definido como um conjunto de reações que estão ocorrendo continuamente dentro da célula de um organismo. Em primeiro lugar, ocorre o aproveitamento de nutrientes para satisfazer as exigências fundamentais da célula. Os compostos químicos formados, degradados ou transformados são denominados de metabólitos. São divididos em duas classes, primários que são aqueles considerados imprescindíveis para o desenvolvimento de um organismo. Os metabólitos secundários, produtos secundários ou produtos

naturais, são os compostos que não são envolvidos de maneira direta em processos como crescimento, desenvolvimento e reprodução dos organismos. Muitos deles distribuídos entre um número muito limitado de espécies dentro do reino vegetal (TAIZ; ZEIGER, 2006).

Os metabólitos secundários nas plantas podem ser divididos em três grandes grupos: terpenos, compostos fenólicos e alcaloides, sendo os principais componentes desses grupos os flavonoides, alcaloides, cumarinas, agliconas, antraquinônicas, triterpenos, saponinas, polifenóis e taninos (TAIZ; ZEIGER, 2006).

Produtos do metabolismo secundário possuem ação antioxidante, não somente pela sua habilidade em doar hidrogênio ou elétrons, mas também em virtude de seus radicais intermediários estáveis (BRAND-WILLIAMS; CUVELIER; BERSET, 1995).

2.39 Terpenos

Terpenos fazem parte do grupo mais antigo de produtos de pequenas moléculas sintetizado por plantas e, provavelmente, o grupo mais difundido de produtos naturais (AUGUST, 1981). São compostos que ocorrem em todas as plantas e compreendem uma classe de metabólitos secundários com uma grande variedade estrutural (RAVEN; EVERT; EICHHORN, 2001).

Os terpenos são formados pela fusão de unidades isoprênicas de cinco carbonos; quando submetidos a altas temperaturas, podem se decompor em isoprenos, podendo referir-se, ocasionalmente, a todos os terpenos como isoprenoides (TAIZ; ZEIGER, 2006).

2.40 Compostos fenólicos

Os compostos fenólicos são substâncias que são encontradas em grandes quantidades em vegetais e, embora em baixas quantidades, também, em microrganismos (COSTA; CUNHA, 2000). Derivam de ácidos benzoicos e de ácidos cinâmicos, cumarinas, flavonoides, de polímeros (taninos e ligninas) ou são de distribuição restrita como fenóis simples, antraquinonas, xantonas, naftoquinonas (SIMÕES et al., 1999).

Esta classe de compostos apresenta uma grande diversidade e divide-se em flavonoides (polifenóis) e não-flavonoides (fenóis simples ou ácidos) (RICE-EVANS; MILLER; PAGANGA, 1996).

2.41 Flavonoides

Flavonoides são encontrados em todas as partes das plantas. Na maioria dos casos, os flavonoides estão presentes como glicosídeo em vacúolos de flores, folhas, caules ou raízes, as unidades mais conhecidas são as flavonas e isoflavonas. Esses compostos compreendem os pigmentos coloridos de flores, mas também atuam como inibidores de enzimas, como um sistema de defesa contra a exposição à radiação ultravioleta e os insetos, e como agentes de metais que são nocivos para as plantas quelantes. Também, estão envolvidos em fotossensibilização e transferência de energia, morfogênese e determinação do sexo, fotossíntese e regulação de hormônios de crescimento de plantas (DI CARLO et al., 1999).

O maior interesse, contudo, é em suas ações fisiológicas e os seus efeitos benéficos sobre a saúde humana. A propriedade mais bem documentada de quase todos os grupos de flavonoides é a sua capacidade de atuar como antioxidantes podendo interferir com, pelo menos, três diferentes sistemas de

produção de radicais livres (NIJVELDT et al., 2001). Os flavonoides são componentes naturais de plantas e possuem efeitos protetores contra doenças humanas, como câncer e doenças cardiovasculares (HOLLMAN, 2001).

2.42 Alcaloides

Dentro dos metabólitos secundários encontrados em vegetais, existem compostos que possuem em sua estrutura o nitrogênio. Incluindo alguns compostos que agem na defesa das plantas contra os mais variados tipos de injúrias, físicas e químicas, estes compostos são conhecidos como alcaloides. Alcaloides típicos são derivados de fontes vegetais, contêm um ou mais átomos de nitrogênio e possuem, geralmente, uma ação fisiológica no homem ou outros animais. Outros alcaloides, que não correspondam com a definição geral, são aqueles compostos sintéticos não encontrados em plantas (EVANS, 2009).

Os alcaloides são conhecidos pelos seus efeitos farmacológicos em animais vertebrados. Possuem átomos de C e N, em seu anel heterocíclico, sendo solúveis em água (MITHEN et al., 2000). As dicotiledôneas são as espécies onde essas substâncias são mais encontradas (LUCA; LAFLAMME, 2001) podendo estar presentes em diferentes partes do vegetal (SOTTOMAYOR et al., 2004) e em representantes de diversas famílias (LORENCE; NESSLER, 2004).

REFERÊNCIAS

ABD-ELSALAM, K. A. Bioinformatic tools and guideline for PCR primer design. **African Journal of Biotechnology**, Nairobi, v. 2, n. 5, p. 91-95, May 2003.

ABDELNUR, P. V. **Metabolômica e espectrometria de massas**. Brasília: EMBRAPA Agroenergia, 2011. 4p.

ALBERICI, R. M. et al. Ambientmass spectrometry: bringing MS into the “real world”. **Analytical and Bioanalytical Chemistry**, Heidelberg, v. 398, n. 1, p. 265-294, Sept. 2010.

ALTSCHUL, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, Oxford, v. 25, p. 3389-3402, 1997.

AMBION. **Real-Time PCR (qPCR) guide**. Disponível em: <<http://www.ambion.com/techlib/tn/81/813.html>>. Acesso em: 20 nov. 2013.

AMORIM, L. et al. (Ed.). **Manual de fitopatologia: doenças das plantas cultivadas**. 3. ed. São Paulo: Agronômica Ceres, 1997. v. 2, 757 p.

ARDREY, R. E. **Liquidchromatography-mass spectrometry: an introduction**. Chichester: J. Wiley, 2003. 276 p.

AUGUST, B. **Organic chemistry**. New York: Blackie, 1891. 1490 p.

AUSUBEL, F. M. et al. (Ed.). **Current protocols in molecular biology**. New York: J. Wiley, 1998. 600 p.

BARBER, M. et al. Fast atom bombardment of solids (F. A. B): a new ion source for mass spectrometry. **Journal of the Chemical Society, Chemical Communications**, London, v. 7, p. 325-327, 1981.

BARCELOS, E. **Características genético-ecológicas de populações naturais de caiaué (*Elaeis oleifera* (H. B. K.) Cortés na Amazônia brasileira**. 1986. 108 f. Dissertação (Mestrado em Ecologia) - Instituto Nacional de Pesquisa do Amazonas, Manaus, 1986.

BARCELOS, E. et al. **A cultura do dendê**. Brasília: EMBRAPA-SPI, 1995. 68 p. (Coleção Plantar, 32).

BASTOS, T. X. Aspectos agroclimáticos do dendezeiro na Amazônia Oriental. In: VIÉGAS, I. J. M.; MÜLLER, A. A. (Ed.). **A cultura do dendezeiro na Amazônia Brasileira**. Belém: EMBRAPA Amazônia Oriental, 2001. p. 47-59.

BEECHER, C. **Metabolic profiling: its role in biomarker discovery and gene function analysis**. Dordrecht: Kluwer Academic, 2003. 352 p.

BILLOTTE, N. et al. Identification of a candidate marker for shell presence in the oil palm using amplified fragment length polymorphism (AFLP) by bulked segregant analysis (BSA). In: PLANT AND ANIMAL GENOMES CONFERENCE, 7., 1999, San Diego. **Proceedings...** San Diego: PAGC, 1999. p. 187.

BIRTIC, S.; KRANNER, I. Isolation of high-quality from polyphenol, polysaccharide and lipid-rich seeds. **Phytochemical Analysis**, Kew, v. 17, n. 3, p. 144-148, 2006.

BORÉM, A.; VIEIRA, M. **Glossário de biotecnologia**. Viçosa, MG: UFV, 2005. 183 p.

BRAND-WILLIAMS, W.; CUVELIER, M. E.; BERSET, C. Use of a free radical method to evaluate antioxidant activity. **Lebensmittel-Wissenschaft Technologie**, London, v. 28, n. 1, p. 25-30, July 1995.

BRASILEIRO, A. C. M.; CARNEIRO, V. T. C. (Ed.). **Manual de transformação genética de plantas**. Brasília: EMBRAPA-SPI/EMBRAPA-Cenargen, 1998. 309 p.

BUDZIKIEWICZ, H.; GRIGSBY, R. D. Massspectrometry and isotopes: a century of research and discussion. **Mass Spectrometry Reviews**, New York, v. 25, n. 1, p. 146-157, Jan. /Feb. 2006.

BURLINGAME, A. L. **Methods in enzymology, biological mass spectrometry**. New York: Academic, 2005. 478 p.

CAMPOS, N. A. et al. Otimização de protocolos de extração de RNA de raiz de milho (*Zea mays*) visando estudos moleculares de alta sensibilidade. In: CONGRESSO NACIONAL DE MILHO E SORGO, 28., 2010, Goiânia. **Anais**. . . Goiânia: Associação Brasileira de Milho e Sorgo, 2010. 1 CD-ROM.

CARDILLO, A. B.; GIULIETTI, A. M.; MARCONI, P. L. Analysis and sequencing of h6hmRNA, last enzyme in the tropane alkaloids pathway from anthers and hairy root cultures of *Brugmansia candida* (*Solanaceae*). **Electronic Journal of Biotechnology**, Cambridge, v. 9, n. 3, p. 196-198, 2006.

CHEN, B. Y.; JANES, H. W. **PCR cloning protocols**. 2nd ed. Totowa: Humana, 2002. 439 p.

CLOTE, P.; BACKOFEN, R. **Computational molecular biology: an introduction**. West Sussex: LTD, 2000. 286p.

COLE, R. B. **Electrospray ionization mass spectrometry: fundamentals instrumentation and applications**. New York: J. Wiley, 1997. 300 p.

COLLADO, V. M. et al. Analysis of Ion dynamics and peak shapes for delayed extraction time-of-flight mass spectrometers. **Review of Scientific Instruments**, Rochester, v. 75, p. 2163-2170, June 2004.

COLLINS, C. H.; BRAGA, G. L.; BONATO, P. S. **Fundamentos de cromatografia**. Campinas: UNICAMP, 2006. 452 p.

CONCEIÇÃO, H. E. O.; MULLER, A. A. Botânica e morfologia do dendezeiro. In: VIÉGAS, I. J. M.; MULLER, A. A. (Ed.). **A cultura do dendezeiro na Amazônia brasileira**. Belém: EMBRAPA Amazônia Oriental, 2000. p. 31-44.

CORLEY, R. H. V.; BARETT, J. H.; JONES, L. H. Vegetative propagation oil palm via tissue culture. In: MALAYSIAN INTERNATIONAL AGRICULTURE OIL PALM CONFERENCE, 38., 1976, Kuala Lumpur. **Proceedings**. . . Kuala Lumpur: Malaysian International Agriculture Oil Palm, 1976. p. 1-7.

CORLEY, R. H. V.; TINKER, P. B. **The oil palm**. 4thed. Oxford: Blackwell Science, 2003. 562 p.

COSTA, A. F.; CUNHA, A. P. da. **Farmacognosia**. 3. ed. Lisboa: Calouste Gulbenkian, 2000. 1032 p.

COTTER, R. J. Laser desorption time-of-flight mass-spectrometry of high mass molecules. **Analytical Chemistry**, Washington, v. 56, p. 1662-1667, 1984.

CUNHA, R. N. V. da et al. Domesticação e melhoramento do caiaué. In: BORÉM, A.; LOPES, M. T. G.; CLEMENT, C. R. (Ed.). **Domesticação e melhoramento: espécies amazônicas**. Viçosa, MG: UFV, 2009. p. 275-296.

CUNHA, R. N. V. da et al. Domestication and breeding of the American oil palm. In: BORÉM, A. et al. (Ed.). **Domestication and breeding: Amazonian species**. Viçosa, MG: UFV, 2012. p. 275-296.

DAHL, P. F. **Flash of the cathode rays: a history of J. J. Thomson's electron**. Bristol: Institute of Physics, 1997. 526 p.

DI CARLO, G. et al. Flavonoids: old and new aspects of a class of natural therapeutic drugs. **Life Science**, Elmsford, v. 65, p. 337-353, June 1999.

DOYLE, J. J.; DOYLE, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. **Phytochemistry**, Saint Paul, v. 19, p. 11-15, 1987.

DRANSFIELD, J. et al. A new phylogenetic classification of the palm family, Areaceae. **Kew Bulletin**, London, v. 60, n. 4, p. 559-569, 2005.

ERLICH, H. A.; GEFLAND, D.; SNINSKY, J. J. Recent advances in the polymerase chain reaction. **Science**, New York, v. 252, n. 5013, p. 1643-1651, June 1991.

EVANS, W. C. **Trease and evans pharmacognosy**. 16th ed. Nottingham: Saunders, 2009. 616 p.

FENN, J. B. et al. Electrospray ionization for mass spectrometry of large biomolecules. **Science**, New York, v. 246, p. 64-71, 1989.

FERREIRA, M. E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética**. 3. ed. Brasília: EMBRAPA, 1998. 35 p.

FERWERDA, J. D. Ecophysiology of the African oil palm. In: ALVIM, P. de (Coord.). **Ecophysiology of tropical crops**. Ilhéús: CEPLAC, 1975. v. 2, p. 1-49.

FIEHN, O. Metabolomics: the link between genotypes and phenotypes. **Plant Molecular Biology**, Dordrecht, v. 48, n. 1, p. 155-171, Jan. 2002.

FINNIGAN, R. E. Quadrupole mass spectrometers from development to commercialization. **Analytical Chemistry**, Washington, v. 66, n. 19, p. 969-975, Oct. 1994.

GEUNA, F.; HARTINGS, H.; SCIENZA, A. A new method for rapid extraction of high quality RNA from recalcitrant tissues of grapevine. **Plant Molecular Biology Reporter**, Athens, v. 16, n. 1, p. 61-67, Mar. 1998.

GHEQUIÈRE, M. et al. Polymorphisme enzymatique chez *Elaeis oleifera* H. B. K. (*E. melanococca*). **Oléagineux**, Paris, v. 42, n. 4, p. 143-150, 1997.

GONÇALVES, A. C. R. Dendezeiro (*Elaeis guineensis* Jacq.). In: CASTRO, P. R. C.; KLUGE, R. A. (Ed.). **Ecofisiologia de culturas extrativas**: cana-deaçúcar, seringueira, coqueiro, dendezeiroeoliveira. Cosmópolis:Stollerdo Brasil, 2001. p. 95-112.

HARKANSSON, P. et al. Iodine-127-plasma desorption mass spectrometry of insulin. **Journal of the American Chemical Society**, Easton, v. 104, p. 2948-2949, 1982.

HARRIS, D. C. **Análise química quantitativa**. 7. ed. Rio de Janeiro: LTC, 2008. 868p.

HIGUCHI, R. et al. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. **Biotechnology Nature**, London, v. 11, n. 11, p. 1026-1030, Nov. 1993.

HOFFMANN, E.; STROOBANT, V. **Massspectrometry**: principles and applications. 3rd ed. Chichester: J. Wiley, 2007. 502 p.

HOLLMAN, P. C. H. Evidence for health benefits of plant phenols: local or systemic effects? **Journal of the Science of Food Agriculture**, London, v. 81, n. 9, p. 842-852, July 2001.

HORNING, E. C.; HORNING, M. G. Human metabolic profiles obtained by GC and GC/MS. **Journal of Chromatographic Science**, Niles, v. 9, p. 129-140, 1971.

HUANG, E. C. et al. Atmospheric pressure ionization mass spectrometry. **Analytical Chemistry**, Washington, v. 62, n. 13, p. 713-725, July 1990.

HUGHEY, C. A.; RODGERS, R. P.; MARSHALL, A. G. Resolution of 11, 000 compositionally distinct components in a single electrospray ionization Fourier transform ion cyclotron resonance mass spectrum of crude oil. **Analytical Chemistry**, Washington, v. 74, n. 16, p. 4145-4149, Aug. 2002.

INVITROGEN. **Concert™ cytoplasmic RNA reagent**: catalog n° 12322-012. San Diego, 2001. Disponível em: <http://tools.invitrogen.com/content/sfs/manuals/cytorna_man.pdf>. Acesso em: 22 nov. 2013.

JALANI, B. S. et al. Improvement of oil through breeding and biotechnology. **Journal of the American Chemistry Society**, Washington, v. 74, n. 11, p. 1451-1455, 1997.

KANCIAN, D. L. et al. Purification of the DNA polymerase of AMV. **Biochimica et Biophysica Acta**, Alberta, v. 246, p. 365-383, 1971.

KARAS, M.; HILLENKAMP, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. **Analytical Chemistry**, Washington, v. 60, p. 2299-2301, 1988.

KINTER, M.; SHERMAN, N. E. **Protein sequencing and identification using tandem mass spectrometry**. New York: Wiley, 2000. 160 p.

KOSTIAINEN, R. et al. You have free access to this content Liquid chromatography/atmospheric pressure ionization-mass spectrometry in drug metabolism studies. **Journal of Mass Spectrometry**, New York, v. 38, n. 4, p. 357-372, Apr. 2003.

LI, J. et al. Conservation of function between mammalian and plant steroid 5 α reductases. **Proceedings of the National Academy of Sciences USA**, Washington, v. 94, n. 8, p. 3554-3559, Apr. 1997.

LIN, S. et al. LC/MS-based non-targeted metabolomics for the investigation of general toxicity of 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin in C57BL/6J and DBA/2J mice. **International Journal of Mass Spectrometry**, Amsterdam, v. 301, n. 1/3, p. 29-36, Mar. 2011.

LINHART, C.; SHAMIR, R. The degenerated primer design problem. **Bioinformatics**, Oxford, v. 18, n. 1, p. 172-180, 2002. Supplement.

LIPMAN, D. J.; PEARSON, W. R. Rapid and sensitive protein similarity searches. **Science**, New York, v. 227, n. 4693, p. 1435-1441, 1985.

LORENCE, A.; NESSLER, E. Molecules of interest: camptothecin over four decades of surprising findings. **Phytochemistry**, Saint Paul, v. 65, n. 20, p. 2735-2749, Oct. 2004.

LORENZI, H. et al. **Palmeiras do Brasil: nativas e exóticas**. Nova Odessa: Plantarum, 1996. 303 p.

LUBRANO, C.; ROBIN, J. R. Major compounds study in fruit pulp oils of six Guiana Palms species. **Acta Botanica Gallica**, Chatenay Malabry, v. 144, n. 4, p. 495-499, 1997.

LUCA, V.; LAFLAMME, P. The expanding universe of alkaloid biosynthesis. **Physiology and Metabolism**, Bethesda, v. 4, n. 3, p. 225-233, June 2001.

MACKAY, I. M. et al. Real-time PCR: history and fluorogenic chemistries. In: MACKAY, I. M. (Ed.). **Real-Time PCR in microbiology, from diagnosis to characterization**. Caister: Academic, 2007. p. 1-40.

MARRAS, A. E. S. Interactive fluorophore and quencher pairs for labeling fluorescent nucleic acid hybridization probes. **Molecular Biotechnology**, Totowa, v. 38, n. 3, p. 247-255, 2008.

MARSHALL, A. G.; HENDRICKSON, C. L.; SHI, S. D. H. Peer reviewed: scaling MS plateaus with high-resolution FT-ICRMS. **Analytical Chemistry**, Washington, v. 74, n. 9, p. 253A-259A, 2002.

MENDHAM, J. et al. **Análise química quantitativa**. 6. ed. Rio de Janeiro: LTC, 2002. 462 p.

MEUNIER, J.; BOUTIN, D. L' *Elaeis melanococca* et l'hybride *Elaeis melanococca* x *Elaeis guineensis*. **Oléagineux**, Paris, v. 30, n. 1, p. 5-8, 1975.

MIRANDA, I. P. A. et al. **Frutos da palmeiras da Amazônia**. Manaus: MCT-INPA, 2001. 120 p.

MITHEN, R. F. et al. The nutritional significance and bioavailability of glucosinolates in human foods review. **Journal of the Science of Food and Agriculture**, London, v. 80, n. 7, p. 967-984, 2000.

MORAES, M. C. B.; LAGO, C. L. Espectrometria de massas com ionização por “electrospray” aplicada ao estudo de espécies inorgânicas e organometálicas. **Química Nova**, São Paulo, v. 26, n. 4, p. 556-563, jul. /ago. 2003.

MULLIS, K. B. Target amplification for DNA analysis by the polymerase chain reaction. **Annales de Biologie Clinique**, Paris, v. 48, n. 8, p. 579-582, 1990.

MURRAY, K. K. et al. Definitions of terms relating to mass spectrometry: IUPAC recommendations 2013. **Pure and Applied Chemistry**, Oxford, v. 85, n. 7, p. 1515-1609, 2013.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. **Program selection tables of the Blast NCBI web site**. Disponível em: <http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=ProgSelectionGuide>. Acesso em: 18 nov. 2013.

NIESSEN, W. M. A. **Liquid chromatography-mass spectrometry**. New York: M. Dekker, 1998. 632 p.

NIJVELDT, R. J. et al. Flavonoids: a review of probable mechanisms of actions and potential applications. **American Journal of Clinical Nutrition**, New York, v. 74, n. 4, p. 418-425, Oct. 2001.

OLIVER, S. G. et al. Systematic functional analysis of the yeast genome. **Trends in Biotechnology**, Amsterdam, v. 16, n. 10, p. 373-378, Oct. 1998.

OOI, C. K. et al. Recovery of carotenoids from palm oil. **Journal of the American Oil Chemistry Society**, Chicago, v. 71, n. 4, p. 423-426, Apr. 1994.

PASSAGLIA, L. M. P.; ZAHA, A. **Técnicas de DNA recombinante:biologia molecular básica**. 3. ed. Porto Alegre: Mercado Aberto, 2001. 331 p.

PELT-VERKUIL, E.; BELKUM, A. van; HAYS, J. P. **Principles and technical aspects of PCR amplification**. Berlin: Springer, 2008. 332 p.

PENG, J.; GYGI, S. P. Proteomics: the move to mixtures. **Journal of Mass Spectrometry**, New York, v. 36, n. 10, p. 1083-1091, 2001.

PRIMROSE, S. B.; TWYMAN, R. M. **Principles of genome analysis and genomics**. York: Blakwell, 2003. 263 p.

PROMEGA. **Access RT-PCR system instructions for use of products A1250, A1260 AND A1280, Quick RT-PCR protocol**. Disponível em: <<http://www.promega.com.br/~media/Files/Resources/ProtCards/Access%20RT-PCR%20System%20Quick%20Protocol.pdf>>. Acesso em: 20 nov. 2013.

RAJANAIDU, N. The oil palm (*Elaeis guineensis*) collections in Africa. In: SOH, A. C.; RAJANAIDU, N.; NASIR, M. (Ed.). **Oil palm germplasm and utilization**. Bangi: Palm Oil Research Institute of Malaysia, 1986. p. 59-83.

RAMOS, E. J. A. Potencial produtivo de híbridos interespecíficos entre dendezeiro e caiauezeiro nas condições do nordeste paraense. In: RAMOS, E. J. A.; VEIGA, A. S.; FURLAN JUNIOR, J. (Ed.). **Potencial produtivo de híbridos interespecíficos entre dendezeiro e caiauezeiro nas condições do nordeste paraense**. Belém: EMBRAPA Amazônia Oriental, 2006. p. 9-23. (EMBRAPA Amazônia Oriental. Documentos, 247).

RAVEN, P. H.; EVERT, R. F.; EICHHORN, S. E. A composição molecular das células vegetais. In: RAVEN, P. H.; EVERT, R. F.; EICHHORN, S. E. (Org.). **Biologia vegetal**. 6. ed. Rio de Janeiro: Guanabara Koogan, 2001. p. 17-39.

RICE-EVANS, C. A.; MILLER, N. J.; PAGANGA, G. Structure-antioxidant activity relationships of flavonoids and phenolic acids. **Free Radical Biology and Medicine**, New York, v. 20, n. 7, p. 933-956, 1996.

RIOS, S. A. A. et al. **Recursos genéticos de palma de óleo (*Elaeis guineensis* Jacq.) e caiaué (*Elaeis oleifera* (H. B. K.) Cortés)**. Manaus: EMBRAPA Amazônia Ocidental, 2012. 39 p. (EMBRAPA Amazônia Ocidental. Documentos, 96).

ROSE, T. M. et al. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. **Nucleic Acids Research**, Oxford, v. 26, n. 7, p. 1628-1635, Apr. 1998.

SAIKI, R. K. et al. Primer-directed enzymatic amplification of DNA with a Thermostable DNA Polymerase. **Science**, New York, v. 239, n. 4839, p. 487-491, Jan. 1988.

SALZMAN, R. A. et al. An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. **Plant Molecular Biology Reporter**, Athens, v. 17, n. 1, p. 11-17, Mar. 1999.

SAMBROOK, J. E.; RUSSEL, D. W. **Molecular cloning: a laboratory manual**. 3rded. New York: Cold Spring Harbor Laboratory, 2002. 2700 p.

SAYERS, E. W. et al. Database resources of the National Center for Biotechnology Information. **Nucleic Acids Research**, Oxford, v. 37, n. 1, p. D5-15, Jan. 2009.

SCHALLEY, C. A. Supramolecular chemistry goes gas phase: the mass spectrometry examination of nanocovalent interactions in host-guest chemistry and molecular recognition. **Journal of Mass Spectrometric**, New York, v. 194, n. 1, p. 11-39, Jan. 2000.

SCHMITTGEN, T. D.; LIVAK, K. J. Analyzing real-time PCR data by the comparative CT method. **Nature Protocols**, London, v. 3, n. 6, p. 1101-1108, June 2008.

SCHWENDIMAN, J.; PALLARES, P.; AMBLARD, P. Premiers exams des accidents de fertilité chez l'hybride intersoécifique de palmier à huile *Elaeis melanococca x Elaeis guineensis*. **Oléagineux**, Paris, v. 37, p. 861-869, 1982.

SILVERSTEIN, R. M.; WEBSTER, F. X.; KIEMLE, D. J. **Spectrometric identification of organic compounds**. 7thed. Danvers: J. Wiley, 2005. 512 p.

SIMÕES, C. O. et al. **Farmacognosia da planta ao medicamento**. 5. ed. Porto Alegre: UFRGS; Florianópolis: UFSC, 1999. 1104 p.

SKOOG, D. A.; HOLLER, F. J.; NIEMAN, T. A. **Princípios de análise instrumental**. 5. ed. Porto Alegre: Bookman, 2002. 1024 p.

SNYDER, L. R.; KIRKLAND, J. J. **Introduction to modern liquid chromatography**. 2nded. Hoboken: J. Wiley, 1979. 960 p.

SOTTOMAYOR, M. et al. Peroxidases and the biosynthesis of terpenoid indole alkaloids in the medicinal plant *Catharanthus roseus* (L.) G. Don. **Phytochemistry Reviews**, Berlin, v. 3, n. 1/2, p. 159-171, Jan. 2004.

SPANAKOS, G. et al. Development of a PCR based method for diagnosis of *Leishmania* in blood samples. **Molecular and Cellular Probes**, London, v. 16, n. 6, p. 415-420, Dec. 2002.

SUGDEN, D.; WINTER, P. **Molecular biomethods handbook**. Berlin: Springer, 2008. 1124 p.

SUGGS, S. V. et al. Use of synthetic oligonucleotides as hybridization probes: isolation of cloned cDNA sequences for human beta 2-microglobulin. **Proceedings of the National Academy of Sciences USA**, Washington, v. 78, p. 6613-6617, 1981.

SURRE, C.; ZILLER, R. **La palmera de aceite**. Barcelona: Blume, 1969. 231 p. (Colección Agricultura Tropical).

SWARTZ, M. E. **Ultra Performance Liquid Chromatography (UPLC):** an introduction. Massachusetts: Separation Science Redefined, 2005. 45 p.

TAIZ, L.; ZEIGER, E. **Plant physiology.** 4thed. Sunderland: Sinauer Associates, 2006. 700 p.

TAN, G. Y. Cytology and cytogenetics. In: CORLEY, J. J. H. R. H. V.; WOOD, B. (Ed.). **Oil palm research.** Amsterdam: Elsevier Scientific, 1976. p. 61-62.

VESSECCHI, R. et al. Nomenclaturas de espectrometria de massas em língua portuguesa. **Química Nova**, São Paulo, v. 34, n. 10, p. 1875-1887, 2011.

VIEIRA, D. P. **Técnicas de PCR:** aplicações e padronização de reações. Disponível em: <http://www.fea.br/Arquivos/Biotecnologia/Material%20Prof%C2%AA%20Cristina%20%20Biologia%20Celular/Principios_da_PCR.pdf>. Acesso em: 29 out. 2013.

VITZTHUM, F. et al. Quantitative fluorescence-based microplate assay for the determination of double-stranded DNA using SYBR Green I and a standard ultraviolet transilluminator gel imaging system. **Analytical Biochemistry**, Washington, v. 276, n. 1, p. 59-64, 1999.

WATSON, J. T.; SPARKMAN, O. D. **Introduction to mass spectrometry:** instrumentation, applications and strategies for data interpretation. London: Wiley, 2007. 860 p.

YAMASHITA, M.; FENN, J. B. Electrospray ion source: another variation on the free-jet theme. **Journal of Physical Chemistry**, Washington, v. 88, p. 4451-4459, 1984.

ZHANG, Z. et al. Protein sequence similarity searches using patterns as seeds. **Nucleic Acids Research**, Oxford, v. 26, n. 17, p. 3986-3990, Sept. 1998.

SEGUNDA PARTE - ARTIGOS**ARTIGO 1 Prospecção e caracterização de genes tecido específico
(folha) de *Elaeis oleifera* (H. B. K.) Cortés**

Luiz Henrique Galli Vargas*
José Antônio de Aquino Ribeiro¹
Daniel Nogoceke Sifuentes¹
Manoel Teixeira Souza Júnior¹
Clenilson Martins Rodrigues¹
Patrícia Verardi Abdelnur¹

**Artigo preparado de acordo com a NBR 6022 (ABNT, 2003), conforme
orienta o Manual de Normalização da UFLA.**

* Universidade Federal de Lavras, Campus Universitário, CEP 37200-000 Lavras, MG. E-mail: zealotrs@gmail.com

¹ Embrapa Agroenergia, Parque Estação Biológica - PqEB s/n° - W3 Norte (final), CEP 70770-901, Brasília, DF. E-mail: jose.ribeiro@embrapa.br, daniel.sifuentes@embrapa.br, clenilson.rodrigues@embrapa.br, patricia.abdelnur@embrapa.br, manoel.souza@embrapa.br

RESUMO

Objetivou-se neste trabalho realizar a prospecção e validação de genes que possuam expressão de suas características em tecido específico (folha) visando ao auxílio a estudos futuros na busca de promotores, para que os mesmos sejam testados e validados em planta modelo (*Setaria* spp.). Para isto, foram realizadas buscas *in silico* em bancos de dados genômicos públicos e os genes candidatos selecionados foram alinhados, utilizando-se a ferramenta blast (*Basic Local Alignment Search Tools*) com o *draft* do genoma da espécie *Elaeis oleifera* (caiaué), usada como fonte de genes de interesse. A avaliação da expressão específica relativados genes candidatos em folha de caiaué foi realizada com a técnica da Reação em Cadeia da Polimerase em Tempo Real (*Real Time Quantitative PCR*- qPCR), utilizando-se para isso dois tipos de tecidos (folha e raiz). Como resultado, obteve-se indicação de maior expressão relativa do gene MyBHV5 no tecido folha quando comparado ao tecido raiz e confirmação da presença deste gene no genoma do caiaué.

Palavras-chave: Expressão relativa. Caiaué. qPCR. Promotores.

1 INTRODUÇÃO

A espécie *Elaeis guineensis* (palma de óleo) destaca-se entre as demais espécies oleaginosas por sua alta capacidade de produção de óleo, tendo produzido no mundo mais de 50 milhões de toneladas em 2012 (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS - FAO, 2012). Em relação à palma de óleo, a espécie *Elaeis oleifera* (caiaué) apresenta menor produtividade de óleo (OOI et al., 1994), mas possui características de interesse agrônômico, como tolerância ao amarelecimento fatal (AF), maior causador de prejuízos a dendeicultura mundial (BARCELOS; CUNHA; NOUY, 1998). Os trabalhos realizados visando à exploração da cultura do caiaué são focados, basicamente, na produção de híbridos entre ocaiaué e a palma de óleo para que possam expressar as características comerciais desejadas. Com esses cruzamentos deseja-se, principalmente, a obtenção de alta produtividade vinda da palma de óleo e a resistência ao principal problema que é o amarelecimento fatal como característica proveniente do caiaué (BARCELOS; NUNES; CUNHA, 2000).

Desta forma, fica evidente a importância do caiaué como fonte de variabilidade a ser empregada no programa de melhoramento da palma de óleo. Contudo, existe uma escassez de estudos que busquem a identificação de genes de interesse presentes no genoma do caiaué.

A prospecção de genes de interesse é uma via diferenciada em relação aos trabalhos que vem sendo realizados até hoje com o caiaué. Esta espécie pode ser uma fonte de genes de interesse, visto que o *draft* de seu genoma encontra-se nos servidores da Embrapa Agroenergia

(dados não publicados), podendo ser explorado de maneira bastante abrangente.

Com base na identificação de um determinado gene, seja ele constitutivo ou tecido específico, é possível que haja a detecção de seu promotor. Esses promotores poderão ser isolados, caracterizados e potencialmente utilizados para agregação de valor nas mais diversas culturas. A grande importância na obtenção de promotores está no fato que as plantas transgênicas comercializadas hoje em dia possuem transgenes comandados por promotores constitutivos que fazem com que sua expressão ocorra em todos os tecidos da planta ao em vez de somente no local onde é necessária. Uma forma de atingir a expressão pontual do transgene de interesse é utilizar promotores de tecidos específicos (SANTOS et al., 2007).

O presente estudo foi conduzido visando à prospecção de genes que demonstrem expressão em tecido específico (folha) de caiaué, utilizando-se para isto ferramentas de bioinformática e biologia molecular. Os genes tecido específicos identificados aqui, serão usados em trabalhos posteriores, onde serão buscados os seus promotores, para que os mesmos sejam testados e validados em planta modelo como a *Setaria* spp.

A etapa de procura dos genes candidatos foi realizada em bancos de dados públicos e os genes selecionados tiveram a avaliação do perfil de expressão em folhas de caiaué foi realizada com a utilização da técnica de qPCR (*Real Time Quantitative PCR*).

2 MATERIAL E MÉTODOS

Foram realizadas buscas por genes que possuíam indícios de expressão especificamente em folhas nos bancos de dados públicos (NCBI, KEGG, UNIPROTKB/SWISS-PROT, DOE JGI, DDBJ e EBI) utilizando-se palavras chaves “*leaf tissue specific*”; “*leaf specific*” de sequências que indicavam expressão em folha na espécie *Elaeis oleifera* (caiaué) ou em táxons próximos (*Elaeis*; *Elaeidinae*; *Cocoseae*; *Arecoideae* e *Commelinids*).

Os genes foram organizados em um quadro e separados por táxon, especificidade, tipo de tecido, tamanho do gene, *link* para acesso e experimentos realizados para a comprovação. Em seguida, os genes candidatos selecionados foram comparados com a base de dados do genoma do caiaué disponível nos servidores da Embrapa Agroenergia (dados não publicados), para que, assim, pudessem ser identificadas as similaridades e homologias existentes com os genes candidatos.

Nesta etapa foram utilizadas as ferramentas blastn (BASIC LOCAL ALIGNMENT SEARCH TOOL - BLAST, 2014) para comparação das sequências de nucleotídeos de entrada com a sequência de nucleotídeos da base de dados do genoma do caiaué e, também, foi realizado o tblastn, onde foi feita a comparação da sequência dos aminoácidos dos genes candidatos com a sequência de nucleotídeos da base de dados do genoma do caiaué.

Por meio da obtenção dos maiores valores possíveis do escore (*score bits*) os quais dependeram do tamanho do alinhamento, do número de *matches/mismatches/gaps* e da matriz de comparação de sequências

utilizadas e, também, resultados do *E-value* que indica a possibilidade das sequências não terem sido alinhadas ao acaso. Para o *E-value* foram utilizados valores obtidos mais próximos de zero na obtenção dos resultados, apesar de não existir um consenso sobre o valor considerado ideal, porém aceitam-se valores entre $\leq 1e^{-5}$ (0, 00001) e $\leq 1e^{-10}$ (0, 00000000001) (PROSDOSCIMI et al., 2002). Foram selecionados três genes candidatos após as etapas de blast: class I KNOX-like 1 protein (*Elaeis guineensis*), phosphoenolpyruvate carboxylase 1 (*Zea mays*) e MybHv5 (*Hordeum vulgare*).

Transcorridas as etapas de procura, identificação e seleção dos genes candidatos mediante ferramentas de bioinformática, foram realizadas as etapas de síntese de *primers*. Este passo visa à utilização dos *primers* nas etapas de PCR e qPCR. A ferramenta, para confecção dos *primers*, foram Primer3Plus (PRIMER3PLUS, 2014) e a ferramenta disponibilizada pela IDT Techonologies® (INTEGRATED DNA TECHNOLOGIES, 2014; ROZEN; SKALETSKY, 2000).

Para o desenho dos *primers*, o tamanho dos *amplicons* ficou definido entre 80 pb e 150 pb, sendo estes valores atribuídos no momento de confecção dos mesmos. A temperatura de *melting* (T_m) dos *primers* ficou estabelecida na faixa de 50°C a 65°C (SUGGS et al., 1981).

O conteúdo de guanina e citosina (G/C) foi definido entre 45 e 55 % e o tamanho dos *primers* foram definidos para ficarem entre 18 pb e 22 pb. Foram selecionados *primers* para controle negativo, um gene de *Elaeis guinnensis* (palma de óleo) denominado EgAGL2, escolhido por meio da referência: *MADS Box Genes in Oil Palm (Elaeis guineensis): Patterns in the Evolution of the SQUAMOSA, DEFICIENS, GLOBOSA,*

AGAMOUS, and *SEPALLATA* Subfamilies. Este gene é expresso somente em inflorescência de palma de óleo, não podendo ocorrer expressão em folha e raiz. O gene denominado EgEfMPOB00119 40S ribosomal protein S23 mRNA, complete cds, mRNA sequence presente em *Elaeis guineensis*, foi escolhido como controle positivo (gene constitutivo), e o gene 40S ribosomal protein S15 (rps15) mRNA, complete cd, presente em *Elaeis oleifera*, como controle positivo (gene constitutivo). As seqüências dos *primers* sintetizados podem ser visualizadas no Quadro 1.

Gene alvo	Código do iniciador	Seq. do iniciador (5' – 3')	Espécie
EgAGL2 (controle negativo)	EgAGL3	F:AGGAAAGCAACCAGGCTACT R:GGAGTGCAAGTGCTTTGTC	<i>Elaeis guineensis</i>
EgEfMPOB00119 (controle positivo) (IDT)	EgEfMPOB1 EgEfMPOB2	F:CCAGGGTTCAGCTGATTAAG R:TCGTCCAAATCCAGCAATC F:AGAAGATTGCCGCCTTTG R:CTTGACGACCTTGAACCTTAC	<i>Elaeis guineensis</i>
EgEfMPOB00119 (controle positivo) (Primer 3 Plus)	EgEfMPOB3	F:AGCCATCTTGAAACGAATG R:CACTTTCGGATGGCAGAGTT	
RPS15 (controle positivo) (IDT)	RPS151 RPS152	F:CAAAGGAACCAGAGGAACAG R:CCAGAAATGAAGCACCAGATA F:CACACCTTCGGAACATGATAA R:CGCCAAGTAGTGACCAATC	<i>Elaeis oleifera</i>
RPS15 (controle positivo) (Primer 3 Plus)	RPS153	F:TTGGCGGAGTTCTCAATCTC R:CTGGTTCCTTTGCCTGACAT	
transcription factor STM (IDT)	C1KL1 C1KL2	F:GGCGAAGGTAGTACTTA R:CCGTCTATCGCTCTCTTTA F:CCGGTACTCTCCGTACTA R:CGTGTCTGTCCGTCTATC	<i>Elaeis guineensis</i>
transcription factor STM (Primer 3 Plus)	C1KL3	F:CGCTCCCATGACTCCTACTC R:GGTTGGGGTGTGTACAGGTT	
phosphoenolpyruvate carboxylase 1 (IDT)	PPC1 PPC2	F: CCGAAATAGTGGTGTTC R: GGTCTCCTTCGTTGTAATG F: GAGTACCGAGGAAGTGT R: AACCACGGATGTAGTAGAG	<i>Zea mays</i>
phosphoenolpyruvate carboxylase 1	PPC3	F: CGCTCGTAACATTCCTTGT R: CCCGTCCCTACTCAATGAAA	

(Primer 3 Plus)			
MybHv5 (IDT)	MybHv1	F: GACTAAGGAGGAGGATGAG R: TCAGGACGGAGGTAATTG	<i>Hordeum vulgare</i> subsp. Vulgare
	MybHv2	F: TCAATTACCTCCGTCCTG R: GCAGCTATCAAAGACCATC	
MybHv5 (Primer 3 Plus)	MybHv3	F: GGCAACTTCACTGAGGAGGA R: CCTCGCAGCTATCAAAGACC	

Quadro 1 - Nomes dos genes, código dos *primers*, sequência de nucleotídeos e espécies originárias dos *primers* sintetizados

Foram coletadas raízes e folhas de caiaué em casa de vegetação, localizada no Núcleo de Apoio de Culturas Agroenergéticas (NAPE). O material é proveniente de sementes, coletadas na Estação Experimental do Rio Urubu, localizada a, aproximadamente, 140 km de Manaus (GUILLAUMET; RODRIGUES; MIRANDA, 2003). As plantas foram geradas há dois anos, por meio do resgate de embriões zigóticos, os quais foram cultivados *in vitro* e, posteriormente, as plântulas aclimatadas em casa de vegetação (VARGAS et al., 2013).

O método de coleta realizado foi por resfriamento imediato, em nitrogênio líquido (N₂), e o material foi imediatamente transportado até o Laboratório de Genética e Biotecnologia (LGB) do Centro Nacional de Pesquisa em Agroenergia (CNPAE - Embrapa Agroenergia) onde os tecidos foram armazenados em *ultrafreezer* (- 80°C) até a extração de ácido ribonucleico (RNA).

Para extração de RNA total, foi utilizando o protocolo comercial Concert™ (Invitrogen®), modificado, especificamente, para palma de óleo (VARGAS et al., 2012).

Foram macerados, aproximadamente, 100 mg de tecido em nitrogênio líquido (N₂), com auxílio de pistilo e cadinho e transferidos

para microtubos de 1,5 ml, juntamente com 500 μ L do reagente Concert™ gelado (4°C). Na sequência, os tubos foram invertidos suavemente por cinco minutos para homogeneização, permanecendo por mais cinco minutos na posição horizontal, em temperatura ambiente. Após esse período, as amostras foram submetidas à centrifugação em microcentrífuga, durante cinco minutos, a uma temperatura de 4°C, com rotação de 13200 rpm. Na sequência, o sobrenadante foi transferido para um novo tubo de polipropileno de 1,5 ml, adicionando-se 100 μ l de NaCl (cloreto de sódio) 5M e foi realizada uma homogeneização com agitação em *vórtex* por 5 segundos, após foram adicionados 300 μ l de clorofórmio em temperatura ambiente.

As amostras foram, então, homogeneizadas com auxílio de um *vórtex* por três segundos, seguidas de inversões suaves durante cinco minutos. Posteriormente, os tubos foram centrifugados, a uma velocidade de 13200 rpm, a 4°C, durante dez minutos, e a fase superior transferida para outro tubo de polipropileno de 1,5 ml, ao qual foram adicionados 500 μ l de isopropanol gelado e homogeneizando as amostras com *vórtex* por cinco segundos. Os tubos foram deixados sob temperatura ambiente por dez minutos e centrifugados a 13200 rpm, sob temperatura de 4°C, por dez minutos e o sobrenadante resultante foi descartado, mantendo-se os *pellets* formados nos tubos.

Foi adicionado 1 ml de etanol 75% (gelado) para lavagem dos *pellets*, descartando-se em seguida o sobrenadante e as amostras deixadas por quinze minutos a uma temperatura de 37°C (banho seco) para secagem. Finalmente, os *pellets* foram ressuspensos em 20 μ l de água ultra pura e deixados a 4°C (geladeira) para completa solubilização e,

após 24 horas, foram armazenados em *ultrafreezer* (- 80°C) para posterior análise de qualidade.

A quantificação das amostras de RNA foi realizada por espectrofotometria, utilizando-se o equipamento NanoDrop® ND-1000 sendo a concentração calculada em ng. μL^{-1}

A qualidade do RNA foi analisada com base nas razões da absorbância A_{260}/A_{280} e A_{260}/A_{230} , para certificação da obtenção de um RNA de elevado grau de pureza. Os RNA's obtidos e selecionados para as etapas posteriores possuem valores entre 1,8 e 2,0, para ambas as razões, garantindo, assim, qualidade assegurada (WILLIAM; KAROL; PIOTR, 1997).

Para verificação da integridade do RNA extraído, foram realizadas análises em gel de agarose 1, 2%, com 0,2 mg mL^{-1} de brometo de etídeo, imerso em tampão de ácido bórico de sódio 1 x (SB), em uma cuba de eletroforese (BRODY; KERN, 2004). A eletroforese foi realizada a 100 V por cerca de uma hora. Os géis obtidos foram visualizados e fotografados sob luz ultravioleta em sistema digital.

Para a eliminação total de DNA, utilizou-se o *kit* comercial Deoxyribonuclease I, Amplification Grade - Invitrogen®), seguindo-se o protocolo e recomendações do fabricante (INVITROGEN, 2014).

Os tubos contendo os componentes da reação (menos EDTA 25 mM) foram incubados em termociclador (Veriti, Applied Biosystems®), em uma temperatura de 25°C, por quinze minutos, em seguida foi adicionado 1 μL de EDTA (ácido etilenodiamino tetra-acético 25 mM), em cada amostra preparada, as quais foram mantidas a uma temperatura de 65°C, durante dez minutos e, na sequência, resfriadas a 4°C. O RNA

tratado foi identificado e armazenado sob uma temperatura de -80°C até a sua utilização.

A etapa seguinte, consistiu na fase de síntese de cDNA, onde se utilizou o *kit* comercial SuperScript®VILO™ MasterMix (Invitrogen®) seguindo-se o protocolo e recomendações do fabricante (INVITROGEN, 2011). Os tubos, contendo os componentes da reação, foram homogeneizados, utilizando-se um vórtex por 5 segundos e incubados em termociclador (Veriti, Applied Biosystems), em uma temperatura 25°C , por dez minutos, 42°C por sessenta minutos e, finalmente, 85°C por cinco minutos. A quantificação das amostras de cDNA foi realizada por espectrofotômetro NanoDrop® ND 1000, sendo a concentração calculada em ng. mL^{-1} .

Após a extração de RNA de folhas e raízes de caiaué, confirmação da qualidade das amostras, síntese de cDNA e RT – PCR, foram realizadas as reações de PCR (*Polymerase Chain Reaction*). Para a reação de PCR foi utilizado o termociclador Applied Biosystems, modelo Veriti.

Utilizou-se o *kit* comercial Platinum® Taq DNA Polymerase (Invitrogen®), (INVITROGEN, 2010a) seguindo-se as recomendações do fabricante. Os componentes (Taq DNA polimerase, oligonucleotídeos iniciadores, MgCl_2 e cDNA molde) foram adicionados a placas de 96 poços para PCR, as quais sempre foram mantidas em gelo durante mistura.

Na sequência, foi realizada uma breve homogeneização dos componentes, usando *vórtex* e uma centrifugação rápida (*spin*). A placa de 96 poços foi incubada em um termociclador (Veriti, Applied Biosystems), a uma temperatura de 94°C , por três minutos, para

desnaturação completa do DNA molde. Foram realizados 40 ciclos de amplificação, a etapa de desnaturação foi realizada em uma temperatura de 94°C por quarenta e cinco segundos. Já, a etapa de anelamento foi realizada de acordo com cada temperatura de *melting* com decréscimo de dois graus de cada *primer* testado, variando entre 49, 5 e 62°C, por trinta segundos e, finalmente, a etapa de extensão foi realizada em temperatura de 72°C por um minuto e meio. As amostras foram incubadas por dez minutos adicionais a 72°C para extensão final do *amplicon*.

Após o final da reação, as amostras foram armazenadas a uma temperatura de - 20°C. Como etapa final foi realizada análise por espectrofotometria, eletroforese em gel de agarose e visualização sob luz ultravioleta em sistema digital (INNIS et al., 1990).

Para a realização da PCR quantitativa em tempo real (qPCR), utilizou-se o *kit* comercial SYBR® GreenER™ qPCR SuperMix Universal (Invitrogen®) (INVITROGEN, 2010b). O protocolo foi modificado em relação ao volume da reação, utilizando-se 10 µl em substituição aos 25 µl recomendados.

O cDNA obtido anteriormente foi diluído serialmente 1:100; 1:50 e 1:25 (20 ng µl⁻¹, 40 ng µl⁻¹ e 80 ng µl⁻¹ respectivamente) os quais foram utilizados para avaliar a eficiência de amplificação dos *primers* de interesse. Os experimentos foram montados em triplicatas para todos os genes de interesse, para os *primers* do gene de controle negativo (expresso somente em inflorescência) e para os *primers* dos genes de controle positivo (genes constitutivos) em todas as três concentrações de cDNA (20 ng µl⁻¹, 40 ng µl⁻¹ e 80 ng µl⁻¹) e em ambos os tecidos (folha e raiz). Três amostras com água em substituição ao cDNA foram utilizadas

como controle negativo de amplificação para cada *primer*, denominadas NTC (*Non Template Control*) para ambos os cDNAs dos tecidos testados (folha e raiz).

Os componentes presentes no *kit* comercial foram adicionados a placas de qPCR de 96 poços do tipo universal (para reações do tipo *standard* e *fast*). O preparo da reação foi realizado sob gelo, durante a adição dos componentes, bem como em ambiente esterilizado e livre de qualquer contaminação.

As amostras foram inseridas no equipamento StepOnePlus™ Real-Time PCR e o programa utilizado para as reações e análises foi o StepOne™/StepOnePlus™*Software* v2. 3 - Life Technologies. As configurações utilizadas no *software* para amplificação foram para uma reação tipo *standard* e o método analítico utilizado para curva de *melting*, análise de amplificação e expressão gênica foi a metodologia $\Delta\Delta Ct$. As configurações utilizadas para a reação foram de: um ciclo de dois minutos e temperatura de 50°C, um ciclo de dez minutos e temperatura a 95°C, quarenta ciclos de quinze minutos a 95°C, quarenta ciclos de um minuto a 60°C e um ciclo de quinze segundos a 95°C.

A análise da curva de eficiência da diluição serial do *pool* de cDNA das amostras, além de determinar a eficiência dos *primers*, também, serviu como referencial para se estabelecer a *baseline* e o *threshold* ideal entre todos os *primers* testados para ambos os tecidos.

Depois de estabelecidas as concentrações mais eficientes dos cDNAs das amostras, o procedimento foi repetido para que se pudesse obter os dados referentes às taxas de expressão de todos os *primers* para todos os genes e em ambos os tecidos testados (folha e raiz), sendo assim

finalizada a etapa de PCR em tempo real, passando para o próximo estágio, ou seja, a análise dos dados obtidos.

As leituras da fluorescência foram realizadas pelo StepOnePlus™ Real-Time PCR a cada ciclo de amplificação e, posteriormente, analisados pelo *software* StepOne™/StepOnePlus™*Software* v2.3 - Life Technologies.

Todas as reações foram submetidas às mesmas condições, o resultado, expresso em valor de Ct, se refere ao número de ciclos de PCR necessários para que o sinal fluorescente atinja o limiar de detecção.

Os resultados individuais expressos em valores de Ct foram analisados pelo próprio *software*.

3 RESULTADOS E DISCUSSÃO

A prospecção de genes está diretamente relacionada à bioinformática a qual se torna importante para a análise de dados biológicos, contidos em bancos de dados, disponíveis na rede mundial de computadores, permitindo, assim, identificar genes de interesse com base na união das mais diversas áreas da ciência e dados disponibilizados de projetos de sequenciamento de DNA e proteínas (BINNECK, 2004). Considerando este pressuposto, com o auxílio de palavras chaves tais como (*leaf tissue specific*; "*leaf specific*") nos táxons *Elaeis*; *Elaeidinae*; *Cocoseae*; *Arecoideae* e *Commelinids*; e buscas *in silico* nos domínios públicos (NCBI, KEGG, UNIPROTKB/SWISS-PROT, DOE JGI, DDBJ e EBI) foram encontradas vinte e seis sequências que possuíam indicação de expressão em tecido específico de folha nas mais variadas espécies.

Todas as vinte e seis sequências, por meio de suas sequências FASTA de nucleotídeos, foram submetidas, inicialmente, a um blastn para comparação das sequências de nucleotídeos de entrada com a sequência de nucleotídeos. Dessa maneira, foi possível obter uma indicação de possível similaridade destas sequências com o *draft* do genoma do caiaué presente nos servidores da Embrapa Agroenergia (dados não publicados). As sequências que obtiveram os maiores valores possíveis dos parâmetros analisados (escore - *score bits*, *matches/mismatches/gaps*, matriz de comparação de sequências) foram normalizadas por variáveis estatísticas e, também, por resultados do valor *E-value*. Este resultado indicou a possibilidade das sequências não terem

sido alinhadas ao acaso e, baseado nisso, foram escolhidos os três genes candidatos.

Três sequências (genes) com melhores resultados (class I KNOX-like 1, phosphoenolpyruvate carboxylase 1 e MybHv5), após serem submetidas à ferramenta tblastn, indicaram possuírem homologia com sequências contidas no genoma do caiaué. Haja vista que os valores obtidos para os parâmetros citados anteriormente foram considerados satisfatórios e os *E-values* obtidos foram próximos de zero. Embora não exista um consenso sobre o valor considerado ideal, aceitam-se valores entre $\leq 1e^{-5}$ (0, 00001) e $\leq 1e^{-10}$ (0, 00000000001) (PROSDOSCIMI et al., 2002).

Para as etapas de extração de RNA dos tecidos folha e raiz de *Elaeis oleifera* o protocolo comercial Concert™ (Invitrogen®), modificado especificamente para palma de óleo (VARGAS et al., 2012), mostrou-se eficiente. Um bom RNA deve estar íntegro e puro, ou seja, livre das impurezas mais comumente encontradas, quando é realizada uma extração de má qualidade tais como: sais, etanol, lipídeos, polissacarídeos, proteínas, fenol, clorofórmio (WANG; TIAN; LI, 2008).

As amostras obtidas, após serem analisadas por espectrofotometria, utilizando-se o equipamento NanoDrop® ND-1000 sendo a concentração calculada em ng. μL^{-1} ficaram dentro dos valores considerados ideais. Para RNA de elevado grau de pureza as razões das absorbâncias A 260 e 280nm (A260/A280) que devem ser entre 1, 8 e 2, 2 (WILLIAM; KAROL; PIOTR, 1997), demonstram, assim, a eficiência do protocolo utilizado para ambos os tecidos (folha e raiz) de *Elaeis oleifera*,

complementando os dados publicados por Vargas et al. (2012) onde o tecido utilizado foi folha jovem (flecha) de *Elaeis guineensis*.

As integridades dos RNAs extraídos puderam ter sua qualidade analisada, também, pela análise em gel de agarose (BRODY; KERN, 2004), onde as bandas 28S, 18S e 5S das amostras de RNA puderam ser observadas, indicando RNA's íntegros. A extração de RNA intacto, também, é difícil, em virtude da degradação enzimática pela RNA presente nos tecidos (RUBIO-PIÑA; VÁZQUEZ, 2008). Comparando com dois protocolos de extração de RNA testado por Habib, Saud e Kausar (2014), Easy-Blue™ Total RNA Extraction kit - iNtRON Biotechnology, Inc. Korea e Rochester, Winer e Shah (1986) modificado, para os tecidos folha e raiz de palma de óleo, nas análises de espectrofotometria, as análises quantitativas e qualitativas, obtidas pelo protocolo utilizado neste trabalho, mostraram-se semelhantes.

Portanto, para as etapas de extração de RNA dos tecidos folha e raiz de *Elaeis oleífera* o protocolo comercial Concert™ (Invitrogen®), modificado especificamente para palma de óleo (VARGAS et al., 2012), mostrou-se eficiente.

A etapa subsequente abordou o tratamento das amostras de RNA dos tecidos folha e raiz de caiaué, objetivando a eliminação total de DNA genômico, com a utilização do kit comercial Deoxyribonuclease I, *Amplification Grade* (Invitrogen®), onde foi seguindo o protocolo do fabricante.

Quando se objetiva a realização de uma etapa de conversão do RNA em cDNA (DNA complementar) e amplificação (RT – PCR, *Reverse Transcription Polymerase Chain Reaction*), muitas vezes, a

quantidade de RNA inicial obtida não é suficiente para a aplicação satisfatória do protocolo experimental desejado (GELDER et al., 1990). A qualidade de RNA desempenha um papel importante na geração de resultados quantitativos precisos de experiências de análise da expressão genética. Um cDNA gerado, com base em um RNA que sofreu degradação, não será amplificado ou marcado com a mesma eficiência de um cDNA derivado de RNA de boa qualidade (GINGRICH; RUBIO; KARLAK, 2006). Para a síntese de cDNA, o *kit* comercial SuperScript®VILO™ MasterMix (Invitrogen®), com a utilização do protocolo do fabricante, mostrou-se eficiente para ambos os RNA's extraídos de folha e raiz (duas amostras de cada) de caiaué. Após a realização das etapas de RT – PCR, as amostras de RNA de folha e raiz de caiaué puderam ter sua qualidade e quantificação analisadas e obtendo resultados considerados satisfatórios. As razões de absorvância ficaram ente 1, 8 e 1, 91 para as faixas A 260 e 280 nm e 1, 97 e 2, 10 para as faixas A 260 e 230 nm, como também tiveram sua amplificação constatada.

Para a etapa de testes por PCR convencional, o *kit* comercial Platinum® Taq DNA Polymerase (Invitrogen®) teve sua eficiência alcançada após variações de temperatura de anelamento dos *primers* e concentração de cDNA utilizado. Os valores padrões indicados pelo fabricante para os demais componentes foram mantidos e se mostraram eficientes.

No que se refere às temperaturas de anelamento utilizadas nos experimentos de PCR convencional, as temperaturas de anelamento dos *primers* foram diminuídas 2° até 5°C da *tm* consensual dos mesmos. E,

como resultado final, optou-se utilizar a t_m -2°C e a concentração de cDNA utilizada ficou definida em $100 \text{ ng. } \mu\text{l}^{-1}$.

A temperatura de *melting* é definida como a temperatura na qual a metade dos fragmentos de DNA está na forma desnaturada e a outra metade na forma de dupla fita (FIXMAN; FREIRE, 1977), a t_m fornecida pelos fabricantes dos iniciadores serviu como parâmetro para a obtenção da temperatura de anelamento ideal, que diz respeito à temperatura na qual ocorre a junção dos *primers* ao cDNA alvo (*template*).

As informações como código dos primers, sequência de nucleotídeos, temperatura de anelamento (TA), tamanho de *amplicon* dos *primers* utilizados podem ser visualizadas na Tabela 1. Todos os *primers* foram testados para ambos os tecidos (folha e raiz).

Tabela 1 - Código dos iniciadores, sequência de nucleotídeos, temperatura de anelamento (TA) utilizada, tamanho de *amplicon* esperado

Cód. <i>Primer</i>	Sequência do primer (5' – 3')	TA - °C	<i>Amplicon</i>
C1KL2	F:CCGGTACTCTCCGTA R:CGTGTCTGTCCGTCTATC	49,5	137 pb
C1KL1	F:GGCGAAGGTAGTGACTTA R:CCGTCTATCGCTCTCTTTA	50	91 pb
PPC1	F:CCGAAATAGTGGTGTCTC R:GGTCTCCTTCGTTGTAATG	51	98 pb
PPC2	F:GAGTTACCGAGGAAGTGT R:AACCACGGATGTAGTAGAG	52	94 pb
MybHv1	F:GACTAAGGAGGAGGATGAG R:TCAGGACGGAGGTAATTG	54,9	141 pb
EgAGL3	F:AGGAAAGCAACCAGGCTACT R:GGAGTGCAAGTGTCTTTGTC	55	292 pb
EgEfMPOB2	F:AGAAGATTGCCGCCTTTG R:CTTGACGACCTTGAACCTTAC	55,5	137 pb
RPS151	F:CAAAGGAACCAGAGGAACAG R:CCAGAAATGAAGCACCAGATA	55,5	104 pb
RPS152	F:CACACCTTCGGAACATGATAA R:CGCCAAGTAGTGACCAATC	55,5	119 pb
PPC3	F:CGCCTCGTAACATTCCTTGT R:CCCGTCCCTACTCAATGAAA	62	120 pb
MybHv2	F:TCAATTACCTCCGTCCTG R:GCAGCTATCAAAGACCATC	55,95	112 pb
EgEfMPOB1	F:CCAGGGTTCAGCTGATTAAG R:TCGTCCAAATCCAGCAATC	56,5	113 pb
EgEfMPOB3	F:AGCCATCTTGGAACGAATG R:CACTTTCGGATGGCAGAGTT	58	116 pb
RPS153	F:TTGGCGGAGTCTCAATCTC R:CTGGTTCCTTTGCCTGACAT	58	122 pb
C1KL3	F:CGCTCCCATGACTCCTACTC R:GGTTGGGGTGTGTACAGGTT	58	123 pb
MybHv3	F:GGCAACTTCACTGAGGAGGA R:CCTCGCAGCTATCAAAGACC	62,1	87 pb

Após os resultados serem analisados por eletroforese em gel de agarose, pode-se perceber que os *primers* PPC2 (gene transcription factor STM - *Zea mays*), RPS51 (controle positivo - *Elaeis oleifera*), MybHv2 e MybHV3, (gene MybHv5 - *Hordeum vulgare*) EgEfMPOB1 e EgEfMPOB3 (controle positivo - *Elaeis guineensis*) demonstraram possíveis ampliações nos tamanhos esperados dos seus *amplicons*, 94,

104, 112, 87, 113 e 116 pares de bases, respectivamente. Os *amplicons* gerados demonstram, também, amplificação em ambos os tecidos testados (folha e raiz) de caiaué. Enquanto o *primer* EgAGL3 (controle negativo - *Elaeis guineensis*) não demonstrou amplificação em ambos os tecidos, indo de acordo com o resultado obtido por Adam et al. (2006).

Os demais iniciadores apresentaram amplicons que não corresponderam aos tamanhos esperados, com possível existência de auto complementariedade, ou ligações inespecíficas. A autocomplementariedade possibilita a extensão de um dímero, durante a fase de polimerização da PCR, ou, simplesmente, impede que o par de *primers* hibridize com a sequência alvo (KAMPKE; KIENINGER; MECKLENBURG, 2001). No gel de agarose, esses dois tipos de ligações podem ser visíveis como produtos amplificados, geralmente, abaixo dos 100 pares de bases (bp) ou uma banda intensa, e, geralmente, com maior tamanho da sequência alvo esperada (LIFE TECHNOLOGY, 2013).

Obtidos os resultados preliminares para testes de *primers* por PCR convencional, deu-se início à etapa de PCR em tempo real (qPCR). O *kit* comercial SYBR® GreenER™ qPCR SuperMix Universal (Life Technologies™) foi utilizado com adaptações do protocolo, sugerido pelo fabricante, após essas modificações realizadas durante os testes, foram definidas as melhores concentrações dos reagentes (*kit* comercial), concentração de cDNA (folha e raiz) e concentração de *primers*.

O protocolo final adaptado ao total da reação utilizada foi de 10 µl, em substituição à reação de 25 µl sugerida pelo fabricante, sendo assim uma redução significativa de 60% de reagente utilizado, bem como redução significativa na quantidade de cDNA e iniciadores. As

concentrações de cDNA que apresentaram os melhores resultados, após a série de diluições realizadas, foram os realizados com concentração de $80 \text{ ng } \mu\text{l}^{-1}$, com menor erro entre as triplicatas, maior nível de expressão final e com menor desvio padrão. A concentração dos *primers* utilizada e com resultado eficiente foi a recomendada pelo fabricante do *kit* comercial, ou seja, $10 \text{ } \mu\text{M}$. No que se diz respeito à temperatura de anelamento dos *primers*, após testes e variações, constatou-se que a temperatura de 60°C foi mais eficiente, diferentemente da PCR convencional, onde foi utilizada a TA específica para cada um. Para a eficiência da técnica da PCR em tempo real os *primers* geralmente apresentam temperatura de anelamento ideal entre 58 e 60°C (DHEDA et al., 2004).

Nestes estudos, puderam ser identificados, dentre todos os dezesseis pares de iniciadores sintetizados e os dois tecidos (folha e raiz) de caiaué, dois genes que apresentaram diferenças de expressão significativas, ambos os tecidos vegetais por meio da metodologia $\Delta\Delta\text{Ct}$ (quantificação relativa). A quantificação relativa é subdividida em dois métodos, o método da curva padrão e o método do Ct comparativo (APPLIED BIOSYSTEMS, 2010). Nas análises realizadas, foi utilizado o método do Ct comparativo para determinar os diferentes níveis de expressão entre os dois tecidos.

Chegou-se a um resultado onde apenas os *primers* MybHV3 (gene MybHv5 - *Hordeum vulgare*), EgEfMPOB3 (gene EgEfMPOB00119 - controle positivo - *Elaeis guineensis*) tiveram ampliações eficientes, incluindo a não amplificação dos NTC's (*Non Template Control*). Estes

resultados foram observados e analisados em três diferentes pontos, curva de amplificação, curva de *melting* (*melt curve*) e quantificação relativa.

Na Figura 1, podem ser observados os resultados referentes à curva de amplificação (*amplification plot*) onde as três fases existentes (exponencial, linear e platô) podem ser visualizadas.

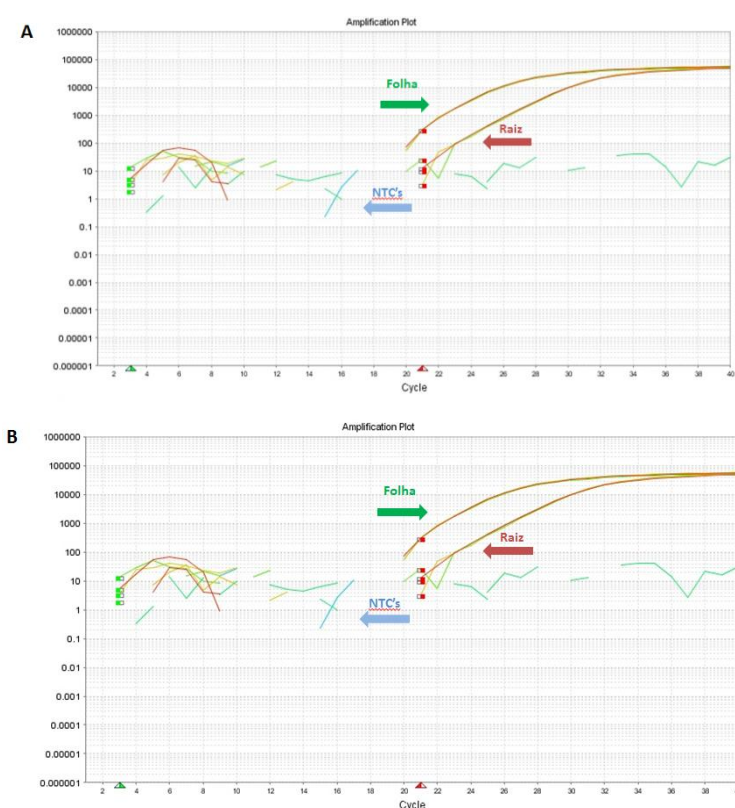


Figura 1 – (A) demonstra a curva de amplificação do gene EgEfMPOB00119 da espécie *Elaeis guineensis* (primer EgEfMPOB3 - controle positivo) com ambos os cDNAs (folha e raiz) a uma concentração de $80 \text{ ng. } \mu\text{l}^{-1}$, onde não ocorreram amplificações das amostras livres de cDNA (NTC's); (B) demonstra a curva de amplificação do gene MybHv5 da espécie *Hordeum vulgare* (primer MybHv3) com ambos os cDNAs (folha e raiz) a uma concentração de $80 \text{ ng. } \mu\text{l}^{-1}$, onde não ocorreram amplificações das amostras livres de cDNA (NTC's)

Observando-se a Figura 1, é notável o pequeno erro existente em ambas as amostras, as linhas de amplificação estão muito próximas, indicando, assim, um baixo erro de pipetagem. Assim como podem ser observadas as três fases existentes na curva de amplificação, a linha exponencial, não havendo amplificação, fase linear onde a quantidade de produtos da PCR dobra a cada ciclo e a fase platô, quando não há mais aumento no número de produtos (NOVAIS; PIRES-ALVES, 2004).

Em relação às análises das curvas de *melting*, os resultados são demonstrados na Figura 2, para os mesmos iniciadores e genes.

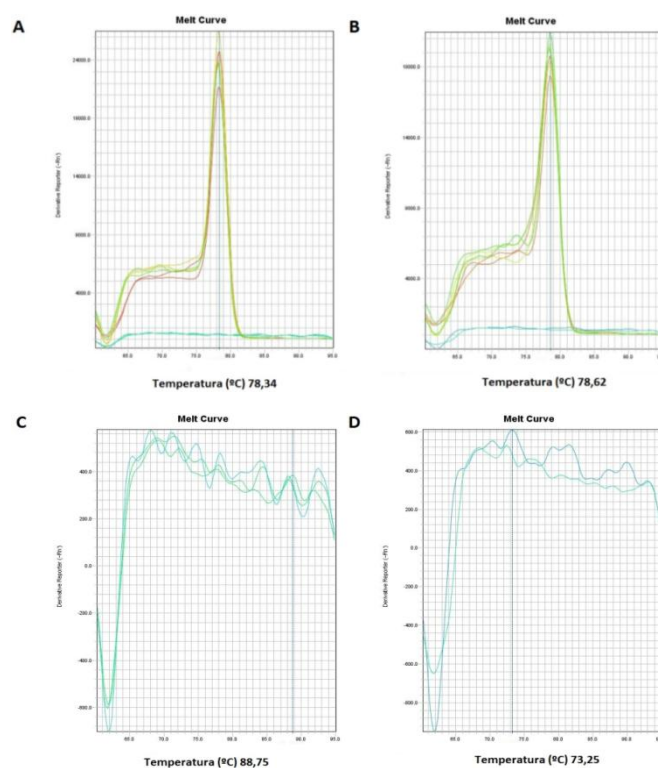


Figura 2 – (A) demonstra a curva de *melting* do gene EgEfMPOB00119 da espécie *Elaeis guineensis* (primer EgEfMPOB3 - controle positivo) com ambos os cDNAs (folha e raiz) a uma concentração de 80 ng. μl^{-1} , onde não ocorreram amplificações das amostras livres de cDNA (NTC's); (B) demonstra a curva de *melting* do gene MybHv5 da espécie *Hordeum vulgare* (primer MybHv3) com ambos os cDNAs (folha e raiz) a uma concentração de 80 ng. μl^{-1} , onde não ocorreram amplificações das amostras livres de cDNA (NTC's), (C) e (D) são as curvas de *melting* dos NTC's dos genes EgEfMPOB00119 e MybHv5, respectivamente

Os gráficos da curva de *melting* representam as alterações na fluorescência que está ligada ao DNA fita dupla, assim que este DNA é aquecido ocorre detecção da diminuição deste fluoróforo, em virtude da dissociação das cadeias de DNA e subsequente liberação do corante. Essa fluorescência é registrada em relação à temperatura e, assim, pode-se

representar a temperatura onde metade da molécula de DNA encontra-se em fita dupla e simples, ou seja, a curva de *melting*.

As avaliações dos resultados por curva de *melting* reduzem a necessidade de análises posteriores em gel de agarose, porém são necessárias para as verificações de ligações inespecíficas e autocomplementariedades. A curva padrão fornece informações importantes sobre a eficiência de amplificação, consistência das replicatas, a inclusão da curva de dissociação é essencial para verificar se houve formação de um único produto ou se produtos inespecíficos, também, foram formados (LIFE TECHNOLOGY, 2013). Os resultados das *tm*'s dos produtos gerados foram 78, 34°C (gene EgEfMPOB00119) e 78, 62°C (gene MybHv5), enquanto os NTC's, para ambos os genes, apresentaram as *tm*'s de 88, 75 e 73, 25°C, respectivamente.

O último dos três pontos analisados neste estudo foi a quantificação relativa dos dois tecidos utilizados, folha e raiz. A expressão relativa é determinada com uma comparação da expressão do gene-alvo (*template*) com genes controle. Esses genes são usados como controles internos da expressão gênica, pois, teoricamente, possuem uma expressão constitutiva, em todos os locais do organismo estudado. Possuindo esta característica, esses genes são utilizados para aferimento da expressão gênica do gene-alvo. Genes ribossômicos costumam ser utilizados para este fim, pois possuem expressão constitutiva (DHEDA et al., 2004).

As análises foram realizadas pela metodologia $\Delta\Delta C_t$ por quantificação relativa. A quantificação relativa descreve mudanças na

expressão do gene alvo em relação a alguma amostra de referência (genes endógenos) demonstrando em ordens de grandeza (WINER et al., 1999).

Na Figura 3, podem ser observados os valores encontrados para a quantificação relativa dos dois tecidos testados para o gene MybHv5 (*primer* MyBHv3).

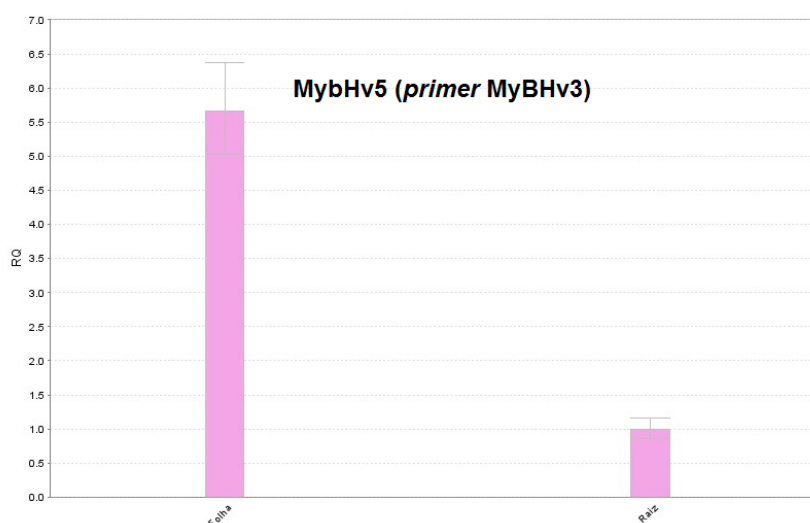


Figura 3 - Gráfico com os valores da quantificação relativa dos tecidos folha e raiz de caiaué, expressos em ordens de grandeza. No eixo X estão os tecidos e no eixo Y os valores de expressão relativa. As barras de erro das repetições biológicas representadas por análise de variância mostraram que resultado foi significativo $<0,05$

Os resultados obtidos demonstram que o gene MybHv5 (*primer* MyBHv3) possuiu uma expressão relativa entre 5,0 e 6,4 vezes, quando comparada ao tecido raiz que teve sua expressão entre 0,8 e 1,2 vezes. A maior expressão atingida pelo gene MybHv5 pode chegar a um valor de 5,6 vezes maior em folha em relação à raiz de caiaué. Enquanto o menor

valor obtido de expressão pode ser de até 3, 8 vezes maior em folha em relação à raiz de caiué.

4 CONCLUSÃO

O gene EgEfMPOB00119, da espécie *Elaeis guineensis* (primer EgEfMPOB3 - controle positivo), mostrou-se eficiente para normalização dos resultados obtidos, teve sua presença confirmada no *draft* do genoma de caiaué.

Indicação de maior expressão relativa do gene MyBHV5 (primer MyBHv3) no tecido folha quando comparado ao tecido raiz e confirmação da presença deste gene no *draft* do genoma do caiaué.

Search and characterization of tissue-specific genes (leaf) of *Elaeis oleifera* (HBK) Cortes.

ABSTRACT

The aim of this study was to perform exploration and validation of genes that have their expression in specific tissue characteristics (leaf) in order to aid future studies in the search for promoters, so that they are tested and validated in model plant (*Setaria* spp.). For this search was performed in silico on public genomic databases, and data selected candidate genes were aligned using the blast tool (Basic Local Alignment Search Tools) as draft of the species *Elaeis oleifera* (caiaué) genome as a source of genes of interest. The evaluation of specific relative expression of candidate genes in leaf caiaué was performed with the technique of Polymerase Chain Reaction in Real Time (Real Time Quantitative PCR - qPCR), using it for the two tissue types (leaf and root) . As a result was obtained indicating highest relative MyBHV5 gene expression in leaf tissue compared to root tissue and confirmation of the presence of this gene in the genome caiaué.

Keywords: Relative expression. Caiaué. qPCR. Promoters.

REFERÊNCIAS

ADAM, H. et al. MADS box genes in oil palm (*Elaeis guineensis*): patterns in the evolution of the Squamosa, Deficiens, Globosa, Agamous, and Sepallata Subfamilies. **Journal of Molecular Evolution**, Berlin, v. 62, n. 1, p. 15-31, Jan. 2006.

APPLIED BIOSYSTEMS. **Applied biosystems 7500/7500 Fast Real-Time PCR system relative standard curve and comparative CT experiments: getting started guide**. Disponível em: <http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_050330.pdf>. Acesso em: 10 jan. 2014.

BARCELOS, E.; CUNHA, R. N. V. da; NOUY, B. Recursos genéticos de dendê (*Elaeis guineensis*) e *Elaeis oleifera* disponíveis na Embrapa e sua utilização. In: FURLAN JÚNIOR, J.; MULLER, A. A. (Ed.). **Agronegócio do dendê: uma alternativa social, econômica e ambiental para o desenvolvimento sustentável da Amazônia**. Belém: EMBRAPA, 1998. p. 131-143.

BARCELOS, E.; NUNES, C. D.; CUNHA, R. N. V. **Melhoramento genético e produção de sementes comerciais de dendezeiro**. In: Viégas, I. J.; Müller, A. A. A cultura do dendezeiro na Amazônia Brasileira. Belém. Embrapa Amazônia Oriental. p. 145-174, 2000.

BASIC LOCAL ALIGNMENT SEARCH TOOL. **BLAST assembled genomes**. Disponível em: <<http://blast.ncbi.nlm.nih.gov/Blast.cgi>>. Acesso em: 10 fev. 2014.

BINNECK, E. As ômicas: integrando a bioinformação. **Biociência e Desenvolvimento**, Brasília, n. 32, p. 28-37, 2004.

BRODY, J. R.; KERN, S. E. Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis. **BioTechniques**, Natick, v. 36, n. 2, p. 214-216, Feb. 2004.

DHEDA, K. et al. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. **Biotechniques**, Natick, v. 37, n. 1, p. 112-114, July 2004.

FIXMAN, M.; FREIRE, J. J. Theory of DNA melting curves. **Biopolymers**, New York, v. 16, p. 2693-2704, 1977.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **The statistics division of FAO**. Disponível em: <<http://faostat.fao.org/site/636/DesktopDefault.aspx?PageID=636#ancor>>. Acesso em: 13 jan. 2014.

GELDER, R. N. V. et al. Amplified RNA synthesized from limited quantities of heterogeneous cDNA. **PNAS**, Washington, v. 87, n. 5, p. 1663-1667, Mar. 1990.

GINGRICH, J.; RUBIO, T.; KARLAK, C. Effect of RNA degradation on data quality in quantitative PCR and microarray Experiments. **Bio-Rad Bulletin**, Cambridge, n. 5452, p. 1-6, 2006.

GUILLAUMET, J. L.; RODRIGUES, M. R. L.; MIRANDA, I. P. A. A estação experimental do rio Urubu: Embrapa Amazônia Ocidental. In: MIRANDA, I. P. A. et al. (Ed.). **Ecosistemas florestais em áreas manejadas na Amazônia**. Manaus: Instituto Nacional de Pesquisas da Amazônia, 2003. p. 30-64.

HABIB, S. H.; SAUD, H. M.; KAUSAR, H. Efficient oil palm total RNA extraction with a total RNA extraction kit. **Genetics and Molecular Research**, Ribeirão Preto, v. 13, n. 2, p. 2359-2367, 2014.

INNIS, M. A. et al. **PCR protocols: a guide to methods and applications**. San Diego: Academic, 1990. 482 p.

INTEGRATED DNA TECHNOLOGIES. Disponível em: <<http://www.idtdna.com/primerquest/Home/Index>>. Acesso em: 10 fev. 2014.

INVITROGEN. **Deoxyribonuclease I, amplification grade protocol**. San Diego, 2014. Disponível em: <<http://tools.lifetechnologies.com/content/sfs/manuals/18068015.pdf>>. Acesso em: 22 mar. 2014.

INVITROGEN. **Platinum® Taq DNA Polymerase**. San Diego, 2010a. Disponível em: <<http://www.lifetechnologies.com/content/dam/LifeTech/migration/files/pcr/pdfs.par.26652.file.dat/platinumtaq-pps.pdf>>. Acesso em: 20 fev. 2014.

INVITROGEN. **SuperScript® VILO™ MasterMix**. San Diego, 2011. Disponível em: <http://tools.lifetechnologies.com/content/sfs/manuals/SuperScript_VILO_MasterMix_man.pdf>. Acesso em: 22 nov. 2013.

INVITROGEN. **SYBR® GreenER™ qPCR SuperMix Universal**. San Diego, 2010b. Disponível em: <http://tools.lifetechnologies.com/content/sfs/manuals/sybrgreener_supermix_uni_man.pdf>. Acesso em: 27 jan. 2014.

KAMPKE, T.; KIENINGER, M.; MECKLENBURG, M. Efficient primer design algorithms. **Bioinformatics**, Oxford, v. 17, n. 3, p. 214-225, Mar. 2001.

LIFE TECHNOLOGY. **Real-time PCR handbook**. Disponível em: <<http://find.lifetechnologies.com>>.

com/Global/FileLib/qPCR/RealTimePCR_Handbook_Update_FLR.pdf>. Acesso em: 15 dez. 2013.

NOVAIS, C. M.; PIRES-ALVES, M. PCR em tempo real. **Revista Biotecnologia Ciência & Desenvolvimento**, Brasília, v. 33, n. 1/3, p. 10-13, jul. /dez. 2004.

OOI, C. K. et al. Recovery of carotenoids from palm oil. **Journal of the American Oil Chemistry Society**, Chicago, v. 71, n. 4, p. 423-424, Apr. 1994.

PRIMER3PLUS: pick primers from a DNA sequence. Disponível em: <<http://primer3plus.com/>>. Acesso em: 10 fev. 2014.

PROSDOSCIMI, F. et al. Bioinformática: manual do usuário. **Revista de Biotecnologia**, Goiânia, v. 29, n. 5, p. 12-25, nov. /dez. 2002.

ROCHESTER, D. E.; WINER, J. A.; SHAH, D. M. The structure and expression of maize genes encoding the major heat shock protein, hsp70. **EMBO Journal**, Oxford, v. 5, p. 451-458, 1986.

ROZEN, S.; SKALETSKY, H. J. Primer3 on the WWW for general users and for biologist programmers. In: MISENER, S.; KRAWETZ, S. A. (Ed.). **Bioinformatics methods and protocols: methods in molecular biology**. Totowa: Humana, 2000. p. 365-386.

RUBIO, P. J. A.; VÁZQUEZ, F. F. A. Isolation of functional total RNA from *Argemone Mexicana* tissues. **Electronic Journal of Biotechnology**, Valparaíso, v. 11, n. 4, p. 15-16, Oct. 2008.

SANTOS, D. B. M. et al. Identificação de ESTs tecido-específicas no banco de dados do genoma funcional de *Coffea* spp. In: SIMPÓSIO DE PESQUISA DE CAFÉS DO BRASIL, 5., 2007, Águas de Lindóia. **Anais. . . Águas de Lindóia: SBC**, 2007. p. 12-13.

SUGGS, S. V. et al. Use of synthetic oligonucleotides as hybridization probes: isolation of cloned cDNA sequences for human beta 2-microglobulin. **Proceedings of the National Academy of Sciences USA**, Washington, v. 78, p. 6613-6617, 1981.

VARGAS, L. H. G. et al. Cultivo *in vitro* e aclimatização de embriões zigóticos de subamostras de palma de óleo (*Elaeis spp.*) mantidas em banco de germoplasma. In: CONGRESSO INTERNACIONAL DE BIOENERGIA, 8., 2013, São Paulo. **Anais. . .** São Paulo: Bioenergia, 2013. p. 123-129.

VARGAS, L. H. G. et al. Estabelecimento de um protocolo eficiente para extração de rna a partir de folhas jovens de *Elaeis guineensis* Jacq. In: WORKSHOP AGROENERGIA, 6., 2012, Ribeirão Preto. **Anais. . .** Ribeirão Preto: FAPESP, 2012. p. 110-113.

WANG, X.; TIAN, W.; LI, Y. Development of an efficient protocol of RNA isolation from recalcitrant tree tissues. **Molecular Biotechnology**, Totowa, v. 38, n. 1, p. 57-64, Jan. 2008.

WILLIAM, W.; KAROL, M.; PIOTR, C. Effect of pH and ionic strength on the spectrophotometric assessment of nucleic acid purity. **BioTechniques**, Natick, v. 22, n. 3, p. 474-481, Mar. 1997.

WINER, J. et al. Development and validation of real-time quantitative reverse transcriptase-polymerase chain reaction for monitoring gene expression in cardiac myocytes *in vitro*. **Analytical Biochemistry**, New York, v. 270, n. 1, p. 41-49, May 1999.

ARTIGO 2 **Extração e análise de metabólitos secundários de folhas de *Elaeis guineensis* Jacq por cromatografia líquida de Ultra Alta eficiência acoplada à espectrometria de massas**

Luiz Henrique Galli Vargas*
José Antônio de Aquino Ribeiro¹
Daniel Nogoceke Sifuentes¹
Manoel Teixeira Souza Júnior¹
Clenilson Martins Rodrigues¹
Patrícia Verardi Abdelnur¹

Artigo preparado de acordo com a NBR 6022 (ABNT, 2003), conforme orienta o Manual de Normalização da UFLA.

* Universidade Federal de Lavras, Campus Universitário, CEP 37200-000 Lavras, MG. E-mail: zealotrs@gmail.com

¹ Embrapa Agroenergia, Parque Estação Biológica - PqEB s/n° - W3 Norte (final), CEP 70770-901, Brasília, DF. E-mail: jose.ribeiro@embrapa.br, daniel.sifuentes@embrapa.br, clenilson.rodrigues@embrapa.br, patricia.abdelnur@embrapa.br, manoel.souza@embrapa.br

RESUMO

Apesar da palma de óleo (*Elaeis guineensis*) destacar-se entre as demais espécies oleaginosas por sua alta capacidade de produção de óleo, há poucos metabólitos secundários identificados e descritos na literatura no gênero *Elaeis*. Assim, como há um pequeno número de trabalhos referentes à utilização da metabolômica no estudo deste gênero objetivou-se neste trabalho identificar os metabólitos secundários presentes em folhas de *Elaeis guineensis* Jacq (palma de óleo), utilizando a abordagem metabolômica. Para isto, foi desenvolvido um protocolo de análise dos metabólitos, baseado na extração e detecção dos metabólitos por técnicas analíticas avançadas, baseadas em cromatografia líquida de ultra alta eficiência, acoplada à espectrometria de massas (UHPLC-MS). Pela análise dos resultados possibilitou-se a identificação de metabólitos com base na detecção das respectivas razões m/z , as quais geraram fórmulas moleculares correspondentes, que foram associadas a compostos descritos na literatura. Os metabólitos presentes nas duas frações analisadas (polar e lipídica) deverão ser validados por experimentos de espectrometria de massas *tandem* (MS/MS).

Palavras-chave: *Elaeis guineensis*. Palma de óleo. Metabolômica. Metabólitos secundários. Espectrometria de massas. UHPLC-MS. *Electrospray*. *Time of flight*.

1 INTRODUÇÃO

O gênero *Elaeis* possui duas espécies: *Elaeis guineensis* Jacq. (palma de óleo) e *Elaeis oleifera* (Kunth) Cortés (caiaué) (CUNHA; LOPES; BARCELOS, 2009). Os plantios comerciais no Brasil são quase, em sua totalidade, referentes à palma de óleo e híbridos interespecíficos, que são cruzamentos entre a palma de óleo e a espécie americana caiaué.

A palma de óleo destaca-se entre as demais espécies oleaginosas por sua alta capacidade de produção de óleo (*oil palm*) (FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS - FAO, 2012). A palma de óleo nos plantios comerciais pode chegar a uma produtividade de até 5 mil Kg ha⁻¹ de óleo e apresenta maior produtividade, quando comparada com a soja (*Glycine max* (L.) Merr.), que é a cultura mais plantada no Brasil (KALTNER et al., 2004). O óleo de palma, extraído da polpa (mesocarpo) e o óleo de pamiste, extraído das nozes, podem ser destinados para os mais diversos fins, desde a produção de biodiesel, indústria alimentícia, medicinal e cosmética (MYA; SOMVONG; CHATCHAWAN, 2010). No entanto, a palma de óleo possui susceptibilidade ao principal problema, causador de prejuízos nessa cultura, o amarelecimento fatal (AF). A hibridação interespecífica entre o caiaué e a palma de óleo, tem sido explorada com o objetivo de desenvolver cultivares com resistência a pragas e doenças, principalmente, ao AF (BARCELOS et al., 2002).

As plantas são as espécies que têm contribuído de forma mais significativa para o fornecimento de metabólitos secundários, existindo potencialidades exploratórias para esses compostos, pois possuem um alto

valor econômico em virtude da sua utilização na produção de medicamentos, cosméticos, alimentos e agroquímicos (PINTO et al., 2002). A técnica que estuda o metabolismo, primário ou secundário de um organismo, seja de forma qualitativa ou quantitativa, é chamada de metabolômica (HALL, 2006).

Neste contexto, o gênero *Elaeis* possui poucos metabólitos secundários identificados e descritos na literatura, assim como há um pequeno número de trabalhos referentes à utilização da metabolômica no estudo deste gênero.

Objetivou-se neste trabalho identificar os metabólitos secundários presentes em folhas de palma de óleo, utilizando a abordagem de metabolômica. Para tanto, foi necessário estabelecer um protocolo de análise eficiente e confiável para a extração, detecção e identificação destes compostos. Para isto, utilizaram-se técnicas analíticas avançadas, baseadas em cromatografia líquida de ultra eficiência, acoplada à espectrometria de massas (UHPLC-MS), com ionização por *electrospray* (ESI) e analisador de massas de alta resolução tipo Q-TOF (*Quadrupole-Time-Of-Flight*).

2 MATERIAL E MÉTODOS

2.1 Obtenção do material biológico

Foram coletadas cinco amostras (repetições técnicas) da folha +17 de cinco diferentes plantas (repetições biológicas) de *Elaeis guineensis*, localizadas no campo experimental de dendezeiro do Centro de Pesquisas Agropecuárias do Cerrado – CPAC, em Planaltina - DF. As plantas foram escolhidas de forma aleatória.

Para coleta do material vegetal, foi utilizado um alicate de extração de disco foliar com diâmetro padrão de 6,35 mm. Foram coletados 16 discos foliares por amostra, totalizando 100 mg de material por amostra. Os materiais foram inseridos em tubos individuais de polipropileno de 2,0 mL, já identificados e previamente resfriados. As amostras foram imediatamente resfriadas em nitrogênio líquido (N₂) e transportadas até o laboratório da Central de Análises Químicas e Instrumentais (CAQ) do Centro Nacional de Pesquisa em Agroenergia (CNPAAE - Embrapa Agroenergia) e armazenadas em *ultrafreezer* (-80°C).

2.2 Extrações

O protocolo para extração de metabólitos secundários seguido foi baseado no protocolo publicado pelo *Max-Planck-Institute of Molecular Plant Physiology* (MAX-PLANCK-INSTITUTE OF MOLECULAR PLANT PHYSIOLOGY, 2014).

Foram preparadas duas soluções estoques, denominadas solução de extração 1 e solução de extração 2. A solução estoque de extração 1 foi preparada com 25 mL de metanol (Panreac- *UPLC grade*) e 75 mL de *tert-Butyl methyl ether* (MTBE – Sigma – grau ACS) e homogeneizada com auxílio de vórtex. A solução foi pré-resfriada a -20°C por 1 hora antes de seu uso. A solução estoque de extração 2 foi preparada com 75 mL de água Milli-Q (*UPLC grade*) e 25 mL de metanol (Panreac - *UPLC grade*) sendo homogeneizada com auxílio de vortex e utilizada em temperatura ambiente no momento da extração. O procedimento de extração foi realizado nos laboratórios da Central de Análises Químicas e Instrumentais (CAQ) e Laboratório de Genética e Biotecnologia (LGB) da Embrapa Agroenergia.

No material vegetal resfriado em nitrogênio líquido e já inserido em tubo de polipropileno de 2 mL, foram adicionadas 2 esferas de metal de 3 mm aos microtubos e processadas no equipamento moinho de bolas (Biospec) pelo período de 45 segundos, objetivando a lise das células das folhas. Imediatamente após, foi adicionado 1 mL da solução estoque de extração 1, gelada por amostra, incubando-se na sequência os tubos em um “banho seco” (*drybath* - Eppendorf®) a uma temperatura de 4°C por 10 minutos. Terminado o processo, foi realizada uma “sonicação” em banho ultrassônico, contendo água e gelo, durante 10 minutos, adicionando-se na sequência 500 µL da solução estoque da mistura de extração 2. Ao final da sonicação, as amostras foram centrifugadas em temperatura ambiente sob rotação de 15000 rpm em uma centrífuga refrigerada (Eppendorf®) por 5 minutos a uma temperatura de 4°C.

Foi observada a formação de duas fases: uma orgânica de coloração verde (fração lipídica) e uma fase polar de coloração clara (fração polar), as quais foram separadas com auxílio de micropipetas e transferidas para tubos de polipropileno de 1,5 mL distintos. As amostras foram concentradas em concentrador a vácuo (LabConco[®] - *speedvac*), por 24 horas, passado o período, as amostras foram identificadas e armazenadas em *ultrafreezer* (Sanyo[®] – VIP) sob temperatura de -80°C até serem analisadas.

2.3 Recuperação dos extratos

A fração lipídica foi ressuspendida com 500 µL de mistura de isopropanol e acetonitrila (ACN) na proporção 30:70 (v/v) + 0,1% de ácido fórmico. Os microtubos, contendo as amostras, foram submetidos a um *vortex* para a obtenção de uma mistura homogênea. Na sequência, as amostras foram “sonicadas” em um banho ultrassônico, durante 5 minutos e centrifugadas em temperatura ambiente a 13400 rpm por 10 minutos. Com auxílio de uma micropipeta, foram retirados cerca de 400 µL da solução e transferidos para frascos tipo *vial* de 2 mL e analisadas imediatamente.

A fração polar foi ressuspendida com 500 µL de água ultra pura (Milli-Q) e os microtubos, contendo as amostras, foram vortexados para a obtenção de uma mistura homogênea. Na sequência, foram submetidos a um banho ultrassônico, durante 5 minutos e centrifugados em temperatura ambiente a 13400 rpm por 10 minutos. Com auxílio de uma micropipeta

foram retirados cerca de 400 μ L da solução e transferidos para frasco tipo *vial* de 2 ml e analisados imediatamente.

2.4 Métodos analíticos

Para a determinação estrutural de diferentes biomoléculas, principalmente as presentes no metabolismo secundário em folhas de palma de óleo, foi utilizada cromatografia líquida de ultra alta eficiência, acoplada à espectrometria de massas (UHPLC-MS), com fonte de ionização do tipo ESI e analisador híbrido tipo Q-TOF.

As amostras foram submetidas à análise por UHPLC-MS. Para separação cromatográfica da fração lipídica, foi empregada uma coluna de fase reversa modelo Acquity BEH C8 (2. 1 x 150 mm, 1. 7 μ m), mantida à temperatura de 45°C. Como fase móvel, foi empregado um gradiente de eluição composto por Fase Móvel A (ácido fórmico 0, 1% (V/V) em água ultrapura) e Fase Móvel B (mistura 70:30 (V/V) de acetonitrila:isopropanol + 0, 1 % de ácido fórmico). Um fluxo de injeção de 0, 400 mL. min⁻¹ com um total de 0, 2 ml de volume injetado.

Para separação cromatográfica da fração polar, foi empregada uma coluna de fase reversa modelo Acquity BEH C8 (2. 1 x 150 mm, 1. 7 μ m), mantida à temperatura de 45°C. Como fase móvel foi empregado um gradiente de eluição composto por Fase Móvel A (ácido fórmico 0, 1% (V/V) em água ultrapura) e Fase Móvel B (mistura Acetonitrila 100 % + 0, 1 % de ácido fórmico). Um fluxo de injeção de 0, 400 mL. min⁻¹ com um total de 1 mL injetado.

As análises foram realizadas, utilizando um cromatógrafo líquido de Ultra Alta Eficiência (UHPLC) (marca Nexera X2 (Shimadzu)), composto por uma bomba quaternária, auto amostrador e forno de colunas, acoplado a um espectrômetro de massas, (Maxis 4 G[®] UHR-TOF MS (Bruker Daltonics[®])), com nitrogênio sendo utilizado como gás de nebulização, dessolvatação e colisão, para realização de experimentos de dissociação induzida por colisão (*Collision Induced Dissociation* – CID). Os espectros de massas foram obtidos utilizando os modos de ionização positivo (ESI (+)-MS) e negativo (ESI (-)-MS).

Os parâmetros instrumentais otimizados e utilizados nas análises de MS foram: capilar: + 3500 V para o modo negativo e + 4000 V para o modo positivo, temperatura de solvatação: 200°C, *dry gas*: 9 L min⁻¹, *nebulizer*: 4 bar e *end plate offset* a 500 V.

Os cromatogramas e espectros de massas foram adquiridos, utilizando os softwares HyStar e otofControl (versões 3. 2), integrantes do pacote Compass 1. 5 for OTOF series (Bruker Daltonics[®]).

Os cromatogramas obtidos para as frações lipídica e polar foram analisados utilizando o *software* Data Analysis (Bruker Daltonics[®]) versão 4. 1, integrante do pacote Compass 1. 5 para OTOF series. Já, para a Aquisição de dados, utilizaram - se os *softwares* otofControl 3. 2 e HyStar 3. 2. Os dados de massas moleculares exatas, em função do tempo de retenção, foram compilados em um banco de dados e os resultados das diferentes replicatas biológicas foram comparados para avaliação da repetibilidade da técnica.

2.5 Análise de Dados

A análise dos dados teve início por meio da utilização da função “*Dissect function*” (ferramenta do Data Analysis): a qual encontra compostos em um cromatograma (LC-MS) mesmo quando picos estão praticamente coeluídos (sobrepostos). Foi investigado o íon mais intenso identificado para cada dissecção do cromatograma.

Para a identificação química dos compostos separados por UHPLC-MS, foi empregada a ferramenta SmartFormula™ do *software* Data Analysis, que considera a massa exata dos compostos e respectivos padrões isotópicos para determinação da fórmula química linear aproximada. A identificação das possíveis fórmulas moleculares foi realizada com base em dois principais parâmetros - (1) erro (dado em ppm) calculado em relação à massa exata medida e (2) mSigma, que está relacionado ao padrão de distribuição isotópica do composto. Estes dois parâmetros associados geram um escalonamento das fórmulas moleculares encontradas pela ferramenta, em uma escala de 0 a 100. A tolerância de erro utilizada para o *SmartFormula* foi de 2 ppm, foram ignoradas as fórmulas moleculares com mSigma superior a 20.

Para determinação da fórmula linear, foi utilizado o *software* Compound Crawler 2. 0 que realiza pesquisas em bancos de dados de acesso livre, onde foram escolhidos os bancos de dados KEGG (*Kyoto Encyclopedia of Genes and Genomes*), ChEBI (*Chemical Entities of Biological Interest*) e Metlin (*Scripps Center for Metabolomics*). Com a utilização das ferramentas citadas, buscou-se, desse modo, delinear a composição metabólica das folhas da palma de óleo.

3 RESULTADOS E DISCUSSÃO

No momento da escolha das folhas para análise química, foram necessários alguns cuidados na coleta das amostras. A amostragem é a fase onde ocorrem os erros que mais dificultam a interpretação dos resultados da análise foliar, ocasionando erros que podem comprometer todo o experimento. Com base nos estudos realizados por Knecht, Ramachandran e Narayanan (1977), na fase jovem da palma de óleo (até o segundo ou terceiro ano após o plantio), a amostragem é feita na folha + 9. A partir do terceiro ou quarto ano de plantio, a amostragem é feita sobre a folha + 17, em decorrência da posição e estágio fisiológico da mesma, sendo considerada como a que melhor expressa o estado fisiológico ideal do dendezeiro (RODRIGUES et al., 2006). O campo experimental da Embrapa Cerrados possui plantas com idade fisiológica superior a três anos, portanto, foram coletadas amostras da folha +17, as quais puderam ser identificadas por meio da filotaxia da planta (Figura 1).

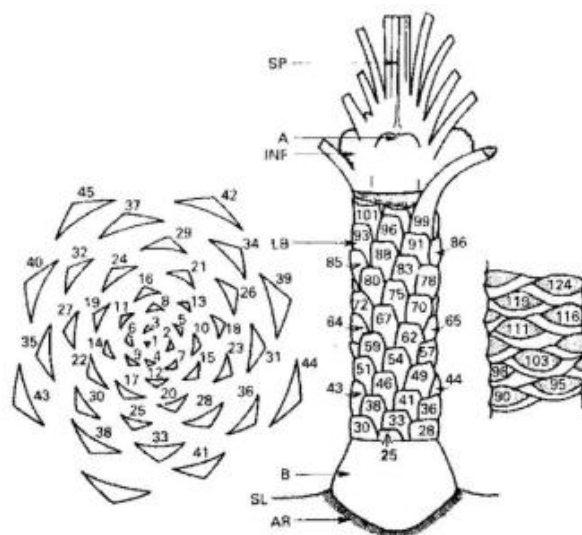


Figura 1 Filotaxia do dendzeiro, adaptado de Henry (1955).

A coleta do material vegetal por resfriamento imediato em N_2 mostrou-se eficiente e de fácil aplicação em campo. Os materiais empregados na coleta foram de simples manuseio e indicaram ser adequados para esta finalidade. O resfriamento por N_2 (criopreservação) visou à conservação de materiais vegetais sob temperatura ultrabaixa, garantindo que todos os processos metabólicos possam ser interrompidos imediatamente (ZHU et al., 2006).

No processo de extração dos metabólitos, os reagentes utilizados visam maximizar a eficiência da extração e garantir o maior ou menor número de metabólitos dependendo do caso em estudo (BARROSNETO; SCARMINIO; BRUNS, 2007). A análise otimizada de metabólitos selecionados ou classes de metabólitos polares e apolares requerem procedimentos de extração específicos em combinação com

instrumentação analítica adequada. No entanto, os protocolos de extração específicos, mais eficientes, muitas vezes, cobrem apenas um número muito limitado de compostos (MAX-PLANCK-INSTITUTE, 2014).

Para uma cobertura metabólica abrangente em plantas, é aconselhável um processo de extração simples e rápido, minimizar o tempo de preparo e análise das amostras, maximizando o rendimento de extração e detectando a maior parte das classes de metabólicos presentes na amostra (MAX-PLANCK-INSTITUTE, 2014).

Mediante a mistura de solventes, utilizada neste trabalho, foi possível extrair metabólicos com diferentes polaridades, com base no processo de extração líquido-líquido. No final da extração, duas fases foram formadas na solução, fase polar (com coloração amarela) e fase orgânica (com coloração verde) (Figura 2). A fase polar deve conter os metabólitos secundários primários, semipolares e polares, enquanto a fase orgânica deve conter lipídios.

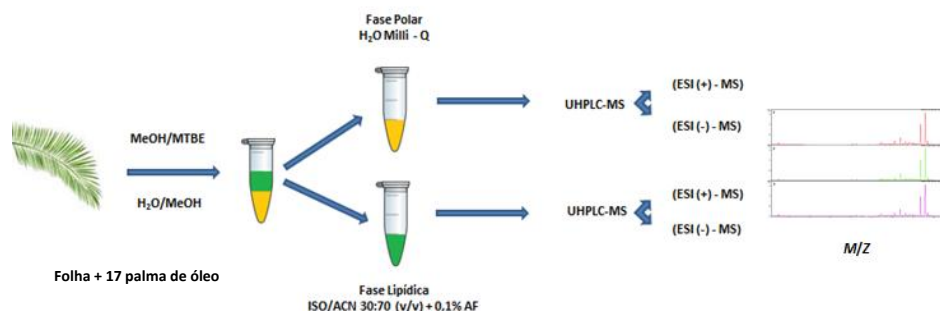


Figura 2 Figura representativa do processo de extração e detecção de metabólitos secundários de folhas de palma de óleo, adaptado de Max-Planck-Institute (2014)

A técnica de UHPLC-MS foi utilizada para a análise dos metabólitos nas frações polar e lipídica (orgânica).

Os cromatogramas obtidos para a fração polar das plantas 1 a 5, utilizando ESI (+)-MS e ESI (-)-MS estão ilustrados nas Figuras 3 e 4, respectivamente. Não houve uma diferença significativa entre os perfis dos metabólitos detectados entre as replicatas biológicas. Houve apenas uma pequena diferença na intensidade dos íons detectados utilizando o modo de ionização positivo (ESI (+)-MS (Fig. 3).

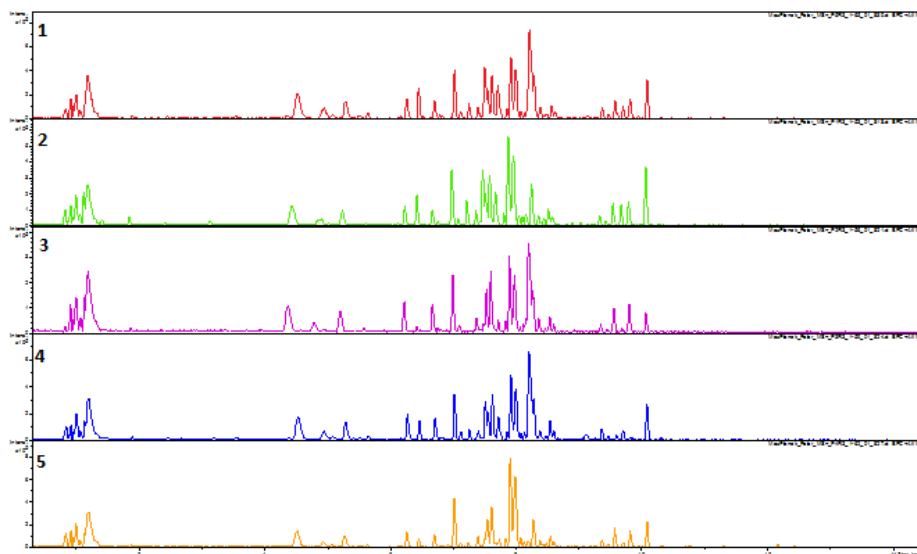


Figura 3 Cromatogramas do pico base (BPC) das frações polares das plantas 1 a 5, respectivamente, utilizando ESI (+)-MS

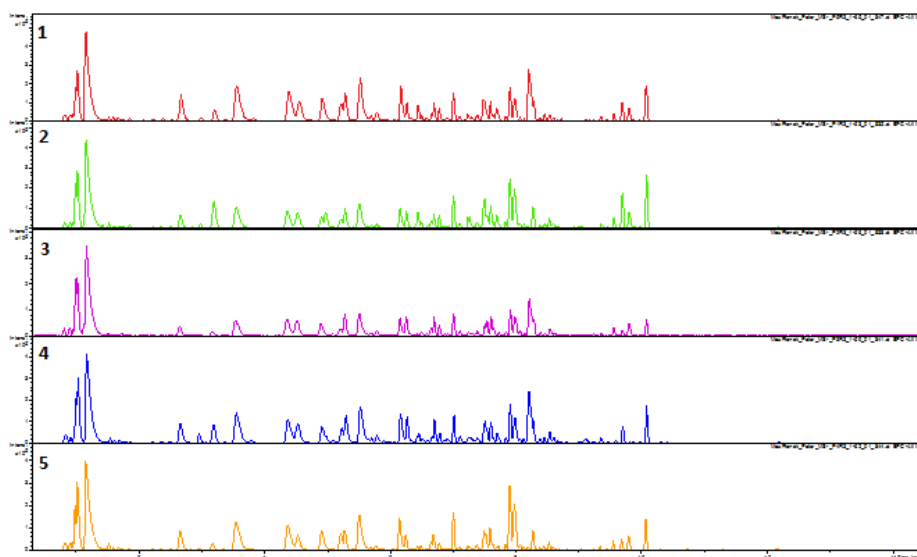


Figura 4 Cromatogramas do pico base (BPC) das frações polares das plantas 1 a 5, respectivamente, utilizando ESI (-)-MS

Os cromatogramas obtidos para a fração lipídica das plantas 1 a 5, utilizando ESI (+)-MS e ESI (-)-MS, estão ilustrados nas Figuras 5 e 6. Não foram observadas diferenças visuais significativas em relação aos perfis químicos gerados quando comparadas às repetições biológicas nos dois modos de ionização (positivo e negativo).

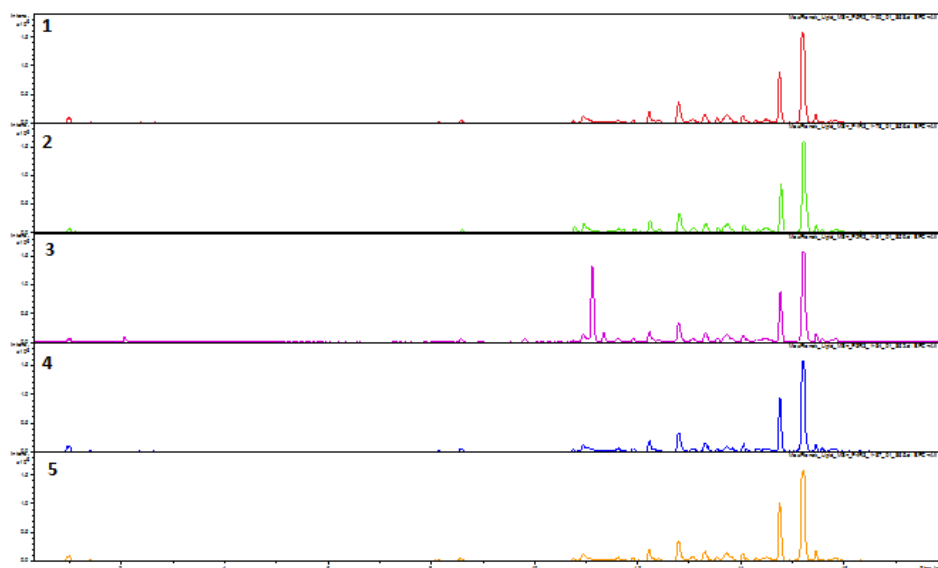


Figura 5 Cromatogramas do pico base (BPC) das frações lipídicas das plantas 1 a 5, respectivamente, utilizando ESI (+)-MS

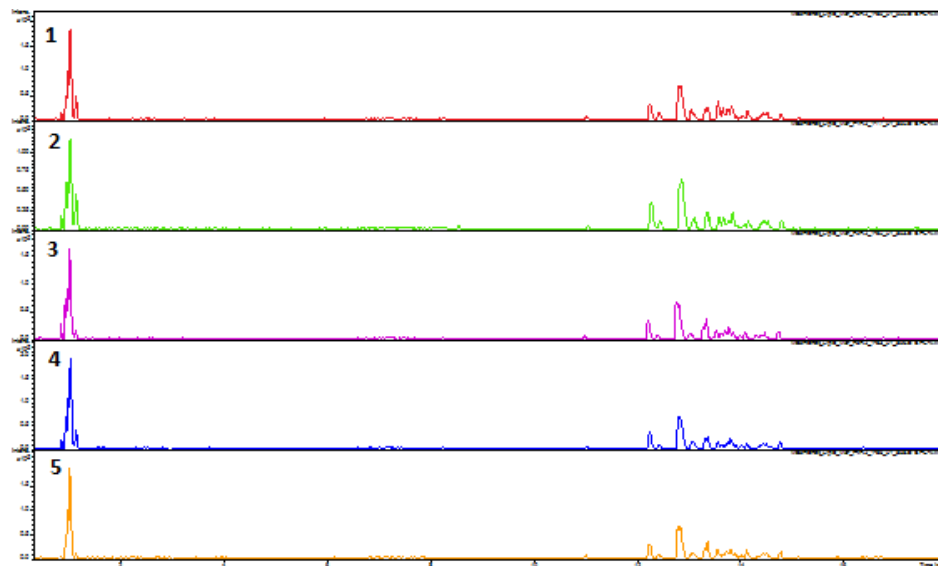


Figura 6 Cromatogramas do pico base (BPC) das frações lipídicas das plantas 1 a 5, respectivamente, utilizando ESI (-)-MS

No entanto, uma menor quantidade de compostos químicos foi detectada na fração lipídica em comparação com a fração polar, utilizando o modo positivo (ESI (+)-MS).

A planta 3 apresentou um aumento na intensidade do sinal em alguns íons, no modo negativo (Figura 6), o que pode ser resultado de um erro experimental ou de alguma anomalia da planta.

Em geral, o número de metabólitos detectados foi menor, utilizando o modo negativo do que o positivo, e as replicatas biológicas apresentaram resultados de repetibilidade.

O perfil químico dos metabólitos, obtido para as diferentes frações (polar e lipídica), utilizando os dois modos de ionização (ESI (+)-MS e ESI (-)-MS), foram distintos e complementares, detectando, assim, uma grande quantidade de metabólitos de diferentes polaridades. Alguns metabólitos com mesma razão m/z fórmulas moleculares foram observados em ambas as frações e modos de ionização (positivo e negativo) (Figura 7). Analisando os dados da Figura 7, é possível concluir que a extração utilizada para a obtenção da fração polar e a análise no modo positivo foi a condição que detectou o maior número de compostos.

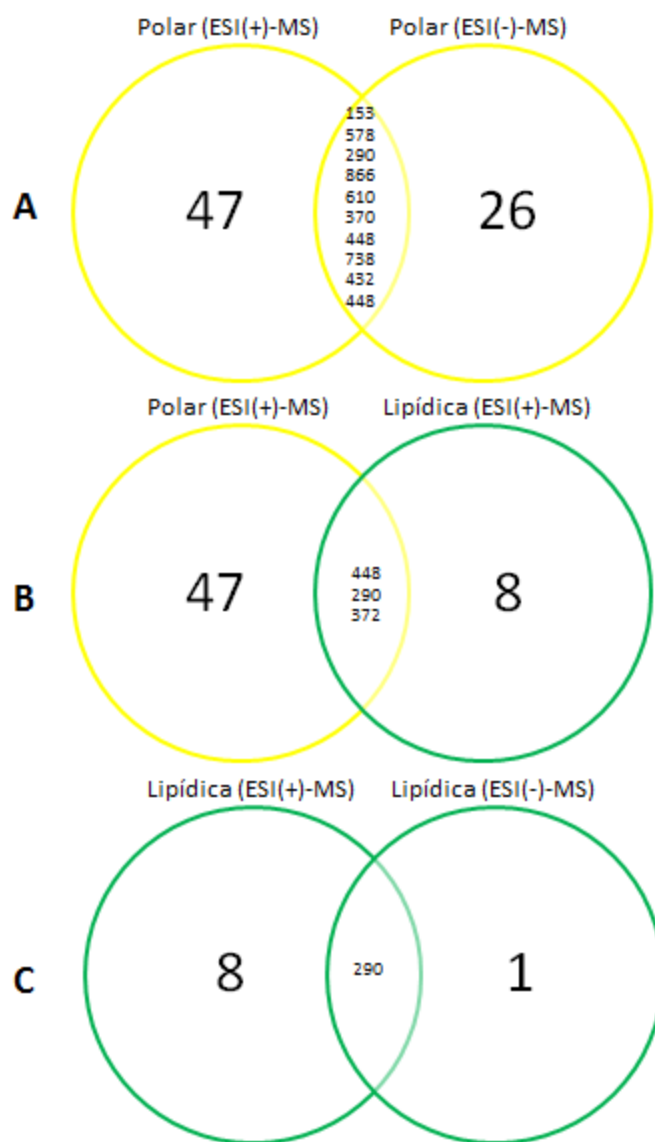


Figura 7 Número de íons detectados em cada condição de análise e íons detectados em ambas as condições. a) Fração polar nos modos de ionização positivo e negativo; b) Frações polar e lipídica no modo de ionização positivo; c) fração lipídica nos modos de ionização positivo e negativo

O experimento, para verificar a reprodutibilidade da técnica, foi realizado utilizando uma planta padrão (4) e triplicata técnica.

Os cromatogramas obtidos para as frações polares e lipídicas nos modos positivo e negativo (Figura 8 - 11) indicaram uma boa reprodutibilidade da técnica, pois apresentaram compostos químicos detectados em tempos de retenção e intensidades muito semelhantes.

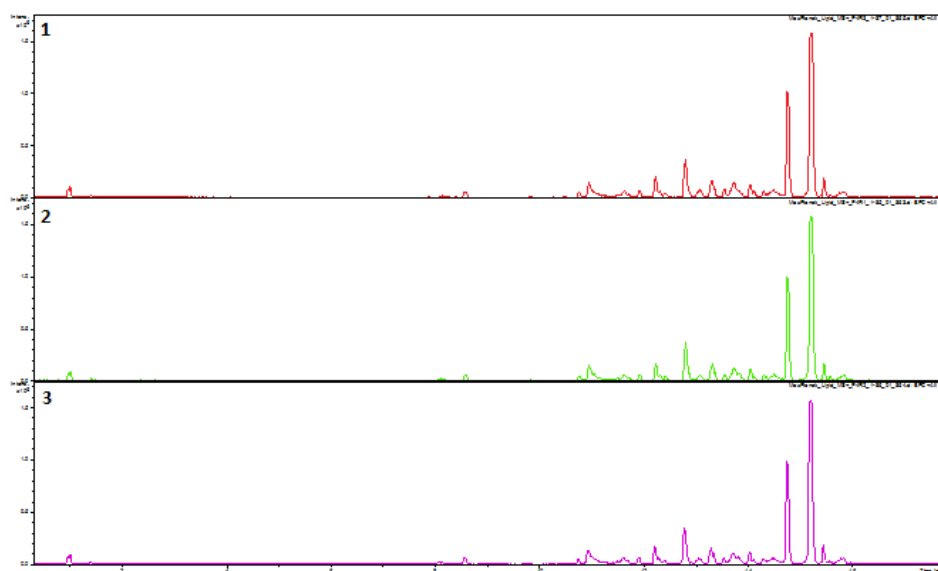


Figura 8 Cromatogramas do pico base (BPC) utilizando ESI (+)-MS para as frações lipídicas da planta 4 (replicatas técnicas)

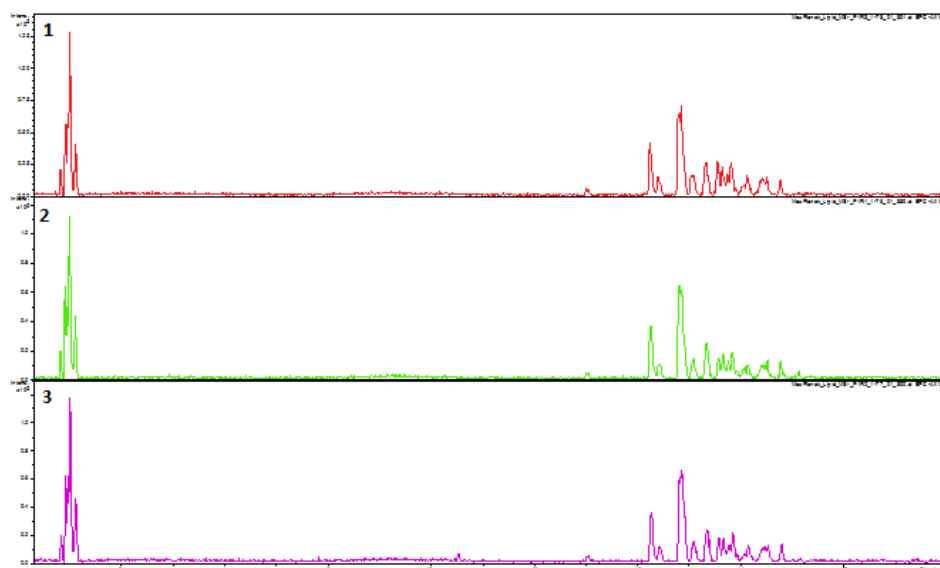


Figura 9 Cromatogramas do pico base (BPC) utilizando ESI (-)-MS para as frações lipídicas da planta 4 (replicatas técnicas)

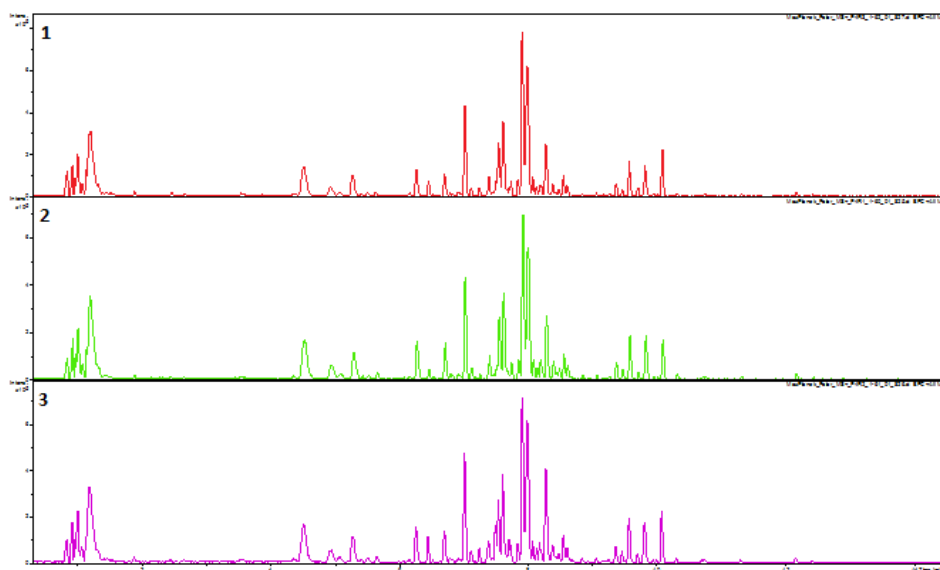


Figura 10 Cromatogramas do pico base (BPC) utilizando ESI (+)-MS para as frações polares da planta 4 (replicatas técnicas)

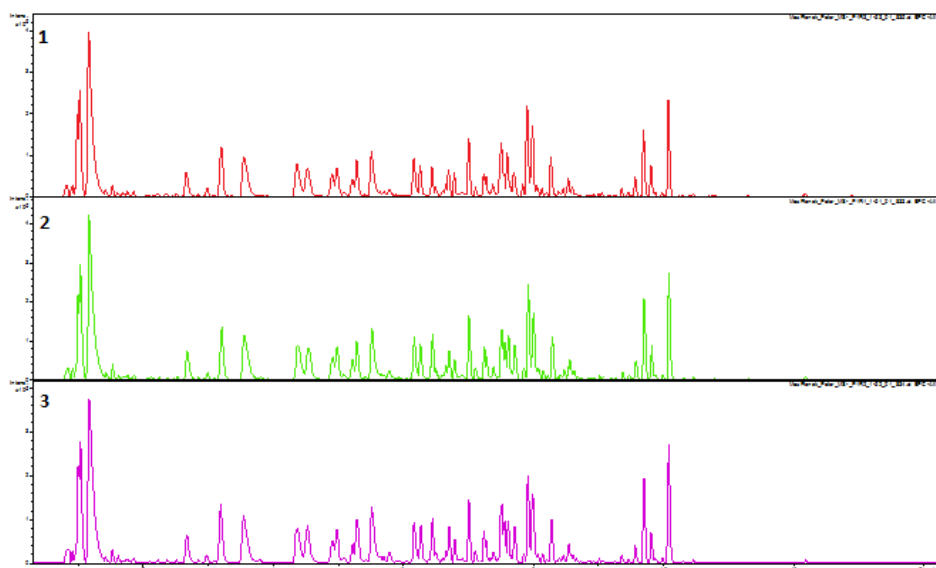


Figura 11 Cromatogramas do pico base (BPC) utilizando ESI (-)-MS para as frações polares da planta 4 (replicatas técnicas)

Todos os cromatogramas obtidos neste trabalho foram obtidos do pico base (BPC). A representação gráfica do sinal de íons representados pelo BPC demonstra os picos detectados e registrados de acordo com a maior intensidade obtida e tempo de retenção (NIESSEN; GROSS; CAPRIOLI, 2006). O BPC é similar ao cromatograma de íons totais (*Total Ion Current Chromatogram –TIC*), porém, apenas o pico mais intenso em cada espectro é analisado. Isto significa que o pico da base do cromatograma representa a intensidade do pico mais intenso em todos os pontos na análise. Os BPC's, geralmente, possuem um visual mais limpo e, portanto, são mais informativos do que cromatogramas TIC (MCNAUGHT; WILKINSON, 1997; MURRAY et al., 2013).

Cada pico do cromatograma foi selecionado e o espectro de massas referente ao tempo de retenção foi obtido. O íon mais intenso do

espectro de massas, para cada tempo de retenção, teve uma lista de fórmulas moleculares atribuídas, de acordo com a correlação entre a massa exata e padrão isotópico dos valores experimentais e teóricos (Smartformula, Bruker Daltonics).

Um exemplo de fórmula identificada para o composto 41 da planta 1, obtido para a fração lipídica, modo positivo pela ferramenta *Smartformula* pode ser visualizado na Figura 12.

SmartFormula Manually

Lower formula:

Upper formula:

Note: for m < 2000 the elements C, H, N, and O are considered implicitly.

Adducts, pos.: Collect adducts

Adducts, neg.:

Measured m/z: Tolerance: ppm Charge:

Meas. m/z	#	Ion Formula	Score	m/z	err [ppm]	Mean err [ppm]	mSigma
449.1079	1	C21H21O11	100.00	449.1078	0.1	-0.0	5.3
	2	C19H9N14O	65.39	449.1078	-0.1	-0.7	12.6
	3	C34H13N2	15.68	449.1073	1.3	-1.2	67.0

Automatically locate monoisotopic peak Maximum number of formulae:

Check rings plus double bonds Minimum: Maximum:

Electron configuration:

Filter H/C element ratio Minimum H/C: Maximum H/C:

Estimate carbon number Generate immediately

Figura 12 Fórmulas moleculares calculadas pelo *SmartFormula* para o composto 41 da planta 1, fração lipídica, modo positivo

As três fórmulas moleculares atribuídas com melhores resultados calculados (exceto aquelas com mSigma superior a 20) foram inseridas em bases de dados (ChEBI, Metlin e Kegg), utilizando a ferramenta CompoundCrawler, e possíveis compostos químicos foram identificados (Quadros 1 a 4). Na Figura 13 é demonstrado o resultado da busca gerada pelo CompoundCrawler para a fórmula $C_{21}H_{21}O_{11}$ $[M+H]^+$ calculada pelo *SmartFormula* para o composto 41 da planta 1.

Pico	T _R (min)	m/z [M+H]	Intensidade (counts)	Fórmula iônica	Massa exata [M+H]	Erro (ppm)	Fórmula empírica	Composto	Base de dados
2	0,88	175, 1192	35254	C ₆ H ₁₅ N ₄ O ₂	175, 1190	-1, 2	C ₆ H ₁₄ N ₄ O ₂	D-arginine	ChEBI: 15816 Metlin: 6924 Kegg: C00792
4	0,94	304, 1501	38259	C ₁₂ H ₂₂ N ₃ O ₆	304, 1503	0, 7	C ₁₂ H ₂₁ N ₃ O ₆	Nicotianamine	ChEBI: 17721 Metlin: 66305 Kegg: C05324
8	1,00	325, 1134	177845	C ₁₂ H ₂₁ O ₁₀	325, 1129	1, 4	C ₁₂ H ₂₀ O ₁₀	Bis-beta-D-fructofuranose 1, 2':2', 3'-dianhydride	ChEBI: 4117
9	1,06	166, 0863	59646	C ₉ H ₁₂ NO ₂	166, 0863	-0, 3	C ₉ H ₁₁ NO ₂	Phenylalanine	ChEBI: 28044
11	1,13	154, 0863	25857	C ₈ H ₁₂ NO ₂	154, 0863	-0, 3	C ₈ H ₁₁ NO ₂	Dopamine	ChEBI: 18243 Kegg: C03758

Quadro 1 - Picos detectados, tempos de retenção (Tr), razão *m/z*, intensidade de sinal, fórmula iônica, massa exata, fórmula empírica, possíveis compostos identificados e rastreamento na base de dados consultada para a fração polar modo positivo (ESI (+) - MS) para folhas + 17 de *E. guineensis*

(...continua...)

12	1, 15	185, 0079	134412	C ₇ H ₅ O ₆	185, 0081	-0, 8	C ₇ H ₄ O ₆	Chelidonic acid	Kegg: C08476
14	1, 24	180, 1018	45397	C ₁₀ H ₁₄ NO ₂	180, 1019	0, 4	C ₁₀ H ₁₃ NO ₂	N-Methyl-L-phenylalanine	ChEBI: 43980
22	1, 87	220, 1330	52956	C ₁₃ H ₁₈ NO ₂	220, 1332	0, 9	C ₁₃ H ₁₈ NO ₂	Darlingine	Kegg: C10857
23	2, 43	206, 0447	23799	C ₁₀ H ₈ NO ₄	206, 0448	0, 4	C ₁₀ H ₇ NO ₄	Xanturenic acid	ChEBI: 10072 Kegg: C02470
25	4, 50	579, 1499	123492	C ₃₀ H ₂₇ O ₁₂	579, 1497	-0, 4	C ₃₀ H ₂₆ O ₁₂	Procyanidin	ChEBI: 27589 Kegg: C10238
27	5, 27	291, 0865	91965	C ₁₅ H ₁₅ O ₆	291, 0863	0, 7	C ₁₅ H ₁₄ O ₆	(-)-epicatechin	ChEBI: 90 Metlin: 3420 Kegg: C09727
28	5, 64	867, 2138	19915	C ₄₅ H ₃₉ O ₁₈	867, 2131	-0, 9	C ₄₅ H ₃₈ O ₁₈	Procyanidin C1	ChEBI: 75643 Kegg: C17624
29	6, 25	579, 1493	120116	C ₃₀ H ₂₇ O ₁₂	579, 1497	0, 7	C ₃₀ H ₂₆ O ₁₂	Kaempferol 7-O- (2-E-p-coumaroyl- α -L-rhamnopyranoside)	ChEBI: 66134
30	6, 45	611, 1605	237582	C ₂₇ H ₃₁ O ₁₆	611, 1607	-0, 2	C ₂₇ H ₃₀ O ₁₆	Rutin	ChEBI: 28527 Kegg: C05625
31	6, 70	291, 0864	103917	C ₁₅ H ₁₅ O ₆	291, 0863	-0, 3	C ₁₅ H ₁₄ O ₆	(+)-Catechin	ChEBI: 15600 Kegg: C06562
33	6, 79	207, 0650	26791	C ₁₁ H ₁₁ O ₄	207, 0652	0, 7	C ₁₁ H ₁₀ O ₄	Eugenin	ChEBI:67374
34	6, 82	867, 2141	13504	C ₄₅ H ₃₉ O ₁₈	867, 2131	-1, 2	C ₄₅ H ₃₈ O ₁₈	ProcyanidinC1	ChEBI: 75643
35	7, 01	595, 1663	287707	C ₂₇ H ₃₁ O ₁₅	595, 1657	-0, 7	C ₂₇ H ₃₀ O ₁₅	Luteolin 7-O-neohesperidoside	ChEBI: 31788 Kegg:C12630
36	7, 12	867, 2142	28952	C ₄₅ H ₃₉ O ₁₈	867, 2131	-1, 2	C ₄₅ H ₃₈ O ₁₈	Procyanidin C2	ChEBI: 75644
37	7, 25	611, 1612	191181	C ₂₇ H ₃₁ O ₁₆	611, 1607	-0, 2	C ₂₇ H ₃₀ O ₁₆	Lucenin-2	ChEBI: 6553 Kegg: C10102

38	7, 36	581, 1504	14579	C ₂₆ H ₂₉ O ₁₅	581, 1501	0, 5	C ₂₆ H ₂₈ O ₁₅	Carlinoside	ChEBI: 3421 Metlin: 48919
39	7, 39	565, 1548	68651	C ₂₆ H ₂₉ O ₁₄	565, 1552	-0, 6	C ₂₆ H ₂₈ O ₁₄	Schaftoside	ChEBI: 9047 Kegg: C10181
40	7, 41	579, 1503	23490	C ₃₀ H ₂₇ O ₁₂	579, 1497	-1, 0	C ₃₀ H ₂₆ O ₁₂	Procyanidin B4	ChEBI: 27589 Kegg: C10238
41	7, 50	449, 1079	410212	C ₂₁ H ₂₁ O ₁₁	449, 1078	0, 1	C ₂₁ H ₂₀ O ₁₁	Orientin*	ChEBI: 7781 Metlin: 48901 Kegg: C10114
42	7, 51	207, 1379	93245	C ₁₃ H ₁₉ O ₂	207, 1380	0, 1	C ₁₃ H ₁₈ O ₂	3, 5, 6-Trimethyleugenol	ChEBI: 59096
44	7, 55	611, 1610	174928	C ₂₇ H ₃₁ O ₁₆	611, 1607	-0, 6	C ₂₇ H ₃₀ O ₁₆	Isorientin 2''-O-rhamnoside	ChEBI: 28596
45	7, 61	565, 1549	277818	C ₂₆ H ₂₉ O ₁₄	565, 1552	0, 5	C ₂₆ H ₂₈ O ₁₄	Apiin	ChEBI: 15932
47	7, 70	449, 1081	235583	C ₂₁ H ₂₁ O ₁₁	449, 1078	-0, 5	C ₂₁ H ₂₀ O ₁₁	Isorientin	ChEBI: 17965
48	7, 73	565, 1553	50003	C ₂₆ H ₂₉ O ₁₄	565, 1552	0, 0	C ₂₆ H ₂₈ O ₁₄	Isoschaftoside	ChEBI: 75589
49	7, 84	565, 1553	55311	C ₂₆ H ₂₉ O ₁₄	565, 1552	-0, 2	C ₂₆ H ₂₈ O ₁₄	Kaempferol 3-O- α -L-arabinopyranosyl-7-O- α -L-rhamnopyranoside	ChEBI: 68878
50	7, 91	595, 1662	371264	C ₂₇ H ₃₁ O ₁₅	595, 1657	0, 7	C ₂₇ H ₃₀ O ₁₅	Vitexin 2''-O- β -D-glucoside	ChEBI: 16631
51	7, 94	227, 1642	14674	C ₁₃ H ₂₃ O ₃	227, 1642	-0, 2	C ₁₃ H ₂₂ O ₃	4, 5-Dihydrovomifoliol	Metlin: 95095
52	7, 95	697, 1618	23606	C ₃₀ H ₃₃ O ₁₉	697, 1611	1, 1	C ₃₀ H ₃₂ O ₁₉	Quercetin-3-O- (2''-O- α -rhamnosyl-6''-O-malonyl)- β -glucoside	ChemSpider: 4478818
53	7, 99	595, 1667	329486	C ₂₇ H ₃₁ O ₁₅	595, 1671	-1, 5	C ₂₇ H ₃₀ O ₁₅	Isovitexin 2''-O- β -D-glucoside	ChEBI: 17379
55	8, 03	739, 2084	44178	C ₃₃ H ₃₉ O ₁₉	739, 2080	0, 5	C ₃₃ H ₃₈ O ₁₉	Malvidin 3-glucoside-5- (6''-malonylglucoside)	Metlin: 47156
56	8, 08	565, 1554	43134	C ₂₆ H ₂₉ O ₁₄	595, 1552	0, 4	C ₂₆ H ₂₈ O ₁₄	Kaempferol 3-O- β -D-apiofuranosyl-7-O- α -L-rhamnopyranoside	ChEBI: 68884

57	8, 13	739, 2083	40368	C ₃₃ H ₃₉ O ₁₉	739, 2080	0, 4	C ₃₃ H ₃₈ O ₁₉	Kaempferol 7-methyl ether 3-[3-hydroxy-3- methylglutaryl- (1->6)]- [apiosyl- (1->2)-galactoside]	Metlin: 51000
58	8, 18	480, 2226	57024	C ₂₄ H ₃₄ NO ₉	480, 2228	-0, 4	C ₂₄ H ₃₃ NO ₉	1-O-Desmethyltetraabenazine glucuronide	Metlin: 2806
59	8, 21	565, 1551	28984	C ₂₆ H ₂₉ O ₁₄	565, 1552	0, 1	C ₂₆ H ₂₈ O ₁₄	7-O-[β-D-arabinopyranosyl- (1→6)-β-D- glucosyl]apigenin	ChEBI: 75556
60	8, 28	433, 1133	211326	C ₂₁ H ₂₁ O ₁₀	433, 1129	-0, 8	C ₂₁ H ₂₀ O ₁₀	Vitexin	ChEBI: 16954
61	8, 31	609, 1815	10112	C ₂₈ H ₃₃ O ₁₅	609, 1814	0, 2	C ₂₈ H ₃₂ O ₁₅	Diosmin	Kegg: C10039 Metlin: 3676
63	8, 48	607, 1655	27502	C ₂₈ H ₃₁ O ₁₅	607, 1657	0, 4	C ₂₈ H ₃₀ O ₁₅	Citrifolinoside	ChEBI: 65637
65	8, 57	449, 1079	49902	C ₂₁ H ₂₁ O ₁₁	449, 1078	0, 1	C ₂₁ H ₂₀ O ₁₁	Quercitrin	ChEBI: 17558
67	8, 62	447, 1291	36915	C ₂₂ H ₂₃ O ₁₀	447, 1286	-1, 2	C ₂₂ H ₂₂ O ₁₀	Biochanin A 7-O-β-D- glucoside	ChEBI: 28751
75	9, 70	371. 2059	156519	C ₂₂ H ₂₂ O ₁₀	371. 2064	1, 5	C ₂₂ H ₂₁ O ₁₀	5-Megastigmen-7-yne-3, 9- diol 9-glucoside	Metlin: 88761
76	9, 81	591. 1711	129376	C ₂₈ H ₃₀ O ₁₄	591. 1708	-1, 6	C ₂₈ H ₂₉ O ₁₄	3'-O-Methylmaysin	Metlin: 49187
78	10. 08	373. 222	361950	C ₁₉ H ₃₂ O ₇	373. 2221	-0, 2	C ₁₉ H ₃₁ O ₇	myriaporone 3	ChEBI:66422

Pico	T _R (min)	m/z [M+H]	Intensidade (counts)	Fórmula iônica	Massa exata [M+H]	Erro (ppm)	Fórmula empírica	Composto	Base de dados
1	0, 83	137, 0595	19685	C ₈ H ₉ O ₂	137, 0597	-1, 7	C ₈ H ₈ O ₂	4-hydroxyphenyl)acetaldehyde	CHEBI:15621
3	0, 97	449, 1086	97149	C ₂₁ H ₂₁ O ₁₁	449, 1078	1, 8	C ₂₁ H ₂₀ O ₁₁	Orientin	ChEBI: 7781
4	1, 0	291, 0865	99463	C ₁₅ H ₁₅ O ₆	291, 0863	0, 7	C ₁₅ H ₁₄ O ₆	(-)-epicatechin	ChEBI:90
5	1, 09	209, 1533	20583	C ₁₃ H ₂₁ O ₂	209, 1536	1, 6	C ₁₃ H ₂₀ O ₂	(3R)-hydroxy-β-ionone	ChEBI: 53173
6	1, 12	373, 2217	38829	C ₁₉ H ₃₃ O ₇	373, 2221	1, 1	C ₁₉ H ₃₂ O ₇	Myriaporone 3	ChEBI: 66422
7	1, 71	171, 1494	30262	C ₉ H ₁₉ N ₂ O	171, 1492	1, 3	C ₉ H ₁₈ N ₂ O	3-Acrylamidopropyl trimethylammonium	Metlin: 34500
10	4, 07	316, 3211	13142	C ₁₉ H ₄₂ NO ₂	316, 3210	-0, 2	C ₁₉ H ₄₂ NO ₂	1-hexadecyl-2-amino-2-deoxy-sn-glycerol	ChEBI: 78541
31	12, 81	613, 4830	351272	C ₃₉ H ₆₅ O ₅	613, 4827	-0, 5	C ₃₉ H ₆₄ O ₅	1, 3-dilinolenoylglycerol	ChEBI: 75852

Quadro 2 - Picos detectados, tempos de retenção, razão *m/z*, intensidade de sinal, fórmula iônica, massa exata, fórmula empírica, possíveis compostos identificados e rastreamento na base de dados consultada para a fração lipídica modo positivo (ESI (+) - MS) para folhas + 17 de *E. guineensis*

Pico	T _R (min)	m/z [M+H]	Intensidade (counts)	Fórmula iônica	Massa exata [M-H]	Erro (ppm)	Fórmula empírica	Composto	Base de dados
3	0,97	225,0613	22908	C ₇ H ₁₃ O ₈	225,0616	1,4	C ₇ H ₁₄ O ₈	N1- (5-phospho-D-ribose)glycinamide (1-)	ChEbi: 58457
4	1,00	683,2250	215166	C ₂₄ H ₄₃ O ₂₂	683,2250	0,2	C ₂₄ H ₄₄ O ₂₂	Galabiose	Metlin: 86577
8	1,14	152,0715	29178	C ₈ H ₁₀ NO ₂	152,0717	-1,2	C ₈ H ₁₁ NO ₂	Dopamine	ChEbi: 18243
14	1,43	272,0561	14159	C ₁₄ H ₁₀ NO ₅	272,0564	1,4	C ₁₄ H ₁₁ NO ₅	Tolcapone	ChEbi: 63630
16	1,62	297,1187	20109	C ₁₁ H ₂₁ O ₉	297,1191	1,5	C ₁₁ H ₂₂ O ₉	α-L-Rhap- (1→3)-D-ribitol	ChEbi: 61090
22	2,67	355,0665	63321	C ₁₅ H ₁₅ O ₁₀	355,0671	1,8	C ₁₅ H ₁₆ O ₁₀	Caffeic acid 3-O-glucuronide	Metlin: 96063
26	4,37	355,0666	71680	C ₁₅ H ₁₅ O ₁₀	355,0671	-1,4	C ₁₅ H ₁₆ O ₁₀	Caffeoyl C1-glucuronide	Metlin: 96067
27	4,53	577,1345	83912	C ₃₀ H ₂₅ O ₁₂	577,1351	-1,1	C ₃₀ H ₂₆ O ₁₂	Procyanidin B1	ChEbi: 75633
28	4,91	355,0664	51373	C ₁₅ H ₁₅ O ₁₀	355,0671	1,8	C ₁₅ H ₁₆ O ₁₀	Caffeic acid 3-O-glucuronide	Metlin: 96063
29	4,98	345,1187	75023	C ₁₅ H ₂₁ O ₉	345,1191	1,1	C ₁₅ H ₂₂ O ₉	3,5-dimethoxy-4-hydroxybenzyl alcohol-4-O-β-D-glucopyranoside	ChEbi:68338
30	5,22	355,0665	41285	C ₁₅ H ₁₅ O ₁₀	355,0671	1,6	C ₁₅ H ₁₆ O ₁₀	Caffeoyl C1-glucuronide	Metlin: 96067
31	5,29	289,0711	97878	C ₁₅ H ₁₃ O ₆	289,0718	2,0	C ₁₅ H ₁₄ O ₆	(-)-epicatechin	ChEbi:90
33	5,52	355,0665	124439	C ₁₅ H ₁₅ O ₁₀	355,0671	0,5	C ₁₅ H ₁₆ O ₁₀	Caffeoyl C1-glucuronide	Metlin: 96067

Quadro 3 - Picos detectados, tempos de retenção, razão *m/z*, intensidade de sinal, fórmula iônica, massa exata, fórmula empírica, possíveis compostos identificados e rastreamento na base de dados consultada para a fração polar modo negativo (ESI (-) - MS) para folhas + 17 de *E. guineensis*

(...continua...)

34	5, 65	865. 1980	16261	C ₄₅ H ₃₇ O ₁₈	865. 1985	1, 2	C ₄₅ H ₃₈ O ₁₈	arecatannin B1	ChEBI:75642
37	6, 17	355. 0665	124439	C ₁₅ H ₁₅ O ₁₀	355. 0671	0, 5	C ₁₅ H ₁₆ O ₁₀	Caffeoyl C1-glucuronide	Metlin: 96067
40	6, 45	609, 1452	103763	C ₂₇ H ₂₉ O ₁₆	609, 1461	-1, 4	C ₂₇ H ₃₀ O ₁₆	Rutin	ChEBI: 28527 Kegg: C05625
41	6, 51	369, 0823	25709	C ₁₆ H ₁₇ O ₁₀	369, 0827	1, 1	C ₁₆ H ₁₈ O ₁₀	5-Hydroxy-6-methoxycoumarin 7-glucoside	Metlin: 94324
43	6, 62	355, 1031	13174	C ₁₆ H ₁₉ O ₉	355, 1035	-0, 9	C ₁₆ H ₂₀ O ₉	6-O- (trans-feruloyl)-D-galactopyranose	ChEBI: 62282
45	6, 71	289, 0712	84452	C ₁₅ H ₁₃ O ₆	289, 0718	1, 8	C ₁₅ H ₁₄ O ₆	(-)-epicatechin	ChEBI:90
57	7, 50	447, 0928	125721	C ₂₁ H ₁₉ O ₁₁	447, 0933	1, 7	C ₂₁ H ₂₀ O ₁₁	Orientin	ChEBI: 7781
67	8, 04	737, 1930	17761	C ₃₃ H ₃₇ O ₁₉	737, 1935	0, 7	C ₃₃ H ₃₈ O ₁₉	Malvidin 3-glucoside-5-(6"-malonylglucoside)	Metilin: 47156
72	8, 28	431, 0981	99941	C ₂₁ H ₁₉ O ₁₀	431, 0984	0, 5	C ₂₁ H ₂₀ O ₁₀	Vitexin	ChEBI: 16954
73	8, 40	463, 0880	12331	C ₂₁ H ₁₉ O ₁₂	463, 0882	-0, 5	C ₂₁ H ₂₀ O ₁₂	Gossypetin 8-rhamnoside	ChEBI: 28086 Kegg: C10050
76	8, 58	447, 0929	17526	C ₂₁ H ₁₉ O ₁₁	447, 0933	0, 8	C ₂₁ H ₂₀ O ₁₁	isoorientin	ChEBI:17965
86	9, 70	415, 1965	194305	C ₂₀ H ₃₁ O ₉	415, 1974	2, 0	C ₂₀ H ₃₂ O ₉	Ethyl 7 epi 12-hydroxyjasmonate glucoside	Metilin: 91484
90	10, 09	417, 2126	267501	C ₂₀ H ₃₃ O ₉	417, 2130	-0, 9	C ₂₀ H ₃₄ O ₉	Auranofin	ChEBI: 2922

Pico	T _R (min)	m/z [M+H]	Intensidade (counts)	Fórmula iônica	Massa exata [M- H]	Erro (ppm)	Fórmula empírica	Composto	Base de dados
3	1,01	289,0716	114269	C ₁₅ H ₁₃ O ₆	289,0718	0,4	C ₁₅ H ₁₄ O ₆	(-)-epicatechin	ChEBI:90

Quadro 4 - Picos detectados, tempos de retenção, razão *m/z*, intensidade de sinal, fórmula iônica, massa exata, fórmula empírica, possíveis compostos identificados e rastreamento na base de dados consultada para a fração lipídica modo negativo (ESI (-) - MS) para folhas + 17 de *E. guineensis*

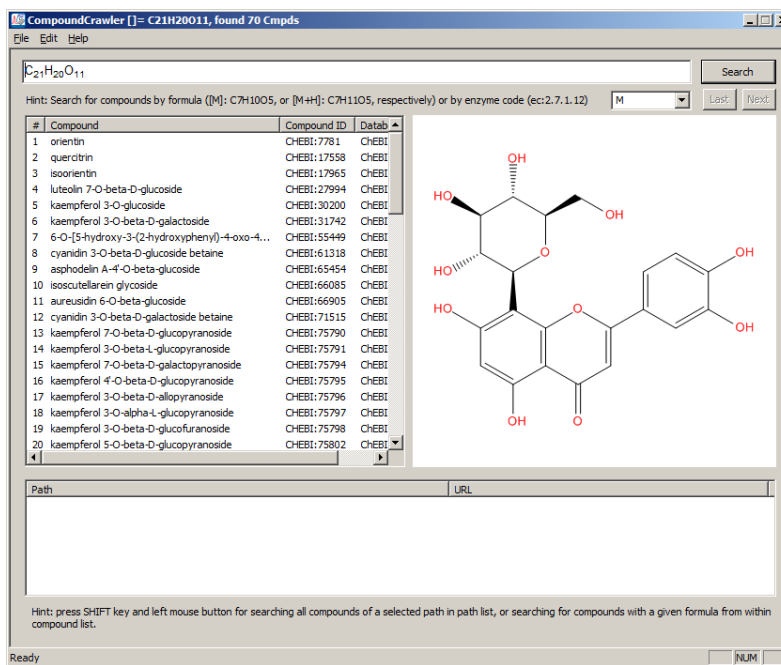


Figura 13 Possíveis compostos identificados nos bancos de base ChEBI, Metlin e Kegg para a fórmula molecular $C_{21}H_{21}O_{11}$ (modo positivo) do composto 41 da planta 1

Noor et al. (2012) identificaram 27 compostos em folhas de palma de óleo, em trabalho publicado na literatura. Neste trabalho 7 fórmulas químicas identificadas são semelhantes ao trabalho publicado. Estas fórmulas e razões m/z estão descritas na Tabela 1.

Tabela 1 Razões m/z (corrigidas) e fórmulas moleculares (empíricas) de compostos identificados neste trabalho

m/z literatura	m/z encontrada	Fórmula literatura	Fórmula encontrada
610, 1520	610, 1607	$C_{27}H_{30}O_{16}$	$C_{27}H_{30}O_{16}$
594, 1517	594, 1567	$C_{27}H_{30}O_{15}$	$C_{27}H_{30}O_{15}$
594, 1552	594, 1567	$C_{27}H_{30}O_{15}$	$C_{27}H_{30}O_{15}$
564, 1409	564, 1552	$C_{26}H_{28}O_{14}$	$C_{26}H_{28}O_{14}$
448, 2883	448, 1078	$C_{21}H_{20}O_{11}$	$C_{21}H_{20}O_{11}$
738, 1991	738, 2080	$C_{33}H_{38}O_{19}$	$C_{33}H_{38}O_{19}$
432, 0979	432, 1129	$C_{21}H_{20}O_{10}$	$C_{21}H_{20}O_{10}$

Fonte: (NOOR et al., 2012).

4 CONCLUSÃO

Uma plataforma analítica baseada em metabolômica para identificar metabólitos em folhas de *Elaeis guineenses* Jacq. (palma de óleo) foi desenvolvida neste trabalho. O método de extração mostrou-se simples e eficaz, uma vez que o processo de extração utilizado foi líquido-líquido, considerando uma mistura de solventes com diferentes polaridades, realizado em uma única etapa. A técnica analítica, UHPLC-MS, apresentou boa reprodutibilidade nas análises, uma vez que os cromatogramas demonstraram um padrão similar, indiferente das replicatas biológicas e replicatas técnicas analisadas.

Ao final deste trabalho foram identificados 47 possíveis compostos presentes na fração polar em modo positivo (ESI (+)-MS), 26 na fração polar em modo negativo (ESI (-)-MS), 8 na fração lipídica em modo positivo (ESI (+)-MS) e 1 na fração lipídica em modo negativo (ESI (-)-MS). No entanto, será necessário o uso da técnica de espectrometria de massas *tandem* (MS/MS) para a elucidação estrutural dos metabólitos identificados e propostos neste trabalho, para comprovação e validação dos mesmos.

Extraction and analysis of secondary metabolites from *Elaeis guineensis* Jacq leaves by liquid chromatography coupled to ultrahigh performance mass spectrometry

ABSTRACT

Although oil palm (*Elaeis guineensis*) stand out among the other oil varieties for its high oil production capacity, there are few secondary metabolites identified and described in the literature on *Elaeis* genus. Just like there is a small number of papers focusing the use of metabolomics in the study of this genus. The aim of this study was to identify secondary metabolites in leaves of *Elaeis guineensis* Jacq (palm oil) using metabolomics approach. Therefore we developed a metabolites analysis protocol based on the extraction and detection of metabolites using advanced analytical techniques based on liquid chromatography coupled to ultra high performance mass spectrometry (UHPLC-MS). The results enabled the identification of metabolites from the detection of their respective m/z reasons, which generate the molecular formulas which have been associated with the compounds described in the literature. The metabolites present in both fractions analyzed (polar and lipid) should be validated by tandem mass spectrometry experiments (MS/MS).

Keywords: *Elaeis guineensis*. Palm oil. Metabolomics. Secondary metabolites. Mass spectrometry. UHPLC-MS. Electrospray time-of-flight.

REFERÊNCIAS

BARCELOS, E. et al. Genetic diversity and relationship in American and African oil palm as revealed by RFLP and AFLP molecular markers. **Pesquisa Agropecuária Brasileira**, Brasília, v. 37, n. 8, p. 1105-1114, ago. 2002.

BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. **Como fazer experimentos: pesquisa e desenvolvimento na ciência e indústria**. 3. ed. Campinas: UNICAMP, 2007. 414 p.

CUNHA, R. N. V. da et al. Domesticação e melhoramento do caiaué. In: BORÉM, A.; LOPES, M. T. G.; CLEMENT, C. R. (Ed.). **Domesticação e melhoramento: espécies amazônicas**. Viçosa, MG: UFV, 2009. p. 275-296.

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS. **Production/Crops**. Rome, 2012. Disponível em: <<http://faostat.fao.org/site/636/DesktopDefault.aspx?PageID=636#ancor>>. Acesso em: 13 jan. 2014.

HALL, R. D. Plant metabolomics: from holistic hope, to hype, to hot topic. **New Phytologist**, Cambridge, v. 169, n. 3, p. 453-468, 2006.

HENRY, P. Leaf growth morphology in *Elaeis*. **Revue Générale de Botanique**, Paris, v. 62, n. 2, p. 66-77, févr. 1955.

KALTNER, F. J. et al. **Viabilidade técnica e econômica de produção de ésteres de óleo de palma, para utilização como substituto de óleo diesel, na Amazônia**. Belém: EMBRAPA Amazônia Oriental, 2004. 54 p. (Documentos, 193).

KNECHT, J. C. X.; RAMACHANDRAN, R.; NARAYANAN, R. Variation des teneurs en éléments nutritifs de la feuille avec l'âge des palmiers dans les prélèvements foliaires. **Oléagineux**, Paris, v. 32, n. 4, p. 139-147, 1977.

MAX-PLANCK-INSTITUTE OF MOLECULAR PLANT PHYSIOLOGY. **Analyzing proteins and metabolites**: all-in-one extraction. Disponível em: <http://www.mpimp-golm.mpg.de/10116/Analyzing_proteins_and_metabolites_All-in-One_Extraction?print=yes>. Acesso em: 10 mar. 2014.

MCNAUGHT, A. D.; WILKINSON, A. **Compendium of chemical terminology**. 2nded. Oxford: Blackwell Scientific, 1997. 450 p.

MURRAY, K. K. et al. Definitions of terms relating to mass spectrometry. **Pure and Applied Chemistry**, Oxford, v. 1, n. 7, p. 1515-1609, June 2013.

MYA, T. A.; SOMVONG, V.; CHATCHAWAN, J. Efficient and rapid plant regeneration of oil palm zygotic embryos cv. 'Tenera' through somatic embryogenesis. **Acta Physiologiae Plantarum**, Copenhagen, v. 33, n. 1, p. 123-128, 2010.

NIESSEN, W. M. A.; GROSS, M. L.; CAPRIOLI, R. M. **The encyclopedia of mass spectrometry**. Oxford: Elsevier, 2006. 1088 p.

NOOR, I. et al. Characterization of Apigenin and Luteolin Derivatives from oil palm (*Elaeis guineensis* Jacq.) leaf using LC-ESI-MS/MS. **Journal of Agricultural and Food Chemistry**, London, v. 45, n. 60, p. 11201-11210, Nov. 2012.

PINTO, Â. C. et al. Produtos naturais: atualidades, desafios e perspectivas. **Química Nova**, São Paulo, v. 25, n. 1, p. 45-61, 2002. Suplemento.

RODRIGUES, M. R. L. et al. **Avaliação do estado nutricional do dendezeiro: análise foliar**. Manaus: EMBRAPA Amazônia Ocidental, 2006. 9p. (Circular Técnica, 26).

ZHU, G. Y. et al. Change in sugar, sterol and fatty acid composition in banana meristems caused by sucrose-induced acclimation and its effects on cryopreservation. **Physiologia Plantarum**, Copenhagen, v. 128, n. 1, p. 80-94, Sept. 2006.

CONSIDERAÇÕES FINAIS

Com os estudos de prospecção de genes tecido específico (folha), pode-se dar um passo importante no que se diz respeito à utilização do *draft* do genoma do caiaué como uma fonte de genes de interesse. Com a realização deste trabalho, foi possível a identificação de um gene presente em seu genoma, o qual teve seu perfil de expressão caracterizado com a utilização da técnica de qPCR. O prosseguimento do trabalho em questão se dará com a identificação de regiões promotoras do gene selecionado, assim como de novos possíveis genes que possam vir a ter sua presença confirmada no genoma do caiaué. Os genes em questão serão testados em planta modelo (*Setaria* spp.) para validação e comprovação de sua expressão específica. Desta maneira, o trabalho será complementado de maneira satisfatória, abrangendo as etapas de prospecção em bancos de dados, identificação e comprovação de presença no genoma do caiaué, comparação de expressão relativa, identificação de regiões promotoras de tecido específico e teste e validação em planta modelo.

No que se diz respeito ao estudo da metabolômica em folhas fisiologicamente maduras de palma de óleo, as técnicas utilizadas para extração de metabólitos secundários se mostraram eficientes e de fácil reprodutibilidade, vários metabólitos foram detectados nas duas frações obtidas (polar e lipídica) assim como nos dois modos de aquisição, positivo (ESI (+)-MS) e negativo (ESI (-)-MS). A utilização da cromatografia líquida de Ultra Alta Eficiência acoplada à espectrometria de massas mostrou-se capaz de produzir resultados confiáveis e com boa reprodutibilidade, os quais puderam ser comprovados por meio dos cromatogramas obtidos. Em estudos futuros, com os dados gerados, será possível a identificação precisa dos metabólitos presentes na folha da palma de óleo mediante a técnica de espectrometria de massas *tandem* (MS/MS).

A união destas duas áreas de trabalho (prospecção de genes e metabolômica) possui importância no que se diz respeito aos estudos com as duas espécies do gênero *Elaeis* spp., pois hoje o governo federal possui diversos incentivos para programas de melhoramento e estudos com as respectivas espécies (*Elaeis guineensis* e *Elaeis oleifera*). O estabelecimento de uma rede de pesquisa no Brasil, que abranja a palma de óleo e o caiaué e que una dados moleculares e informações sobre compostos de interesse, presentes no metabolismo dessas espécies pode, cada vez mais, agregar valor e gerar mais incentivos para a expansão das culturas citadas em nosso país.