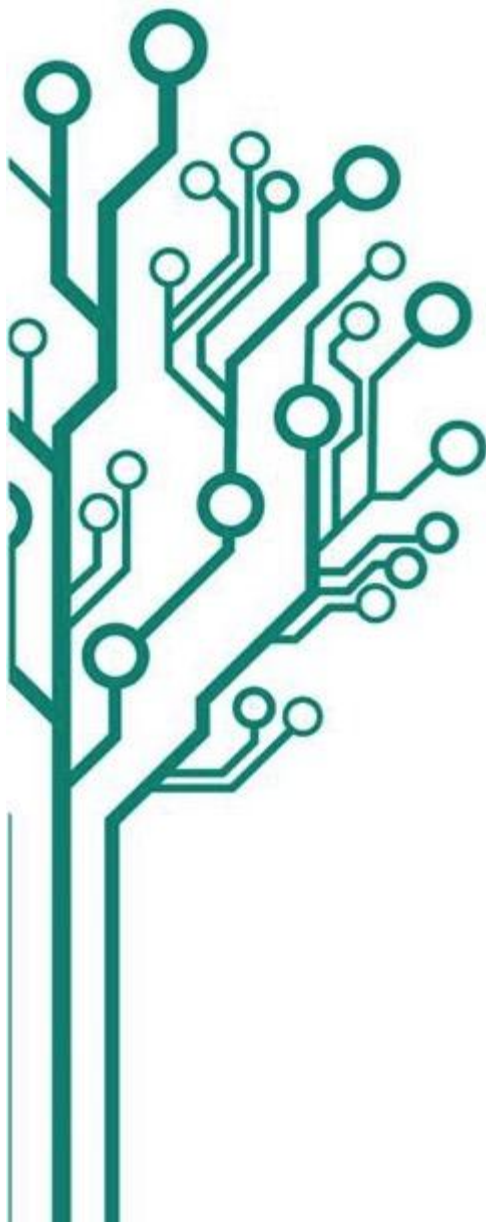


**Mestrado Profissional**  
**Uso Sustentável de Recursos Naturais em Regiões Tropicais**



**BASE MOLECULAR DA  
VARIAÇÃO NATURAL DA  
BIOSSÍNTESE DO COMPOSTO  
VOLÁTIL METIL-CHAVICOL  
NAS ESPÉCIES DE DENDÊ**

**TATIANA MAIA DE OLIVEIRA**

**BELÉM-PA  
2015**

**TATIANA MAIA DE OLIVEIRA**

**BASE MOLECULAR DA VARIAÇÃO NATURAL DA  
BIOSSÍNTESE DO COMPOSTO VOLÁTIL METIL-  
CHAVICOL NAS ESPÉCIES DE DENDÊ**

Dissertação apresentada ao Programa de Mestrado Profissional  
em Uso Sustentável de Recursos Naturais em Regiões  
Tropicais do Instituto Tecnológico Vale (ITV).

Orientador: Sergei Kushnir  
Co-orientador: Nelson Carvalho

BELÉM– BRASIL  
2015

**Dados Internacionais de Catalogação na Publicação (CIP)**  
**Biblioteca do ITV – Belém-PA**

048 Oliveira, Tatiana Maia de

Base molecular da variação natural da biossíntese do composto volátil metil-chavicol nas espécies de Dendê/ Tatiana Maia de Oliveira -- Belém-PA, 2015.

100 f.: il.

Dissertação (mestrado) -- Instituto Tecnológico Vale, 2015.

Orientador(a): Sergei Kushnir

Co-orientador: Nelson Carvalho

1. Gene. 2. Polinizador. 3. Híbrido. Base molecular da variação natural da biossíntese do composto volátil metil-chavicol nas espécies de Dendê

CDD 581.35

**TATIANA MAIA DE OLIVEIRA**

**BASE MOLECULAR DA VARIAÇÃO NATURAL DA  
BIOSSÍNTESE DO COMPOSTO VOLÁTIL METIL-  
CHAVICOL NAS ESPÉCIES DE DENDÊ**

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Ciências Ambientais e Desenvolvimento Sustentável do Programa de Mestrado Profissional em Uso Sustentável de Recursos Naturais em Regiões Tropicais do Instituto Tecnológico Vale (ITV).

Data da aprovação: 02/04/2015

Banca examinadora:

---

Dr. Sergei Kushnir  
Orientador – Instituto Tecnológico Vale (ITV)

---

Dr. Nelson Carvalho  
Co-orientador – Instituto Tecnológico Vale (ITV)

---

Dr. Rafael Valadares  
Instituto Tecnológico Vale (ITV)

---

Dr. Marcelo Murad Magalhães  
Embrapa Amazônia Oriental (EMBRAPA)

BELÉM– BRASIL  
2015

“O homem é ao mesmo tempo criatura e criador do meio ambiente, que lhe dá sustento físico e lhe oferece a oportunidade de desenvolver-se intelectual, moral, social e espiritualmente. A longa e difícil evolução da raça humana no planeta levou-a a um estágio em que, com o rápido progresso da ciência e da tecnologia, conquistou o poder de transformar de inúmeras maneiras e em escala sem precedentes o meio ambiente. Natural ou criado pelo homem, é o meio ambiente essencial para o bem estar e para gozo dos direitos fundamentais, até mesmo o direito à própria vida.”

(Declaração de Estocolmo sobre o Ambiente Humano, 1972)

## Dedico

Aos meus pais Dirceu e Selma e minha irmã Silvia pelo amor incondicional, incentivo, carinho e apoio dedicados durante a minha formação profissional e pessoal.

## AGRADECIMENTO

Primeiramente, a Deus por ter me concedido à oportunidade de realizar mais um sonho, abençoando-me e trilhando os meus caminhos.

Aos meus amados pais e minha irmã Silvinha pelos ensinamentos dos verdadeiros valores da vida, pelos conselhos, pela compreensão e por todo incentivo que sempre me deram nas horas mais difíceis e por estarem sempre presentes em todos os momentos da minha vida. Amo muito vocês!

Aos meus familiares que sempre incentivaram e me apoiaram mesmo estando longe, em especial a minha querida tia Cléia e tio Jurandir e meus primos Marcelo, Mario Sylvio e sua esposa Tâmara e sua filha Karol. Meus sinceros agradecimentos pelo carinho, apoio, torcida e força. Amo vocês!

Minha eterna gratidão ao professor pesquisador Dr. Sergei Kushnir, pela orientação, confiança, paciência e valiosos ensinamentos.

Meus sinceros agradecimentos ao professor Dr. Nelson Carvalho pela co-orientação, confiança e ensinamentos.

Aos pesquisadores Dr<sup>a</sup> Aleksandra Skiryecz, Dr<sup>a</sup> Elena Bablychuk, Dr. Rafael Valadares, Dr<sup>a</sup> Hivana Dall'Agnol e Dr. Ivan Quevas pelos conselhos, incentivos e apoio os meus singelos agradecimentos.

Aos técnicos de laboratório Msc. Cinthia Miléo e Manoel Lopes pela amizade, pelo agradável convívio e pelo constante auxílio na realização dos experimentos.

A todos os professores pesquisadores pelos ensinamentos, sugestões e incentivos ao longo do curso.

A todos os amigos do mestrado em especial a querida Rose, Alexandre, Breno e João pela amizade, companheirismo, ajuda na realização dos trabalhos, horas de descontração e de boas risadas.

Aos amigos da república pela amizade, pelo convívio, pelos momentos de pura descontração e por serem meus guias turísticos.

À querida amiga Mannu, pela amizade, conselhos, incentivo, torcida, horas e mais horas de bate-papo e aos demais colegas que fiz ao longo desse período.

Ao Instituto Tecnológico Vale (ITV) pela realização profissional e pessoal concebida.

À Fundação de Amparo e Desenvolvimento da Pesquisa (FADESP) pela concessão da bolsa de estudos.

À Biopalma pela parceria e fornecimento dos materiais biológicos para a realização deste trabalho.

Aos funcionários do Programa de Pós-Graduação do Mestrado Profissional em Uso Sustentável de Recursos Naturais em Regiões Tropicais.

A todos que fazem e fizeram parte da minha vida, pois tenho certeza de que viemos a esse mundo para sermos felizes e realizarmos os nossos sonhos.

**Muito obrigada!!!!**



## LISTA DE FIGURAS

Figura 1: Fruto do dendê mostrando de fora para dentro, o epicarpo, mesocarpo, endocarpo e o kernel. ....	9
Figura 2: Plantas adultas de <i>Elaeis guineensis</i> , <i>Elaeis oleifera</i> e híbridos interespecíficos (OxG), cultivados no município de Manicoré, estado de Manaus. ....	12
Figura 3: <i>Elaeis guineensis</i> apresentando Amarelecimento Fatal (AF). Plantas mantidas no banco de germoplasma da Biopalma, situada no município de Moju, estado do Pará.....	15
Figura 4: Síntese de metil-chavicol (MeC) nas plantas de Manjeriçã e Petúnia.....	20
Figura 5: Inflorescências femininas e masculinas de <i>E. guineensis</i> liberam compostos voláteis que são atrativos de insetos polinizadores.....	23
Figura 6: Fluxograma das análises <i>in silico</i> e experimental da biossíntese de metilchavicol (MeC) em <i>Elaeis guineensis</i> e <i>Elaeis oleifera</i> .....	28
Figura 7: Modelagem por homologia dos genes chavicol <i>O</i> -metiltransferase ( <i>COMT</i> ) e eugenol sintase ( <i>EGS</i> ). ....	35
Figura 8: Árvore filogenética Neighbour-joining das sequências proteicas de <i>E. guineensis</i> são similares as sequências do manjeriçã para o gene <i>COMT</i> .....	36
Figura 9: Árvore filogenética Neighbour-joining das sequências proteicas de <i>E. guineensis</i> são similares as sequências do manjeriçã para o gene <i>EGS</i> .....	36
Figura 10: Alinhamento das sequências do gene chavicol <i>O</i> - metiltransferase ( <i>COMT2</i> ) no éxon 1 das espécies Africana (EG) e Americana (EO).. ....	38
Figura 11: Deficiência na expressão do gene estrutural para a síntese de metilchavicol (MeC) em <i>E. oleifera</i> .. ....	39
Figura 12: Desenho experimental dos <i>primers</i> AS16 e AS 17 no éxon 1 do gene <i>COMT2</i> em <i>E. oleifera</i> (EO).. ....	40
Figura 13: Eletroforese em gel de agarose 1% do DNA genômico de <i>E. guineensis</i> , <i>E. oleifera</i> e <i>O. basilicum</i> .....	41

Figura 14: Visualização em gel de Agarose 2% do DNA genômico de <i>E. guineensis</i> e <i>E. oleifera</i> .....	42
Figura 15: Imagem de cromatograma dos alinhamentos dos fragmentos gênicos amplificados pela PCR para os <i>primers</i> AS16 e AS17 do gene <i>COMT2</i> no éxon 1 em <i>E. oleifera</i> (EO)....	44
Figura 16: Deficiência de mutação no gene estrutural para a síntese de metilchavicol (MeC) em <i>E. oleifera</i> . ....	45
Figura 17: Deficiência no fluxo metabólico para a síntese de metilchavicol (MeC) em <i>E. oleifera</i> .....	46
Figura 18: Diferença nas sequências de DNA entre os acessos de bancos de dados genômicos. ....	47

## LISTA DE TABELAS

Tabela 1: Volume de reagentes e concentração final na reação de PCR.....	31
Tabela 2: Sequências de <i>primers</i> para a realização da PCR.....	32
Tabela 3: Combinação de <i>primers</i> , número de acesso, tamanho esperado em pares de base (pb) utilizados na amplificação dos genes de <i>Elaeis</i> spp. ....	32

## LISTA DE SIGLAS E ABREVIATURAS

aa- Aminoácido

AF- Amarelecimento Fatal

BLAST- *Basic Local Alignment Search Tool*

COV – Composto Orgânico Volátil

COMT - Chavicol *O*-metiltransferase

DNA- Ácido desoxirribonucleico

dNTPs – Deoxinucleotídeos Trifosfato

ddNTPs – Dideoxinucleotídeos Trifosfato

EDTA-- Ácido etilenodiamino tetra-acético

EGS- Eugenol sintetase

EOMT – Eugenol *O*- metiltransferase

MeC- Metil-chavicol

NCBI- *National Center for Biotechnology Information*

OMT- *O*-metiltransferase

PBS- Tampão fosfato-salino

PCR – Reação em Cadeia da Polimerase

Phe- Aminoácido fenilalanina

RNA –Ácido ribonucleico

Rnase- - Ribonuclease

TBE - Tampão de separação eletroforética de ácidos nucléicos ou proteínas

## SUMÁRIO

RESUMO.....	i
ABSTRACT .....	iii
<b>1. INTRODUÇÃO.....</b>	<b>1</b>
1.1 Objetivos.....	3
1.1.1 Objetivo geral.....	3
1.1.2 Objetivos específicos .....	3
1.2 Problema e hipótese .....	4
1.2.1 Problemas: Questões biológicas.....	4
1.2.2 Hipóteses .....	5
1.2.2.1 Fatores de regulação.....	5
1.2.2.2 Mutações naturais de genes estruturais .....	6
1.2.2.3 Deficiência do fluxo metabólico.....	7
<b>2. REFERENCIAL TEÓRICO E FUNDAMENTAÇÃO CIENTÍFICA .....</b>	<b>8</b>
2.1 Descrição e caracterização botânica das espécies de óleo de palma .....	8
2.1.1 <i>Elaeis guineensis</i> Jacq.....	8
2.1.2 <i>Elaeis oleifera</i> (H.B.K.) Cortés.....	10
2.1.3 Híbridos interespecíficos (OxG) .....	11
2.2 O dendezeiro como a maior fonte de óleo vegetal e industrial.....	12
2.2.1 Amarelecimento Fatal (AF) .....	14
2.3 Biossíntese de Compostos Orgânicos Voláteis (COVs) em óleo de palma....	16
2.3.1 A função dos metabólicos das plantas na ecosfera .....	16
2.3.1.1 O papel dos metabólitos voláteis nas interações entre planta e meio ambiente	17
2.3.1.2 Biossíntese do composto volátil metil-chavicol nas plantas .....	18
2.3.2 Polinizadores .....	22
2.3.2.1 O papel dos polinizadores no óleo de palma.....	22
2.4 Aplicação da Bioinformática na biologia molecular .....	25
<b>3. MATERIAL E MÉTODOS .....</b>	<b>28</b>
3.1.1 Análise de bioinformática na biossíntese de metil-chavicol.....	28

3.1.2	Material Vegetal.....	29
3.1.3	Extração do DNA genômico .....	29
3.1.4	Análise qualitativa e quantitativa do DNA .....	30
3.1.5	Reação em Cadeia da Polimerase (PCR) .....	30
3.1.6	Purificação dos fragmentos de DNA amplificados pela PCR.....	33
3.1.7	Análise dos fragmentos de DNA por sequenciamento de Sanger.....	33
<b>4.</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>34</b>
4.1.1	Identificação de genes ortólogos <i>COMT</i> e <i>EGS</i> no genoma do dendê.....	34
4.1.2	Mutações nos genes <i>EGS</i> e <i>COMT</i> na espécie Americana .....	36
4.1.3	Análise do DNA genômico .....	40
4.1.4	Análises da reação de sequenciamento .....	43
<b>5.</b>	<b>CONCLUSÃO .....</b>	<b>48</b>
<b>6.</b>	<b>REFERÊNCIAS.....</b>	<b>49</b>
	<b>ANÉXOS.....</b>	<b>60</b>

## RESUMO

A palmeira de óleo Africana ou Dendê (*Elaeis guineensis*) é a principal cultura produtora de óleos vegetais no mundo. O rendimento máximo da extração de óleo em seus cultivares depende da polinização efetiva pelos insetos. O composto orgânico volátil metilchavicol (MeC) desempenha um papel fundamental na atração dos insetos polinizadores às flores femininas e masculinas na espécie Africana. O maior obstáculo para o desenvolvimento da indústria de óleo de palma na América Latina está relacionado a uma doença, denominada de Amarelecimento fatal (AF), cujo agente etiológico é desconhecido. A espécie de palma Americana ou Caiué (*Elaeis oleifera*), endêmica das Américas Central e do Sul, é uma variedade do dendezeiro que apresenta resistência ao AF, assim como os híbridos interespecíficos que resultam dos cruzamentos entre Oleifera e Guineenses (O x G). Entretanto, as plantas híbridas O x G apresentam uma baixa taxa de polinização. Esta condição impõe a necessidade de polinização manual assistida o que ocasiona o aumento nos custos de produção (mão de obra) nestes palmares. Em parte, esta baixa taxa de polinização pode ser explicada pela ausência de biossíntese do MeC na espécie *E. oleifera*, ocasionando níveis significativamente baixos do composto MeC nos híbridos O x G. O objetivo deste estudo foi investigar os mecanismos moleculares que distinguem as espécies Africana e Americana quanto a biossíntese do MeC. Através da análise comparativa em bases de dados disponíveis na internet (MPOB e NCBI) foi identificado no genoma da espécie Africana, um conjunto de genes que codificam enzimas que poderiam catalisar a biossíntese do MeC em duas etapas distintas e relacionadas a partir do substrato acetato *p*-cumaril. As análises de alinhamento das sequências gênicas mostraram que os genes chavicol *O*-metiltransferase (*COMT2* e *COMT4*) da palma Americana apresentam mutações do tipo *frame-shift* nas sequências que podem impedir a produção de enzimas da via biossintética do MeC. Neste contexto, sugere-se que a espécie Americana perdeu a capacidade de sintetizar o composto MeC, devido a ocorrência de pressões seletivas negativas desconhecidas ao longo do tempo. Para validar os resultados das análises *in silico*, foram sequenciados fragmentos de 900 e 465 pares de base (pb) do gene *COMT2* e *COMT4* que pudessem confirmar a variação natural presente entre as duas espécies de dendezeiro. O sequenciamento destes fragmentos na espécie *E. oleifera* demonstrou não ser válida a hipótese de que a ocorrência de mutações estaria associada a alterações na biossíntese do MeC. Dessa maneira, os resultados obtidos sugerem a ocorrência de outros fatores responsáveis pela diferença de produção do MeC entre

as espécies de dendê. Este estudo estabelece uma base experimental para futuras investigações referentes aos desafios na compreensão dos fatores relacionados à produção diferenciada deste volátil atraente de polinizadores em *E. guineensis*.

**Palavras-chave:** gene, metil-chavicol, polinizador, sequenciamento de DNA, óleo de palma



## ABSTRACT

African oil palm, *Elaeis guineensis* is the major oil producing crop in the world. The realization of the maximal yield potential of existing cultivars depends on effective pollination by the insects. It is believed that the organic volatile molecule methyl chavicol (MeC) plays a central role in guiding the pollinating insects towards female and male flowers of the African oil palm species. The obstacle for the oil palm industry development in Latin America is a disease fatal yellowing, whose etiologic agent is unknown. The American oil palm, *Elaeis oleifera* that is a wild palm endemic to Central and North of the South America is resistant for fatal yellowing, as well as the interspecific oil palm hybrids known as OxG hybrids. The shortcoming of the OXG palms is low fertility however, it needs the implementation of the manual assisted pollination, increasing the labor intensity of the crop. In part, low fertility can be explained by the fact that *E. oleifera* does not synthesize MeC, whereas OxG have significantly lower MeC production. Thus, the goal of this research project was to investigate the molecular mechanisms that distinguish African and American oil palm species in regard of MeC biosynthetic ability. Using similarity searches in (MPOB and NCBI) we identified in the African oil palm genome a set of candidate genes that encode enzymes that could catalyze the two-step biosynthesis of MeC starting from *p*-coumoyl acetate. Sequence alignment analysis showed that the orthologous genes chavicol *O*-methyltransferase (*COMT2* and *COMT4*) in American oil palm carry frame-shift causing mutations that prevent enzyme production. In that scenario, the American oil palm species lost the ability to synthesize MeC due to as such unknown negative selective pressure in a wild. To validate the results from *in silico* analyzes were sequenced fragments of 900 and 465 base pairs (bp) of the *COMT2* and *COMT4* gene which could confirm this natural variation between the two species. Sequencing of these fragments in *E. oleifera* demonstrated isn't true the hypothesis that the occurrence of mutations would be associated with alterations in the biosynthesis of the MeC. Thus, the molecular analysis obtained suggest the occurrence of other factors responsible for MeC's production of difference between the species of palm. This study establishes an experimental basis for future research related to challenges in understanding the factors related to the differentiated production of volatile attractive to pollinators in *E. guineensis*.

**Key-words:** gene, pollinators, methyl chavicol, DNA sequencing, oil palm

## 1. INTRODUÇÃO

O óleo de palma é considerado a cultura oleaginosa mais produtiva no mundo (Singh et al. 2013), e vem impulsionando a atenção do governo e instituições de pesquisa devido a crescente demanda por óleos vegetais para a indústria alimentícia, cosmética e na produção de combustíveis (Becker, 2010; Homma et al. 2000; Veiga e Furia, 2000).

O gênero *Elaeis* (tribo Cocoeae), família das Arecaceae (Dransfield et al. 2008) é uma das mais antigas famílias datadas no período Cretáceo (Purseglove 1972 in Singh et al. 2013). O gênero consiste em duas espécies, *Elaeis guineensis* Jacq. e *Elaeis oleifera* (Kunth) Cortês, (Zeven 1965 in Singh et al. 2013).

A palmeira Africana ou Dendê (*E. guineensis*) é encontrada no Sul da África e representa uma das mais importantes fontes de óleos vegetais no mundo (Barcelos et al. 2002). A espécie Americana ou Caiuê (*E. oleifera*) é distribuída na América Central e América do Sul e vem sendo considerada como recurso genético promissor para programas de melhoramento devido a resistência de algumas anomalias na América Latina. Nos últimos anos, híbridos interespecíficos (OxG), obtidos do cruzamento entre a palma Americana e a palma Africana, vêm aumentando substancialmente na América Latina por causa de sua aparente resistência à doenças (Torres et al. 2010).

O obstáculo para o desenvolvimento do cultivo do dendezeiro na América Latina está relacionado a uma patologia devastadora, denominada de Amarelecimento Fatal (AF). Tal anomalia vem dizimando os plantios de dendê, forçando produtores e melhoristas à buscarem fontes de resistência nas espécies Americanas que são resistentes ao AF, assim como os híbridos interespecíficos.

As interações entre polinizadores e o dendê são necessárias para uma elevada produção de frutos, conseqüentemente, para a produção de óleo. Numerosas espécies de insetos estão envolvidas na polinização do dendezeiro, mas os principais agentes polinizadores são besouros (Coleoptera, Curculionidae), entre os quais se destaca o *Elaeidobius kamerunicus*, conhecidos popularmente por gorgulhos (Lajis et al. 1985).

O principal composto orgânico volátil (COV), presente nas flores de dendê foi identificado como 1-metoxi-4 (2-propenil) benzeno, também conhecido como metil-chavicol (MeC), desempenha um papel fundamental na atração de polinizadores, principalmente, os gorgulhos. Todavia, os híbridos OxG apresentam deficiência com a baixa polinização,

resultando a necessidade de polinização manual assistida (Corley e Tinker, 2008). A baixa polinização nos híbridos pode ser explicado, devido às inflorescências de *E. oleífera*, não apresentar liberação de MeC suficiente para atração de insetos polinizadores. O conhecimento e o avanço nos estudos de biologia molecular em compreender a biossíntese de MeC nas espécies de palma será um desafio para estudos futuros.

## 1.1 OBJETIVOS

### 1.1.1 **Objetivo geral**

Investigar os mecanismos moleculares que distinguem as espécies Africana e Americana quanto a biossíntese de metil-chavicol.

### 1.1.2 **Objetivos específicos**

- Efetuar análises de bioinformática do genoma do dendezeiro para identificar genes ortólogos envolvidos na biossíntese de metil-chavicol.
- Realizar sequenciamento *de novo* dos genes envolvidos na biossíntese do metil-chavicol.

## 1.2 PROBLEMA E HIPÓTESE

### 1.2.1 Problemas: Questões biológicas

A eficiência na polinização da palmeira Africana (*E. guineensis*) depende de uma co-evolução entre planta e inseto polinizador, como o *Elaeidobius kamerunicus* (Corley e Tinker 2008). Acredita-se que o metil-chavicol (MeC) é o principal composto volátil, emitido das flores masculinas e femininas da palma Africana, responsáveis por atrair espécies de gorgulhos (Lajis et al. 1985; Hussein et al. 1989). Todavia, a palma Americana (*E. oleifera*), não emite quantidade significativa de MeC para atração de insetos polinizadores (Gomes, 2011). Deste modo, quais seriam as mudanças evolutivas no genoma das palmeiras, que poderiam esclarecer a divergência natural na capacidade de ambas às espécies sintetizarem o MeC?

Os híbridos interespecíficos (OxG), provenientes do cruzamento de ambos os progenitores não produzem o MeC nos níveis que caracterizam *E. guineensis* (*pisifera*) (Gomes, 2011). Contudo, são menos atraentes para os gorgulhos (Tan, 1985). Desse modo, a deficiência na síntese de MeC na espécie Americana, apresenta-se como semi-dominante nos híbridos.

As análises de padrões biogeográficos presentes nas palmeiras, da tribo Cocoeae, gênero *Elaeis* spp., sugere que a tribo é de origem Gondwana vicariance, e sua diversificação, pode ter coincido com o rompimento continental a milhares de anos (Hahn, 2002). Estudos palinológicos, registraram um fóssil de pólen no período Mioceno e nas camadas do delta no Nilo semelhante ao pólen de *E. guineensis* (Zeven, 1964).

Zeven (1964), sugeriu que *Elaeis* spp., teve origem na África e estendeu-se para a América do Sul, quando os continentes foram unidos. Ambas as ideias da África e da América do Sul são consistentes uma vez que as primeiras espécies de *Elaeis* sofreram especiação geográfica com o rompimento de Gondwana, há cerca de 60 milhões de anos atrás (Zeven, 1964; Hahn, 2002). É interessante especular que o ancestral de *Elaeis* em Gondwana (*Elaeis gondwaneae*), utilizava o MeC como atrativo de insetos polinizadores, esta consideração é relevante para as questões biológicas abordadas neste estudo. Se *E.*

*gondwaneae* sintetizava MeC e dependia de gorgulhos para a polinização, contudo, as espécies Americanas perderam essa característica de atrair insetos polinizadores.

A co-evolução entre planta e polinizador é pensado como uma interação ecológica/benéfica, que constituem um dos principais fatores da mudança evolutiva (Armbruster, 2014). Na América do Sul, como poderia haver alguma seleção ecológica que substituiria os benefícios da dependência da polinização pelos gorgulhos? Uma alternativa seria que *E. gondwaneae* não produzia o composto MeC como atrativos de polinizadores (característica dominante na espécie Africana), assim, quais seriam as alterações moleculares na evolução desta espécie?

### 1.2.2 Hipóteses

Nesta seção, serão abordados exemplos relevantes de outros grupos taxonômicos de plantas nas causas moleculares para a variação natural.

#### 1.2.2.1 Fatores de regulação

As flores *Clarkia breweri* emitem uma mistura de compostos voláteis, como linalol, enquanto que outra espécie estreitamente relacionada, como a *C. Concinna* não emite esse composto. Notavelmente, os genomas de ambas as espécies contêm um gene para a síntese de linalol, mas o gene é altamente expresso em flores de *C. breweri*, o que não se verifica em *C. concinna*.

A emissão de linalol a partir de flores de *C. breweri* (incluindo outros compostos de aroma, como eugenol e isoeugenol, não presentes em *C. concinna*) é atrativo de polinizadores, entretanto, estes insetos não visitam regularmente as flores de *C. concinna* (Koeduka et al. 2009a).

A razão para a falta de expressão do gene linalol sintase em *C. concinna*, poderia está relacionado a uma mutação, no gene que codifica o fator de regulação na transcrição. Por exemplo, ANTHOCYANIN2 (AN2), definido como MYB é o principal fator de transcrição que determina a cor da flor de *Petunia integrifolia* e *Petunia axillaris*, respectivamente.

As análises de sequenciamento na variação de AN2 em *P. axillaris* apresentou perda da função dos alelos. Os resultados mostraram que o AN2 é um dos principais determinantes na atração de polinizadores. Entretanto, alterações genéticas em um único gene, causa uma grande alteração na biologia da polinização, podendo surgir um novo tipo de polinizador e, conseqüentemente, envolver um número limitado de genes relacionados à síntese de linalol (Hoballah et al. 2007). A divergência neste estudo será referida como uma “deficiência na expressão estrutural do gene”.

### 1.2.2.2 Mutações naturais de genes estruturais

A co-evolução entre o odor e o polinizador na América do Sul em plantas de Petúnia, fornece um exemplo de divergência genética. As flores de *P. axillaris* são brancas e emitem uma mistura de compostos voláteis, como os benzenoides e fenil-propenos, e são polinizadas por uma diversidade de espécies de esfingídeos. Algumas variações na composição de *P. axillaris* foram notadas e, em particular, foram observadas que as flores da subespécie *P. a. parodii*, não emitem eugenol e isoeugenol, enquanto, que as flores de *P. a. axillaris* liberam estes compostos voláteis (Hoballah et al. 2007; Kondo et al. 2007).

Ambos, os compostos isoeugenol e eugenol em *P. hybrida* são biossintetizados a partir de acetato cumaril em reações catalisadas pela isoeugenol sintase (PhIGS1) e eugenol sintase (PhEGS1), respectivamente, através do intermediário meteto-quinona. Análises moleculares de clonagem e sequenciamento do gene mostraram que *P. axillaris subsp. parodii*, apresenta um gene funcional eugenol sintase (*EGS*) que expressa nas flores, mas o seu gene isoeugenol sintase (*IGS*), contém uma mutação, tornando-o inativo (Koeduka et al. 2009a).

A mutação em genes estruturais poderá explicar a variação natural e as diferenças na emissão de voláteis, entre diferentes acessos de uma mesma espécie e/ou espécies estreitamente relacionadas, que será denominado neste estudo como “mutação do gene estrutural”.

### 1.2.2.3 Deficiência do fluxo metabólico

A espécie *Petúnia axillaris subsp. paradii*, apresentam genes que são expressos nas pétalas das flores, tais como o eugenol sintase (EGS) e eugenol *O*- metiltransferase (EOMT). No entanto, as espécies não sintetizam eugenol e metil-eugenol.

Segundo Kondo et al. (2007), o acetato cumaril (substrato do eugenol sintase) é convertido por uma enzima ainda desconhecida, definida como enzima "X" para acetato diidrocuraril. Assim, a enzima (X) compete com EGS para o mesmo substrato, na medida em que as espécies são incapazes de produzir eugenol. Nesse mecanismo, a via do fenil-propeno é encaminhado para a síntese do acetato diidrocuraril, por meio do eugenol.

O acetato cumaril e acetato coniferílico são os precursores na biossíntese de fenil-propenos para a síntese do metil-chavicol, eugenol e lignina, respectivamente. A deficiência do fluxo metabólico, devido a competição da biossíntese de metil-chavicol com a lignina, poderia explicar a deficiência da síntese do MeC na palma Americana.



## 2. REFERENCIAL TEÓRICO E FUNDAMENTAÇÃO CIENTÍFICA

### 2.1 DESCRIÇÃO E CARACTERIZAÇÃO BOTÂNICA DAS ESPÉCIES DE ÓLEO DE PALMA

#### 2.1.1 *Elaeis guineensis* Jacq.

O dendezeiro (*Elaeis guineensis*) Jacq. é uma palmeira originária da Costa Ocidental da África (Golfo da Guiné), sendo encontrada em povoamentos subespontâneos desde o Senegal até a Angola. No Brasil foi introduzido no século XVII pelos escravos e adaptou-se bem ao clima tropical úmido (Figura 2a) (Trindade, 2005; Müller e Andrade 2010).

O sistema radicular do dendezeiro é do tipo fasciculado, formado a partir do bulbo localizado na base do estipe, de onde saem raízes primárias verticais e horizontais. São muito rígidas e não possuem pêlos absorventes (Conceição e Muller 2000; Cunha et al. 2009).

As folhas possuem bases pecioladas que permanecem aderidas à estipe por no mínimo 12 anos (Akani, 2007). São plantas monóicas e suas inflorescências são protegidas por duas brácteas fibrosas (espatas) são espádices compostas por espigas formadas na axila de cada folha. As inflorescências masculinas e femininas são produzidas em ciclos unissexuais. A razão sexual entre o número de inflorescências femininas e o número total de inflorescências pode variar de 3 a 20 anualmente (Adam, 2005).

O dendê apresenta três formas de frutos denominados de *dura*, *pisifera* e *tenera* (Abdullah et al. 2011). O híbrido intraespecífico *tenera* é do tipo *dura* (♀) e *pisifera* (♂) e altamente preferido para o cultivo comercial (Bhore e Shan 2011). Os frutos da *tenera* apresentam maior volume de mesocarpo e maior produção de óleo em relação a *dura*, devido a espessura fina do endocarpo. A maior parte da produção de óleo de palma comercial em todo o sudeste asiático depende do cultivo de híbridos *tenera* (Bhore e Shan 2011).

Os frutos amadurecem entre 5 a 6 meses após a fecundação das flores femininas. São do tipo drupa, séssil, formato ovóide, medindo de 2 a 5 cm de comprimento e pesando de 3 a 30 gramas (Conceição e Muller 2000). Quando cortado transversalmente, o fruto pode ser

distinguido pelo epicarpo (casca do fruto), mesocarpo (polpa), contendo o óleo de palma, e a matriz fibrosa consiste no endocarpo (shell) e o núcleo (kernel/ amêndoa) contém um óleo diferente, assemelha-se ao óleo de coco (Figura 1) (FAO, 2002).



Figura 1: Fruto do dendê mostrando de fora para dentro, o epicarpo, mesocarpo, endocarpo e o kernel.

As amêndoas apresentam forma ovóide e ocupam toda a cavidade do endocarpo composta por um tegumento fino de coloração amarela que se torna negro quando seco. O embrião fica embutido em uma cavidade do endosperma logo abaixo do poro germinativo (Müller e Andrade 2010).

Essa cultura é de suma importância uma vez que apresenta grande potencial na redução de mão-de-obra no campo. Todavia, a sua produção vem sendo afetada por algumas doenças, como o amarelecimento fatal (AF), responsável por milhares de mortes de plantas em países produtores, como Brasil, Colômbia, Equador, Suriname, Costa Rica, Nicarágua e Panamá (Franqueville 2001).

### 2.1.2 *Elaeis oleifera* (H.B.K.) Cortés

Endêmico da zona tropical úmida da América Latina ocorre, em populações espontâneas desde o sul do México até as áreas amazônicas do Brasil e Colômbia (Corley e Tinker 2008). Na Amazônia é muito frequente a ocorrência da espécie sobre manchas de solo denominadas “terra preta de índio” com característica de alta fertilidade em virtude dos elevados teores de fósforo, cálcio e matéria orgânica (Barcelos et al. 2002; Cunha et al. 2009).

A morfologia de *E. oleifera* é caracterizada por apresentar um tronco do tipo estipe, cilíndrico maciço, caracterizado pelo crescimento anual entre 5 a 10 cm, até o máximo de 5 metros de altura. Por volta dos 15 anos de idade, o estipe curva-se sobre o solo permanecendo a parte terminal ereta, mantendo a coroa foliar a 2-3 metros de altura, num estágio denominado de procumbência (Figura 2b) (Corley e Tinker 2008). A plântula nos primeiros dias após a germinação desenvolve uma raiz primária que é envolvida por finas raízes laterais que vão aumentando de espessura e comprimento. As raízes de até um metro de comprimento são formadas em toda a extensão do tronco procumbente (Corley e Tinker 2008).

Nas folhas os folíolos são dispostos em um mesmo plano entre a raque, conferindo aspecto crespo às folhas uma das principais diferenças entre a *E. oleifera* e *E. guineensis* (Corley e Tinker 2008).

São espécies monóicas, ou seja, as inflorescências masculinas e femininas são produzidas na mesma planta. Tanto as inflorescências masculinas como as femininas encontram-se envoltas por duas espatas externas que se rompem expondo a raque floral (Cunha et al. 2009). A raque da inflorescência feminina mede de 15 a 20 cm de comprimento e é nela que estão dispostas em formato espiral as ráquias com 4 a 9 cm de comprimento. A raque masculina mede de 10 a 15 cm de comprimento e suporta entre 40 a 200 ráquias com comprimento variando entre 8 e 14 cm (Corley e Tinker 2008).

O fruto é do tipo drupa de tamanho menor do que os frutos de *E. guineensis*. Os frutos quando maduros podem apresentar coloração variando de amarela a alaranjada, pesando entre 1,7 e 13,0 gramas (Corley e Tinker 2008).

Essa espécie é considerada como recurso genético promissor para programas de melhoramento de óleo de palma para o provimento à tolerância e algumas anomalias na América Latina como melhoria da qualidade do óleo e crescimento lento (Barcelos et al. 2002). No entanto, as suas características de desenvolvimento vegetativo e reprodutivo são desconhecidas, e sua fenologia não tem sido descrita (Hormaza et al. 2012).

### 2.1.3 Híbridos interespecíficos (OxG)

Os híbridos interespecíficos (OxG) obtidos do cruzamento entre a palma Americana e a palma Africana, vêm aumentando na América Latina, devido a sua aparente resistência à doenças (Torres et al. 2010) (Figura 2c).

Os híbridos OxG geralmente apresentam características intermediárias aos dois genitores, (*E. guineensis* e *E. oleifera*) revelando assim a ação predominante aditiva dos genes. Dentre essas características, enfatiza-se a menor taxa de crescimento de tronco, o que aumentaria o período de exploração econômica dos plantios; produção de óleo com teor de ácidos graxos insaturados superiores ao do *tenera*, sendo, portanto, mais líquido nas condições ambientais e apresentando-se mais apropriado ao consumo alimentar. Além, de aspectos relacionados à resistência a doenças e pragas, às quais o dendezeiro é suscetível (Barcelos, 2001).

No entanto, apresentam problemas com a baixa polinização, o que pode ser devido às divisões celulares anormais que são frequentes durante a formação do grão de pólen, assim como a baixa germinação do pólen (Hardon e Tan, 1969). Segundo Corley e Tinker (2008), problemas com o desenvolvimento do grão de pólen, podem está relacionado com as proporções de flores femininas solitárias desenvolverem-se nas inflorescências masculinas (andromórficas); por problemas na frutificação, devido as inflorescências serem menos atraentes para insetos polinizadores ou por apresentarem baixos níveis ou falta de compostos voláteis adequados presentes nas inflorescências.

O cultivo de híbridos é uma importante alternativa para a produtividade na exploração econômica e principalmente por serem resistentes e/ou tolerantes a patologias e pragas como o Amarelecimento Fatal (AF).



Figura 2: Plantas adultas de *Elaeis guineensis*, *Elaeis oleifera* e híbridos interespecíficos (OxG), cultivados no município de Manicoré, estado de Manaus. (A) *E. guineensis* Jacq. apresentando folhas superiores cortadas, devido à propagação clonal; (B) *E. oleifera* expondo tronco procumbente; (C) Híbrido (OxG). Imagens cedidas da coleção de germoplasma da EMBRAPA, Manaus.

## 2.2 O DENDEZEIRO COMO A MAIOR FONTE DE ÓLEO VEGETAL E INDUSTRIAL

O comércio internacional de óleo de palma começou na virada do século XIX e tornou-se a principal mercadoria de navios negreiros após a abolição da escravatura. O estabelecimento do comércio de óleo de palma na África Ocidental foi o principal resultado da Revolução Industrial na Europa. A população começou a levar a sério a questão do saneamento básico, o que gerou um aumento na procura por sabão, consequentemente, a demanda por óleos vegetais (FAO, 2002).

A Nigéria foi o maior exportador até 1934, quando o país foi superado pela Malásia, entretanto, a África liderou o mundo na produção e exportação ao longo da primeira metade do século XX. Os maiores produtores de óleo de palma são a Indonésia e Malásia (FAO, 2002).

O dendê é a oleaginosa com maior produção por unidade de área e existem mais de 12 milhões de hectares plantados de óleo de palma no mundo, com uma produção de mais de 45 milhões de toneladas, o que deverá aumentar substancialmente devido ao crescimento da procura do óleo pela indústria alimentícia e o uso de biodiesel na Europa e alguns países da América do Sul (Fedepalma, 2010). No entanto, apenas 5% da área total mundial é utilizada para o cultivo, embora responda por 33% da produção mundial de óleo vegetal (Singh et al. 2013).

No mercado mundial o óleo de palma é considerado um líder de mercado, constituindo cerca de 35% do comércio mundial de óleos e gorduras (Bhore e Shan, 2011). No Brasil, os cultivos de óleo de palma estão mais concentrados no estado do Pará, com cerca de 80% da cultura no país, e plantios adequadamente manejados produzem em média de 4 a 6 toneladas (t) de óleo/ha/ano (Chia et al. 2009). No estado também estão às maiores empresas nacionais produtoras de óleo de palma, distribuídas pelos municípios de Mojú, Tatilândia, Acará, Tomé Açu, Bonito, Igarapé-Açu, Santo Antônio de Tauá, Santa Izabel do Pará e Castanhal. Quase toda a produção é destinada à indústria alimentícia, embora atualmente, o mercado de combustível seja a mais forte tendência (Becker, 2010).

A importância do cultivo do dendezeiro na Amazônia e, notadamente, para o estado do Pará, está associada com vários aspectos, incluindo: alta produção de óleo vegetal que é utilizado nas indústrias de cosméticos e como combustível alternativo; alta capacidade de fixação de carbono e de proteção do solo contra erosão; importante alternativa para a ocupação das áreas desmatadas propícias para essa cultura (Homma et al. 2000; Veiga e Furia, 2000).

A cultura do dendezeiro apresenta ciclo econômico produtivo de 25 anos e, para uma ótima produção agroindustrial, o ideal seria que a produção de cachos fosse uniformemente distribuída durante o ano (Bastos et al. 2001).

Os principais produtos produzidos por essas palmeiras são os óleos de palma e de palmiste, extraídos da polpa do fruto e da amêndoa, respectivamente. Essa palmeira é a

oleaginosa de maior produtividade conhecida no mundo, permitindo extrair entre 4 e 5 toneladas de óleos de palma e de 1 a 1,5 toneladas de óleo de palmiste por hectaar/ano (Muller e Alves, 1998).

O óleo de palma apresenta coloração amarelada e/ou avermelhada sabor adocicado e aroma forte. É empregado na indústria oleoquímica na produção de sabão, sabonete, velas, tintas, dentre outros. Na indústria, após o refino e desodorização é utilizado como matéria-prima na fabricação de margarina, manteiga vegetal, óleo de cozinha (Kritchevsky, 2000; Nagendran et al. 2000). O óleo de palmiste normalmente é utilizado na indústria de cosméticos (Norizzah et al. 2004; Basiron, 2007).

A composição do óleo de *E. guineensis* e *E. oleifera* difere apenas na composição dos ácidos graxos. Em *E. guineensis* predomina o ácido palmítico, já na *E. oleifera* a maior concentração é de ácido oléico, que é um ácido graxo insaturado e confere maior fluidez ao óleo (Asemota, 2004). A *E. oleifera* apresenta óleo de baixo rendimento e não apresenta preferência para o cultivo comercial, apesar de conter mais de 69% de ácido oléico (Bhore e Shan, 2011).

A temperatura ambiente para o óleo de *E. guineensis* apresenta uma fração líquida próxima de 50%, enquanto, que em *E. oleifera* a fração é superior a 80%. Portanto, o óleo produzido por híbridos interespecíficos (OxG) apresenta uma fluidez maior pela presença do ácido oléico, reduzindo significativamente os gastos com os processos de refinamento do óleo (Asemota, 2004).

### 2.2.1 Amarelecimento Fatal (AF)

O dendê é susceptível a uma diversidade de doenças patogênicas, entre elas o Amarelecimento Fatal (AF), o que tem gerado grande perda no negócio de óleo da palma sul-americana. É considerada uma enfermidade devastadora e misteriosa, devido seu agente etiológico ser desconhecido (Boari, 2008).

Tal anormalidade vem dizimando os plantios de dendê, forçando produtores e melhoristas a buscarem fontes de resistência em espécies relacionadas como o Caiuá. Sua incorporação aos programas de melhoramento genético de Palma de Óleo traz, então,

enormes expectativas para o incremento da produção e qualidade de óleo, e consequentemente, para a sustentabilidade para a palmicultura no país (Rios et al. 2012).

O AF se caracteriza, inicialmente, pelo ligeiro amarelecimento dos folíolos basais das folhas intermediárias (3, 4, 5 e 6) e, mais tarde, pelo aparecimento de necroses nas extremidades dos folíolos que evoluem para a seca total dessas folhas (Boari, 2008).

Um dos principais sintomas do AF é a seca da folha flecha e, eventualmente, pode ocorrer à remissão temporária da planta, seguida do declínio generalizado e morte. Van Slobbe e Souza (1991), relataram que, geralmente, as plantas morrem com 7 a 10 meses após o aparecimento dos primeiros sintomas, quando não ocorre a remissão. A partir da morte da folha flecha, não há mais a produção de cachos. Além disso, o sistema radicular não se desenvolve após o aparecimento dos primeiros sintomas do AF e, se houver remissão de folhas, a produtividade de cachos é muito baixa (Figura 3) (Boari, 2008).



Figura 3: *Elaeis guineensis* apresentando Amarelecimento Fatal (AF). Plantas mantidas no banco de germoplasma da Biopalma, situada no município de Moju, estado do Pará. (A) Remoção da área afetada pela anomalia. (B) Aspecto do sistema radicular saudável (raiz superior) exibindo crescimento de raízes secundárias (em branco) e raízes afetadas pela doença (raiz inferior).

Nos tecidos da estirpe e meristema das plantas com AF, não são observados apodrecimentos ou necroses do sistema vascular (Boari, 2008). O sistema radicular se apresenta necrosado logo no início do aparecimento do amarelecimento dos folíolos das folhas intermediárias. Os cachos originados de folhas não afetadas apresentam-se normais (Ayala, 2001; Bernardes, 2001).



Números de pesquisas em campo demonstram que a incidência, a gravidade e os sintomas do AF estão intimamente associados a condições abióticas (Chinchilla, 2010), além de mudanças abruptas na disponibilidade de água, tanto a seca causada pelo déficit de água em solos argilosos e a anoxia causada pelo excesso de água no solo (Reyes et al. 2013). Isto pode ser proposto devido aos períodos de estresse serem mais acentuados na presença de fenômenos como La Niña/ El Niño, que parecem está relacionados ao AF como foi relatado na Colômbia (Cadena et al. 2004).

Plantas sob estresse abiótico, reajustam o seu metabolismo para tornar eficiente o uso de água e carbono, a fim de evitar a falta de nutrientes. No entanto, o avanço do AF é irreversível e a ausência de minerais significa, naturalmente, um colapso completo de funções da planta, além de ser uma resposta de defesa, e quando as plantas apresentam deficiência nutricional, estas são altamente suscetíveis ao ataque de patógenos (McDowell et al. 2011).

Neste cenário, planta que sofre carência de nutrientes relacionada à presença de fatores abióticos, sucumbiria ao AF causado por combinações de um ou mais agentes patogênicos para diferentes palmeiras. Uma explicação diferente pode estar relacionada com as mesmas origens genéticas de resistência aos estresses abióticos e bióticos (Atkinson e Urwin, 2012).

## 2.3 BIOSSÍNTESE DE COMPOSTOS ORGÂNICOS VOLÁTEIS (COVs) EM ÓLEO DE PALMA

### 2.3.1 A função dos metabólitos das plantas na ecosfera

As plantas desenvolveram um incrível arsenal de mecanismos metabólicos naturais de defesa e interação com o meio ambiente. A biossíntese de metabólitos é realizada por rotas específicas do organismo (Dudareva et al. 2013).

O metabolismo consiste nas modificações de moléculas orgânicas nas células que são catalisadas por enzimas. Convencionalmente é dividido em metabolismo primário e

secundário. Entretanto, as vias metabólicas são interligadas, havendo uma alteração em uma das vias, conseqüentemente, acarretará modificação na outra.

Os principais compostos do metabolismo primário consistem nos carboidratos, proteínas, lipídeos, aminoácidos, ácidos nucleicos (DNA e RNA) que são essenciais à vida do organismo, por meio da via fotossintética. Essa rota metabólica atua como precursor para a biossíntese dos compostos do metabolismo secundário.

O metabolismo secundário envolve um arsenal de metabólitos especializados. Nas plantas atuam em diversas funções, como defesa contra o ataque de herbívoros e patógenos, na atração de polinizadores, dispersores de sementes, estresses abióticos e bióticos, dentre outros (Muhlemann et al. 2014). Os principais grupos constituintes desse metabolismo são os alcaloides ou compostos nitrogenados, os fenóis e terpenos ou terpenoides (Dudareva et al. 2013).

Todas as plantas produzem metabólitos secundários, os quais variam em quantidade, qualidade e local de produção de uma espécie para outra, sendo sua síntese desencadeada por eventuais mudanças no meio em que se encontram (Ferreira e Aquila, 2000).

### **2.3.1.1 O papel dos metabólitos voláteis nas interações entre planta e meio ambiente**

As plantas sintetizam uma diversidade de compostos orgânicos voláteis (COVs), que são envolvidos na defesa contra as condições adversas do ambiente em resposta ao estresse biótico ou abiótico e permitem a comunicação entre a planta e o meio ambiente. A liberação de COVs florais assume função de atração de polinizadores (Koeduka et al. 2009b) e defesa (Gang et al. 2001) contra florívoros e patógenos (Muhlemann et al. 2014).

Os voláteis são um grupo relativamente grande de produtos vegetais naturais, consistem em um líquido lipofílico com baixo peso molecular. A propriedade física destes compostos lhe permite atravessar livremente a membrana celular e ser liberado para o meio ambiente (Pichersky et al. 2006). A biossíntese de COVs ocorre nas raízes, caules, folhas, frutas, sementes e flores (Muhlemann et al. 2014).

Os voláteis vêm sendo identificados em 90 diferentes famílias de plantas ambas as angiospermas e gimnospermas, respectivamente (Knudsen, et al. 2006). Nas plantas, a formação de compostos orgânicos voláteis é especialmente mais regulada. Os COVs são

liberados predominantemente de flores na maioria das plantas com flores (Knudsen et al. 2006), tecidos vegetativos, como folhas e caules em coníferas, e folhas de ervas, tais como hortelã-pimenta e manjeriço, representam a principal fonte de emissão de COV nestas espécies (McConkey et al. 2000; Vassao et al. 2006).

Em todos os órgãos das plantas, a produção e liberação de COV exibem perfis semelhantes de desenvolvimento, aumentando durante as fases iniciais (quando as folhas são jovens e não estão totalmente expandidas, quando os frutos ainda não estão maduros, ou antes, das flores estarem prontas para a polinização) e diminuem ou permanecem relativamente constantes, posteriormente (Dudareva et al. 2000; Vassao et al. 2006).

Nos tecidos vegetais são produzidos pequenas quantidades de COVs que podem ser induzidos por danos mecânicos e herbívoros ou infecção de patógenos (Arimura et al. 2004). Voláteis liberados a partir de frutas e sementes e/ou esporos atraem vários animais dispersores que são essenciais para a dinâmica espacial de populações das plantas nos ecossistemas (Dudareva et al. 2013).

Os voláteis são afetados por fatores abióticos como umidade do solo e ar, intensidade luminosa, fotoperíodo e nível de fertilização (Gouinguéné e Turlings, 2002). Estas variações oferecem aos inimigos naturais informações não somente a respeito da presa, mas também sobre a planta e seu estado fisiológico (Pinto-Zevallos et al. 2013).

Os COVs induzidos após o ataque dos herbívoros têm despertado o interesse na comunidade científica, uma vez que eles não só guiam os inimigos naturais em direção à planta hospedeira da presa, mas também na mediação de várias interações ecológicas que têm potencial para o desenvolvimento de novas estratégias de controle de pragas (Pinto-Zevallos et al. 2013).

### **2.3.1.2 Biossíntese do composto volátil metil-chavicol nas plantas**

A biossíntese de COVs depende da disponibilidade de carbono, nitrogênio e enxofre para o metabolismo primário e é importante para o metabolismo secundário. Com base em sua origem biossintética, todos os COVs são divididos em várias classes, incluindo os

terpenoides, fenil-propenoides/benzenoides, derivados de ácidos graxos saturados, aminoácidos, dentre outros compostos químicos (Dudareva et al. 2013).

Os fenil-propenos são uma classe de compostos voláteis encontrados em todas as gimnospermas e angiospermas (Koeduka et al. 2009b). É um grupo pequeno de moléculas fenólicas, derivadas do gênero fenil-propanoide que é a chave dos elementos flavonoides muito importante de ervas e temperos (Gang et al. 2002).

O fenil-propanoide e os compostos benzênicos representam a segunda maior classe do COV em plantas (Knudsen et al. 2006) e são exclusivamente derivados do composto aromático fenilalanina (Phe) e apresentam uma ampla variedade de funções (Muhlemann et al. 2014).

A fenilalanina é transformada em coenzima A (CoA) e convertida em acetato cinâmico por desaminação, seguida por metilação e hidroxilação para gerar o acetato *p*-cumaril (Figura 4) e outros como o fenil-propano (C<sub>6</sub>-C<sub>3</sub>), que são as duas principais vias na biossíntese de monolignóis e flavonóides (Ferrer et al. 2008).

Os fenil-propenos dependem da estrutura do carbono e são divididos em três subclasses: fenil-propanoides (C<sub>6</sub>-C<sub>3</sub>), benzenoides (C<sub>6</sub>-C<sub>1</sub>) e compostos relacionados aos fenil-propanoides (C<sub>6</sub>-C<sub>2</sub>) (Muhlemann et al. 2014). Os fenil-propenos incluem o eugenol e o metil-chavicol (MeC), sintetizados por organismos vegetais e emitidos pelas flores que servem como atrativos para insetos polinizadores (Koeduka et al. 2009b), e suas propriedades antimicrobianas e aromáticas são utilizados pelos seres humanos como pesticidas e conservantes de alimentos.

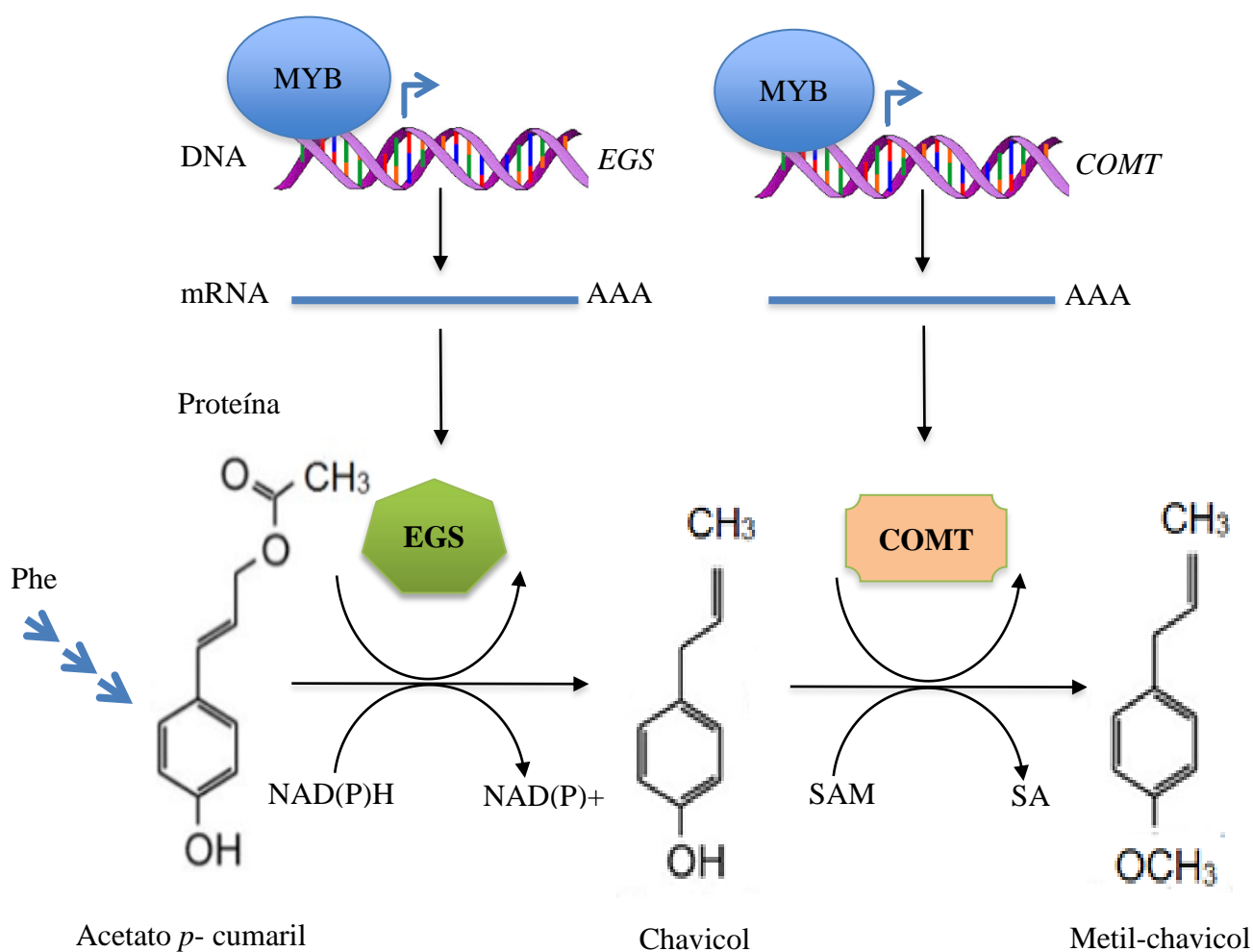


Figura 4: Síntese de metil-chavicol (MeC) nas plantas de Manjeriçao e Petúnia. O fator de transcrição MYB, inicia a expressão dos genes eugenol sintase (*EGS*) e chavicol *O*-metiltransferase (*COMT*) (seta em azul). A biossíntese do MeC depende da fenilalanina (Phe), esse aminoácido é convertido em acetato *p*-cumaril, após várias etapas enzimáticas, neste estudo será dado ênfase na via biossintética de MeC. A partir desse substrato o EGS sofre oxidação, a partir do seu co-fator enzimático NAD(P)H para NAD(P)<sup>+</sup> e forma o composto chavicol. O COMT sofre metilação, por meio do seu co-fator enzimático SAM para SA, e assim forma o composto volátil MeC. NAD(P)H (dinucleótido de nicotinamida e adenina fosfato) e SAM (S-adenosilmetionina).

Outro componente principal dos óleos essenciais é o *t*-anetol presente nas sementes de anis (*Pimpinella anisum*) Apiaceae e erva-doce (*Foeniculum vulgare*), utilizados como fumigante contra mosquitos (Chang e Ahn, 2002).

Vários componentes do fenil-propeno são classificados diferentemente um do outro em dois aspectos: os radicais ligados ao anel fenil e a posição da dupla ligação na cadeia lateral propenil. O chavicol apresenta uma *para*-hidroxil funcional no anel fenil, mas a dupla

ligação na cadeia propenil está entre C<sub>8</sub> e C<sub>9</sub> (Koeduka et al. 2009b). A síntese de alguns fenil-propenos encontrados em plantas, como o *t*-anetol, metil-chavicol, metil-eugenol e metil-isoeugenol, requerem metilação para o grupo *para*-hidroxil em seu anel fenil, respectivamente (Dudareva et al. 2013).

O principal COVs contido nas flores de dendê foi identificado como 1-metoxi-4 (2-propenil) benzeno, também conhecido como metil-chavicol (MeC), responsável por atrair insetos polinizadores (Lajis et al. 1985). A estrutura química do MeC consiste em um anel benzeno substituído por um grupo metoxi e propenilo. O MeC é um isômero de ligação dupla e componente comum nas plantas (Bidri et al. 2012).

A enzima eugenol sintetase (EGS) utiliza o substrato acetato cumaril para convertê-lo em seu respectivo produto por meio da via intermediária da quinona (Koeduka et al. 2008; Koeduka et al. 2006; Louie et al. 2007). O EGS está intimamente relacionado com várias outras redutases dependentes de NADP(H) envolvidos na biossíntese de fenil-propanoide em plantas (Koeduka et al. 2009b).

O eugenol natural produz apenas um produto, o eugenol ou isoeugenol, mas estudos de mutagênese *in vitro* de alguns resíduos no sítio ativo, a estrutura cristalográfica mostrou uma influencia na posição da ligação do substrato em uma enzima que produz uma mistura de regiões isoméricas de fenil-propeno (Koeduka et al. 2008).

A toxicidade de COV em concentrações elevadas para as células em geral, torna-os úteis como compostos de defesa e, conseqüentemente, são encontrados em tecidos vegetativos de várias espécies de plantas. Devido a sua toxicidade eles podem ser sequestrados em estruturas especializadas ou emitidas rapidamente a partir da planta, por exemplo, o eugenol e o metil-chavicol são sintetizados e armazenados nos tricomas glandulares na superfície das folhas de manjeriço (*Ocimum basilicum*) (Gang et al. 2001; Lijima et al. 2004). Foi reportado que nas glandulas de manjeriço duas enzimas (90% idênticas) de chavicol *O*-metiltransferase (COMT) e eugenol *O*-metiltransferase (EOMT) catalisam a formação de metil-chavicol e metil-eugenol para chavicol e eugenol, respectivamente (Gang et al. 2001). As regiões conservadas de DNA ou motivos de EOMT e COMT desempenham importante papel na catálise de metabólitos secundários. COMT apresenta pouca atividade com *t*-anetol, enquanto, o EOMT sua atividade é quatro vezes maior com eugenol do que com o chavicol (Wang e Pichersky, 1998).

Genes que codificam a biossíntese de enzimas de fenil-propenos foram identificados a partir de várias espécies (Koeduka et al. 2009a). Entretanto, as enzimas de óleo de palma envolvidas na biossíntese de MeC e sua expressão na célula são desconhecidas (Koeduka et al. 2009b).

### 2.3.2 Polinizadores

A polinização é caracterizada pela transferência de grãos de pólen de anteras de flores masculinas para o aparelho reprodutor de flores femininas (estigma), permitindo assim a fecundação da planta. Pode ser realizada por fatores abióticos, tais como chuva e vento ou fatores bióticos por animais polinizadores e dispersores de sementes. Muitas angiospermas dependem de polinizadores, principalmente insetos, que estão entre os mais importantes polinizadores pela quantidade de plantas que podem visitar, e conseqüentemente, promover a fecundação e a formação de frutos (Gomes, 2011).

Os polinizadores buscam por algum tipo de recompensa nas flores, seja aumento (pólen e néctar), abrigo para acasalar e/ou ovopositar (Gomes, 2011). O ganho para as plantas e polinizadores conduz a evolução de adaptações de plantas a diferentes polinizadores (Próctor, 1996).

Os insetos estão entre os polinizadores mais importantes de plantas que dependem deles para realizar a fecundação. Responsáveis pela polinização de 86% dos plantios comerciais de frutas, nozes e sementes (Nabhan e Buchmann, 1997; Kevan e Imperatriz-Fonseca, 2002). Aproximadamente 73% das espécies vegetais cultivadas no mundo são polinizadas por abelhas, 5% por besouros e 22% por outros insetos, além de mamíferos e pássaros (FAO, 2004). Plantios no sul da Bahia mostraram uma taxa de fecundação de 28% para o híbrido (OxG) contra 79% para a espécie Africana (Moura et al. 2008).

#### 2.3.2.1 O papel dos polinizadores no óleo de palma

No final dos anos 70 acreditava-se que a reprodução do dendê era anemófila ao invés de entomófila (Turner e Gilbanks, 1974; Hardon e Corley, 1976). A crença de que o dendê

era polinizado pelo vento partiu da suposta falta de evidência de qualquer inseto polinizador eficaz (Candwell et al. 2003). Entretanto, Syed (1979) em seus estudos sugeriu que os insetos apresentam um papel significativo na polinização do dendê, e constatou a presença de um grande número de insetos presentes nas inflorescências durante o período de antese que emitiam um odor parecido com anis o que atraia os insetos. Desses insetos relatou a presença de *Elaeidobious* spp. e *Atheta* spp.

Segundo Hussein et al. (1989), numerosas espécies de insetos estão envolvidos na polinização das palmeiras, mas os principais agentes polinizadores são besouros entre os quais *Elaeidobius kamerunicus* e *Elaeidobius subvillatus* (Coleoptera: Curculionidae) predomina (Figura 5). Estes besouros alimentam-se de partes moles de flores masculinas e pólen, e depositam ovos nas flores masculinas. Os grãos de pólen contêm amido e constituem a maior fonte de alimento desse besouro, também conhecidos por gorgulhos, que habitam as inflorescências masculinas em grande número ( $2000 \pm 3000$  indivíduos), respectivamente (Tandón et al. 2001).

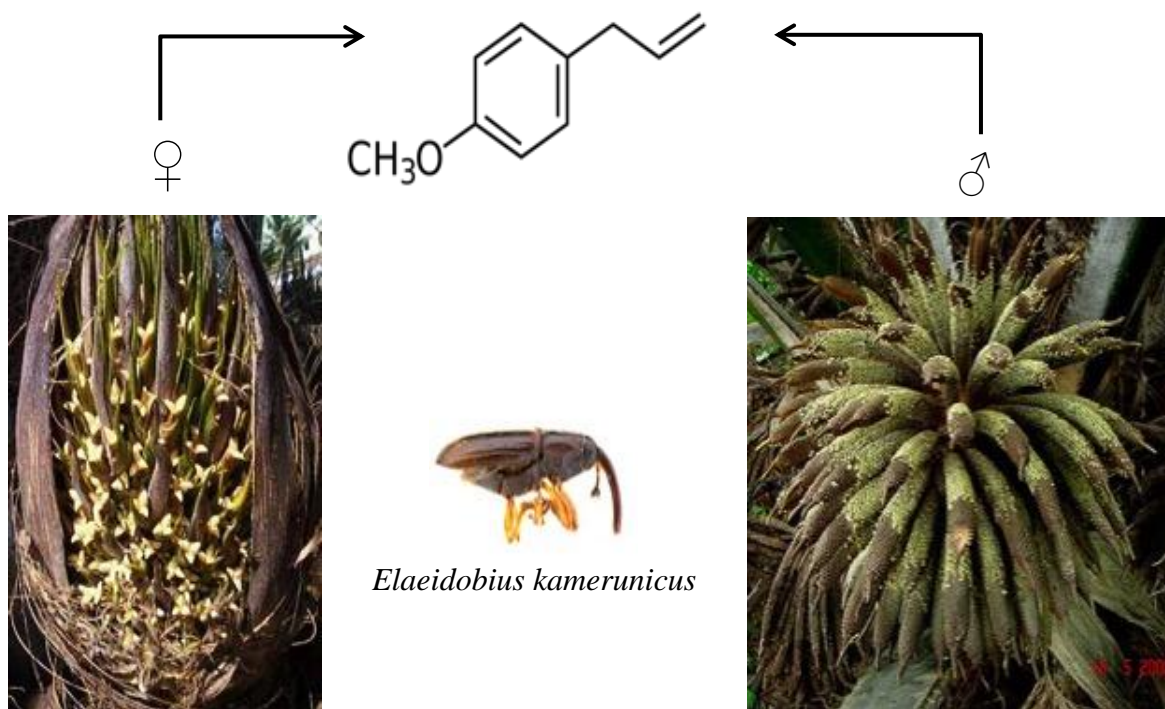


Figura 5: Inflorescências femininas e masculinas de *E. guineensis* liberam compostos voláteis que são atrativos de insetos polinizadores. As flores emitem o metil-chavicol, responsável por atrair insetos adultos de *Elaeidobius kamerunicus*, principal agente polinizador do dendê.



As Inflorescências femininas emitem aroma de anis ou erva-doce semelhante ao odor produzido pelas flores masculinas atraindo besouros da espécie gorgulhos que visitam as flores femininas por acidente em busca das inflorescências masculinas (Tandón et al. 2001).

As flores geralmente emitem os mais altos valores e as mais diversas misturas de COVs (Knudsen et al. 2006) para atração de insetos polinizadores. Os perfis de compostos voláteis emitidos pelas flores são específicos para cada espécie e tendem a refletir o tipo de polinizador (Dobson, 2006). Os voláteis ocasionam a atração de polinizadores às flores (Raguso, 2004), entre as plantas e herbívoros e animais disseminadores de sementes (Pichersky et al. 2006), assim como a localização de plantas hospedeiras por insetos herbívoros (Bruce et al. 2005).

O principal composto volátil das flores do dendê é o metil-chavicol (MeC) (Lajis et al. 1985). Este composto orgânico volátil pode ser responsável por atrair insetos adultos *E. kamerunicus* para a polinização do óleo de palma (Hussein, 1989). Para as espécies de híbridos (OxG) a baixa produtividade, provavelmente, é devido as suas inflorescências serem menos atraentes para os insetos polinizadores como o *E.kamerunicus*. Isto pode ser devido aos baixos níveis ou a falta de COVs apropriados presente nas inflorescências (Corley e Tinker, 2008).

Estudos realizados por Gomes (2011), observou que inflorescências femininas e masculinas de *E. guineensis (tenera)*, *E. oleifera* e o híbrido (OxG), liberam quantidades semelhantes do composto volátil MeC. No entanto, as inflorescências de *tenera* liberam uma quantidade maior desse volátil que as demais, o que justificaria a maior atratividade de polinizadores.

A variação na quantidade de metil-chavicol liberada por plantas de *tenera*, caiaué e híbrido (OxG), pode está relacionada a maior ou menor atratividade de polinizadores. Espécimes de *E. kamerunicus* são atraídos no campo por concentrações de MeC entre 150 a 200µl/armadilha (Hussein, 1989). A maior quantidade desse volátil liberada por *tenera* explica a maior atração de polinizadores (Gomes, 2011).

A liberação de MeC nas inflorescências de *E. oleifera* pode não ser suficiente para atrair adultos de *Elaeidobius*. No entanto, no estado do Amazonas curculionidaeos do gênero *Celestes* e de outras duas espécies, em processo de identificação, além dos nitidulídeos do gênero *Mystrops*, são atraídos, em grandes números pelas inflorescências masculinas e

femininas. O gênero *Celestes*, pertence à subtribo Derelomina a mesma subtribo das espécies de *Elaeidobius* (Gomes, 2011). Este espécime encontra-se sobre as inflorescências femininas de *E. oleifera*, raspando a superfície do estigma da planta, este comportamento, prejudica a germinação de grãos de pólen uma vez que o estigma de *Elaeis* spp., produz uma substância rica em polissacarídeo responsável por promover a germinação do grão de pólen (Tandón et al. 2001).

A baixa disponibilidade de inflorescências masculinas e, conseqüentemente, de pólen conduz a produção de cachos abortivos ou mal formados, devido à baixa polinização natural (Berthaud et al. 2000). Para que ocorra boa taxa de frutificação nos cachos, é necessário que haja, ao mesmo tempo, boa disponibilidade de pólen e uma população de insetos polinizadores em quantidade significativa (Berthaud et al. 2000).

Inúmeras culturas que dependem de polinizadores, provavelmente, emitem baixas concentrações de COVs, no entanto, para o nosso conhecimento não existem programas de melhoramento para selecionar por fragrâncias florais que sejam “amigas de polinizadores”, sugerindo assim que os serviços de polinização podem ser melhorados através da produção de flores mais perfumadas (Dudareva et al. 2013). Além disso, muitos cultivos modernos são deficientes na produção de COV, como resultado da criação para o aumento do crescimento ou rendimento, a qual afeta negativamente o metabolismo secundário dos COVs (Barcelos et al. 2002; Dudareva et al. 2013).

Uma compreensão abrangente de estudos metabólicos em plantas e sua regulação, bem como o conhecimento sobre os compostos voláteis, envolvidos nas interações planta-inseto, irá avançar nos estudos metabólicos de COVs.

## 2.4 APLICAÇÃO DA BIOINFORMÁTICA NA BIOLOGIA MOLECULAR

Os avanços tecnológicos vêm sendo empregados como ferramentas importantes no campo da Biologia. Desta forma, as ciências existentes tiveram que se adequar aos princípios tecnológicos, informatizando-se para poder acompanhar o avanço que a informática proporcionou e vem proporcionando à humanidade. Assim, surgiu a

Bioinformática, utilizada como ferramenta para auxiliar na investigação da biotecnologia, como a interpretação de genomas (Alves, 2013).

O conceito de Bioinformática surgiu na década de 1990 como uma nova área do conhecimento científico, representando a interação da Biologia com a Informática. Entretanto, o reconhecimento da Bioinformática deu-se em 1995, após a leitura do genoma de uma bactéria. Pode-se afirmar que o surgimento da Bioinformática clássica ocorreu com o sequenciamento de biomoléculas. Atualmente, a Bioinformática está bastante avançada. Este avanço pode ser percebido quando se decifra um gene com aproximadamente 12 mil bases em menos de um minuto. E que em 2000, para se realizar essa mesma tarefa era necessários 20 minutos, enquanto que no final da década de 1980 levava-se mais ou menos um ano (Alves, 2013).

A Bioinformática consiste principalmente na análise computacional de sequências de DNA, RNA e proteínas. Essa nova ciência surgiu na última década em função da necessidade de ferramentas sofisticadas para analisar o crescente volume de dados gerados em biologia molecular (Santos, 2004). Utiliza ferramentas computacionais avançadas tanto para o armazenamento quanto para análise de dados biológicos e moleculares. Na concepção de Silveira, (2005), a Bioinformática é uma área interdisciplinar que absorve a ciência da computação, matemática, biologia, física e medicina.

O domínio dessa ciência cresceu em paralelo com a internet, em que a rápida análise de dados e a troca de informações sobre os códigos biológicos e computacionais estão em convivência harmônica, por meio de suas múltiplas ramificações, gerenciando e integrando bancos de dados aplicáveis, e construindo sistemas *in silico* para simulação de formas naturais e modificadas de produtos específicos (Espindola et al. 2010).

A Bioinformática foi desenvolvida com o objetivo de mapear as sequências genéticas. Após, seu desenvolvimento os biólogos moleculares passaram a utilizar métodos estatísticos capazes de analisar grandes quantidades de dados biológicos. Atualmente, com a Bioinformática, é possível predizer funções de genes e demonstrar relações entre genes e proteínas (Alves, 2013).

Através da criação de bancos de dados com as informações já processadas, é possível acelerar a investigação em outras áreas como a Medicina, a Biotecnologia, a Agronomia, entre outras. Ademais, além de oferecer algoritmos, bancos de dados e ferramentas

estatísticas, a Bioinformática auxilia no desenvolvimento de *softwares* para a Biologia Molecular e para web sites de Biotecnologia (Araújo, 2005).

A Bioinformática representa, hoje, um dos grandes desafios para se tentar decifrar o genoma, pois ao mesmo tempo é uma forma de se conseguir informações imediatas para os dados do genoma que vem sendo descobertos, também é a base para um sucesso científico futuro (Wieczorek et al. 2008).

### 3. MATERIAL E MÉTODOS

#### 3.1.1 Análise de bioinformática na biossíntese de metil-chavicol

Para caracterizar as vias metabólicas dos compostos orgânicos voláteis (COVs) da biossíntese do metil-chavicol (MeC) nas espécies de dendê foi realizado um levantamento com base na literatura das sequências de aminoácidos (aa) do gene-alvo chavicol *O*-metiltransferase (*COMT*) e eugenol sintase (*EGS*), no banco de dado genômico NCBI (*National Center for Biotechnology Information*) e MPOB (*Malasyan Palm Oil Board*).

Foi utilizado as folhas de manjeriço (*Ocimum basilicum* L.) para identificar os genes ortólogos envolvidos na síntese do MeC nas espécies de dendê. As sequências de (aa) do gene-alvo obtidos pelo algoritmo BLAST (*Basic Local Alignment Search Tool*) em formato FASTA foram comparados com as sequências de *E. guineensis* e *E. oleifera*, e assim, recuperar as sequências de genes homólogos. Posteriormente, foram alinhados utilizando o *software* CLUSTALW (*Multiple Sequence Alignment*) para a identificação das regiões conservadas (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Os oligonucleotídeos (*primers*) foram desenhados com base em sua especificidade, alguns foram projetados em regiões promotoras do gene (Figura 6).

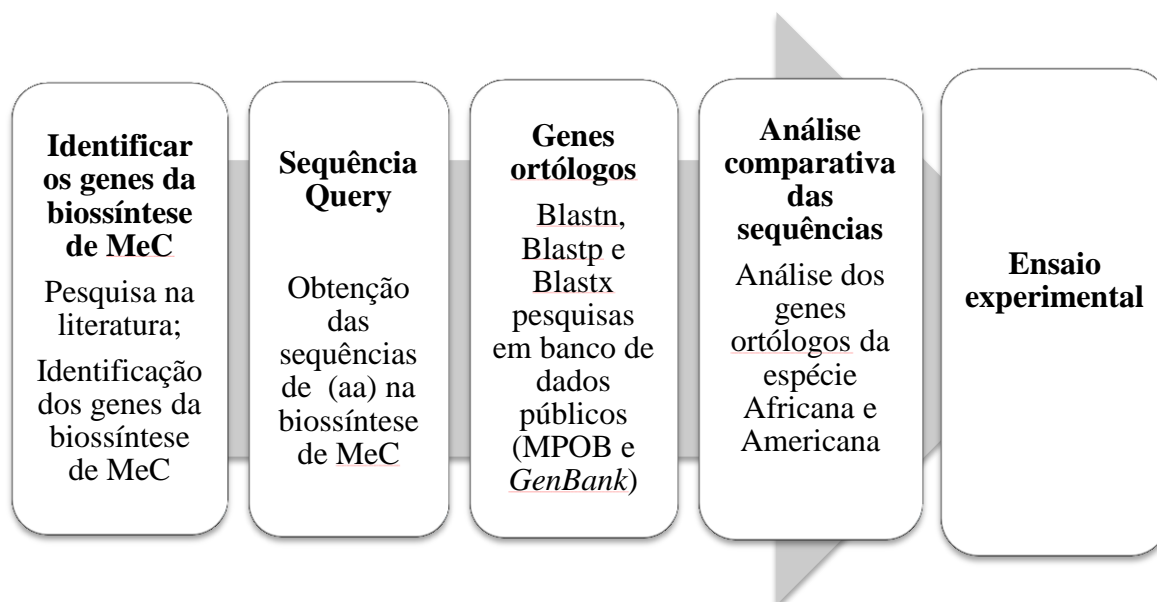


Figura 6: Fluxograma das análises *in silico* e experimental da biossíntese de metilchavicol (MeC) em *Elaeis guineensis* e *Elaeis oleifera*.

### 3.1.2 Material Vegetal

Os materiais vegetais de *Elaeis guineensis* avaliados foram coletados nos acessos da Biopalma subsidiária da Vale, localizado no município de Mojú, situado a uma latitude de 01°53'02" sul e a uma longitude de 48°16'08" nordeste do estado do Pará e no município de Manicoré, no estado de Manaus para *Elaeis oleifera* e híbridos interespecíficos (OxG). Foram avaliados tecidos vegetais frescos de ambas as espécies.

### 3.1.3 Extração do DNA genômico

O DNA genômico foi extraído de folhas frescas de *E. guineensis (tenera)* e *E. oleifera*, conforme o protocolo por centrifugação de lise e homogeneização utilizando pistilo e cadinho segundo as instruções do fabricante (*Axygen Biosciences*).

Inicialmente, 200mg de tecido vegetal fresco foi macerado em cadinho, adicionou-se 350µl de tampão fosfato salino (contendo 8g de NaCl; 0,2g de KCl; 1,44g de Na<sub>2</sub>HPO<sub>4</sub> e 0,24g de H<sub>2</sub>KO<sub>4</sub>P) e 0,9µl de ribonuclease para homogeneizar a amostra. Para o melhor aproveitamento da amostra macerada foi adicionado uma pequena alíquota do tampão.

O tecido macerado foi transferido igualmente para dois microtubos de 2mL e adicionadas 20µl de proteinase K e 150µl de tampão de lise. Em seguida, foram submetidos ao vórtex por um minuto. Por serem amostras fibrosas foram incubados em banho maria a 56°C por 30 minutos, sob agitação. Foram adicionados 350µl de tampão de precipitação de proteínas e submetidos ao vórtex em velocidade máxima por 30 minutos e centrifugados a 12.000 rpm por 10 minutos em temperatura ambiente para sedimentar os restos celulares.

Os sobrenadantes foram transferidos para uma coluna MiniPrep no microtubo de 2mL e o conteúdo das duas amostras do sobrenadante foram transferidos para uma única coluna e centrifugados a 12.000 rpm por um minuto. O filtrado foi descartado e a coluna recolocada no microtubo, logo após foi adicionado 500µl de tampão de lavagem na coluna e centrifugado a 12.000 rpm por um minuto. O filtrado foi descartado e colocado novamente na coluna no microtubo. Adicionou-se 700µl de tampão de dessalinização na coluna e centrifugado a 12.000 rpm por um minuto. Este passo de lavagem foi repetido duas vezes para garantir uma remoção completa de sais, eliminando-se possíveis problemas na reação.

O filtrado foi descartado e a coluna foi recolocada no microtubo e centrifugada por mais um minuto.

Em seguida a coluna foi transferida para um microtubo de 1,5mL e foi adicionado 200µl de eluente pré-aquecido a 65°C no centro da membrana da coluna. Por fim a amostra permaneceu por um minuto em temperatura ambiente e em seguida centrifugada por mais um minuto. O DNA foi mantido congelado a -20°C até o momento da realização da amplificação (PCR).

#### **3.1.4 Análise qualitativa e quantitativa do DNA**

O produto da extração do DNA foi quantificado e visualizado em gel de agarose 1%. Foram adicionados 10µl/mL de Syber®Safe DNA Gel Stain (*Invitrogen*), dissolvidos em 100mL de tampão TBE 1X (10,8g tris; 5,5g de ácido bórico e 0,75g de EDTA) com pH 8. Após, a corrida das amostras em gel foram visualizados e fotodocumentados sob luz UV. O marcador de peso molecular *1Kb Plus DNA Ladder* (*Invitrogen*) foi incluído em todos os géis avaliados (2µl em cada poço).

#### **3.1.5 Reação em Cadeia da Polimerase (PCR)**

A amplificação do DNA foi realizada pela técnica da reação em cadeia da polimerase (PCR), desenvolvido por Kary Mullis em 1985. A reação consistiu no protocolo GoTaq®Flexi DNA Polymerase (Promega), segundo as instruções do fabricante (Tabela 1).

Tabela 1: Volume de reagentes e concentração final na reação de PCR

Reagente	Volume Final ( $\mu$ l)
Água ultrapura	31,25
Tampão de reação	10,0
dNTP 40mM	1,0
Primer <i>foward</i>	1,5
Primer <i>reverse</i>	1,5
MgCl <sub>2</sub> 25mM	2,5
TaqDNA polimerase	0,55
Amostra	2,0
Volume Total	50,0

Inicialmente as reações foram incubadas a 95°C durante cinco minutos e por mais 30 segundos para a desnaturação da dupla fita do DNA. Na sequência, a temperatura foi reduzida para 55°C por 40 segundos, permitindo o anelamento do *primer* a uma região complementar no DNA *template* (Tabela 2). Na etapa seguinte a temperatura foi elevada a 72°C por 55 segundos para o pareamento da enzima TaqDNA polimerase com o DNA, finalizando por mais 10 minutos. Essa enzima tem por finalidade inserir os desoxinucleotídeos (dNTPs) para a formação das cadeias complementares etapa conhecida por extensão. Ao final desse processo fecha o primeiro ciclo da PCR. Para ter uma quantidade suficiente da sequência alvo foram realizadas 35 ciclos de termociclagem. A cada ciclo ocorre aumento exponencial na quantidade de moléculas de DNA amplificado.

Foram utilizados cinco pares de *primers* para cada amostra avaliada, correspondentes à sequências de 500 e 900 pares de base (pb), respectivamente (Tabela 2).



Tabela 2: Sequências de *primers* para a realização da PCR

<i>Primers</i>	Sequências (5'-3')
AS4	GTGCCGCTAAAATTCGAAAAC
AS5	CTGAAACATGTCTCCCGCGA
AS10	CAAACCACGATACTGCAAAT
AS11	CAAGACCATGGTACTGCAAAC
AS12	GGGATGTGCTCGAACATGTG
AS16	GAAAATTTACTCCCTCGAGAGG
AS17	GAAAATTTACTCCCTCGAGAGG
AS21	AATTGGAAAGGTCTCTAAAGACC
AS22	CAGCATTACGAGAGCGAAGGAC

As combinações de *primers*, número de acesso da sequência do gene a ser analisado e o tamanho esperado das bases nitrogenadas utilizados para a amplificação dos fragmentos de *E. guineensis* e *E. oleifera*, encontram-se na Tabela 3.

Tabela 3: Combinação de *primers*, número de acesso, tamanho esperado em pares de base (pb) utilizados na amplificação dos genes de *Elaeis* spp.

DNA Template	Número de acesso do gene	<i>Primers</i>	Tamanho esperado (pb)
EG1	EG4N993	AS11/AS12	841
EG2	EG4N49835	AS21/AS22	464
EG3	EG4N49844	AS16/AS17	861
EG4	EG4N49850	AS4/AS5	877
EO1	EO4N993	AS10/AS12	819
EO2	EO4N49835	AS21/AS22	464
EO3	EO4N49844	AS16/AS17	855
EO4	EO4N49850	AS4/AS5	874

Posteriormente, o produto da amplificação foi submetido à análise de eletroforese em gel de agarose 2% descrito anteriormente.

### 3.1.6 Purificação dos fragmentos de DNA amplificados pela PCR

A análise da purificação dos fragmentos de DNA amplificados pela PCR foi realizada conforme o protocolo de purificação do kit Prep Axy Prep (*Axygen Biosciences*), conforme recomendações do fabricante. Foi adicionado três vezes o volume do tampão de ligação do DNA a amostra. Após, foi acrescentado 700µl de tampão de dessalinização contendo etanol 100%. Esse processo de lavagem foi repetido com uma alíquota de 400µl e centrifugados a 12.000rpm. Para eluir o DNA adicionou-se 20µl de eluente pré-aquecido a 65°C em banho maria. Em seguida foi visualizada em gel de agarose 2%.

### 3.1.7 Análise dos fragmentos de DNA por sequenciamento de Sanger

Os produtos da purificação dos fragmentos da PCR foram submetidos ao sequenciamento de Sanger realizado no sequenciador (3730 DNA Analyser), conforme o protocolo do kit BigDye®XTerminator v3.1 Cycle Sequencing (*Life Technologies®/Applied Biosystems*), realizado no laboratório de Genômica do Instituto Tecnológico Vale (ITV).

A reação de sequenciamento consistiu em 6µl do reagente BigDye Mix, 5ng do amplicon purificado, 1pmol de cada *primers*, 4µl de tampão e 4µl água ultrapura para um volume final total de 20ul. Foi incluído controle positivo composto pelos reagentes: BigDye com volume de 6µl; pGEM (1µl); *primer* (4pmol) e água ultrapura (9µl) para um volume final de 20µl.

O protocolo padrão de termociclagem conteve uma etapa inicial a 96°C por um minuto, seguida por mais dez segundos, 50°C por 5 segundos e 60°C por 55 segundos, correspondentes à 35 ciclos finalizando com 4 minutos.

Os produtos da reação obtidos foram purificados, adicionando-se 2µl de EDTA (125 mM) e 2µl de acetato de sódio (3 M, pH 4,8), seguidos de 50µl de etanol absoluto 100%, em seguida foram incubados em temperatura ambiente por 15 minutos e centrifugados por 45 minutos a 4000 rpm, logo após, foi feito o descarte do sobrenadante.

Na etapa seguinte foram adicionados 70µl de etanol 70% e centrifugados por 10 minutos a 4000 rpm, em seguida o sobrenadante é desprezado. Foi realizado um spin com a placa invertida e o precipitado obtido foi submetido a estufa a 50°C por dez minutos, adicionou-se 10µl de formamida sob agitação por 15 segundos e aquecido a 95°C por dois minutos, seguidos por um choque térmico de dois minutos em gelo. Posteriormente, as amostras foram submetidas ao sequenciador automático (3730 DNA Analyzer). O produto do sequenciamento obtido foi purificado conforme o protocolo BigDye®XTerminator™Purification (Applied Biosystems).

As sequências obtidas pelo sequenciamento de Sanger foram visualizadas pela ferramenta FinchTv. Logo após, foram confirmadas pelo algoritmo Blastn, e em seguida alinhadas pelo software CLUSTALW para obter as sequências consenso entre *E. guineensis* (*tenera*) e *E. oleifera*.

## 4. RESULTADOS E DISCUSSÃO

### 4.1.1 Identificação de genes ortólogos *COMT* e *EGS* no genoma do dendê

A identificação dos genes nas espécies Africana e Americana, envolvidos na produção de metil-chavicol (MeC) foi baseada na via biossintética de *Ocimum basilicum* L (Anexo A). Foi identificado no genoma da espécie Africana, um conjunto de genes ortólogos que codificam enzimas que poderiam catalisar a biossíntese do MeC em duas etapas distintas, a partir do acetato *p*-cumaril, o eugenol sintase (EGS) e o chavicol-*O*-metiltransferase (COMT) (Anexo B).

As folhas de manjeriço possuem em sua superfície dois tipos de tricomas glandulares, denominados de peltados e capitados, glândulas constituintes do óleo volátil de eugenol e metil-chavicol, respectivamente (Gang et al. 2001). Segundo Gang et al. (2002) foram reportadas duas enzimas (OMT) com atividade *O*-metiltransferase definidas como eugenol *O*-metiltransferase (EOMT) e chavicol *O*-metiltransferase (COMT), identificadas em extratos de proteínas de manjeriço capazes de converter o chavicol e eugenol para metil-chavicol e metil-eugenol, respectivamente. Outros estudos com enzimas redutases

dependentes de NADP(H), como o EGS, isolados da Petúnia (*Petunia hybrida*) também estão intimamente envolvidos na biossíntese de fenil-propanoide (Koeduka et al. 2009a).

As sequências nucleotídicas das espécies Africana e Americana, obtidas dos bancos de dados MPOB e NCBI, foram alinhadas pelo algoritmo Blastp, utilizando as sequências COMT e EGS de *O. basilicum*, como query, em seguida, foram realizados alinhamentos múltiplos utilizando o *software* Clustal (Anexos C e D). Os genes *EGS* organizaram-se em três pequenas famílias e *COMT* apresentaram seis genes ortólogos, respectivamente. Indicando que esses genes poderiam mediar a biossíntese de MeC.

Para verificar a funcionalidade de genes ortólogos *EGS* e *COMT*, foram realizadas análises de modelagem por homologia baseadas nas sequências de proteínas de manjeriço, por meio do *software* SwissModeler e raio X para criar um modelo de proteína com conformação em 3D.

A análise AutoDoc identificou que a proteína putativa *O*-metiltransferase da família 2 do gene *COMT* seria capaz de ligar-se ao chavicol (em verde) e, portanto, são homólogas (Figura 7a). Para o gene *EGS* foi utilizada a proteína isoflavona redutase (em verde), identificada com capacidade de ligar-se ao acetato *p*-cumaril (em vermelho), verificando a existência de sequências homólogas (Figura 7b).

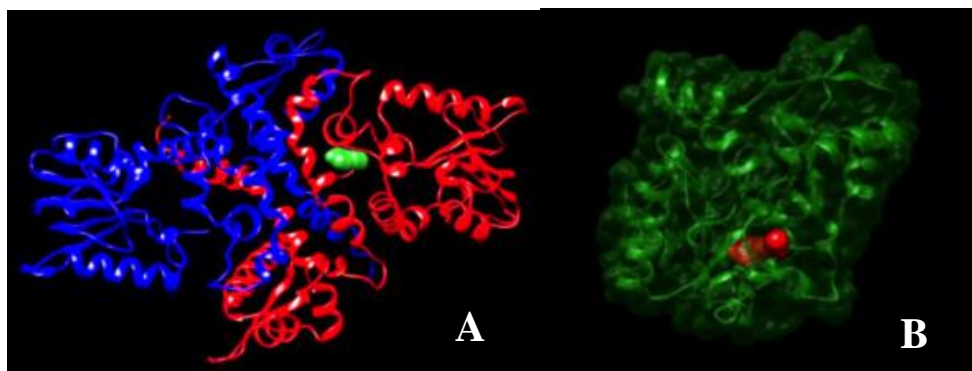


Figura 7: Modelagem por homologia dos genes chavicol *O*-metiltransferase (*COMT*) e eugenol sintase (*EGS*). (A) Análise do gene *COMT* ligando-se ao chavicol (em verde). (B) O gene *EGS* é capaz de ligar-se ao acetato *p*-cumaril (em vermelho).

De acordo com a técnica de modelagem por homologia, pode-se verificar similaridade estrutural e funcional de regiões conservadas nas sequências. Com base, nas

análises filogenéticas observa-se que as sequências de *E. guineensis* são semelhantes ao manjeriço (sp[Q93WU]) para os respectivos genes *COMT* e *EGS* (Figura 8 e 9).

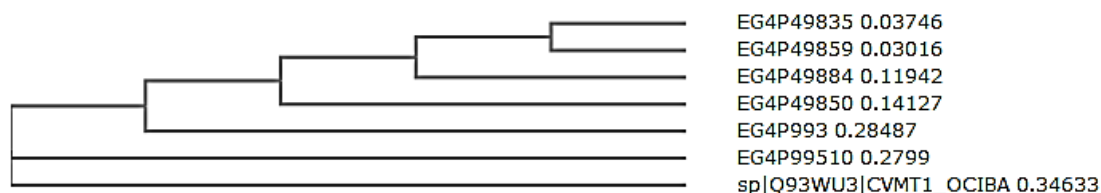


Figura 8: Árvore filogenética Neighbour-joining das sequências proteicas de *E. guineensis* são similares as sequências do manjeriço para o gene *COMT*.

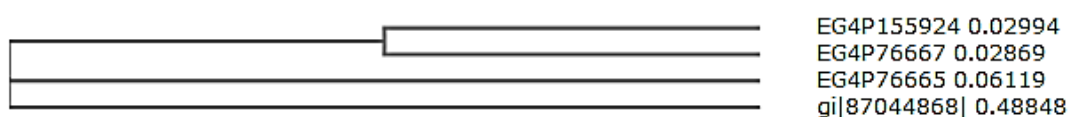


Figura 9: Árvore filogenética Neighbour-joining das sequências proteicas de *E. guineensis* são similares as sequências do manjeriço para o gene *EGS*.

O método de Neighbour-joining ou agrupamento de vizinhos consiste em formar árvores filogenéticas não enraizadas a partir de uma matriz de distância. Podem-se verificar sequências homólogas entre as espécies de palmeira com o manjeriço, respectivamente, corroborando com as análises de modelagem por homologia.

#### 4.1.2 Mutações nos genes *EGS* e *COMT* na espécie Americana

As sequências nucleotídicas de *E. guineensis* dos genes ortólogos *EGS* e *COMT2*, foram utilizadas como query para as sequências genômicas de *E. oleifera* depositadas no MPOB. Os alinhamentos das sequências de ambas as espécies, indicaram variações significativas entre os genes.

As análises de alinhamento das sequências gênicas mostraram que o gene ortólogo *COMT2* para a palma Americana apresentou 15 diferenças (polimorfismos) no éxon 1, sendo cinco *indels*. Para a espécie Africana não apresentou alterações, correspondendo à 783pb (Figura 10a).

A comparação dos alinhamentos dos homólogos de ambas as espécies indicam mutações *frame-shift* para o gene ortólogo *COMT2* para a espécie Americana (Figura 10b). As mutações *frame-shift* altera o modo de leitura na sequência de DNA, devido, a inserção e/ou deleção de um ou mais nucleotídeos. Mutações gênicas originam-se de alterações na sequência de bases nitrogenadas de um determinado gene durante a duplicação da molécula de DNA. Essa alteração pode ocorrer por perda, adição ou substituição de nucleotídeos, o que pode originar um gene capaz de codificar uma proteína diferente da qual deveria ter sido codificada.

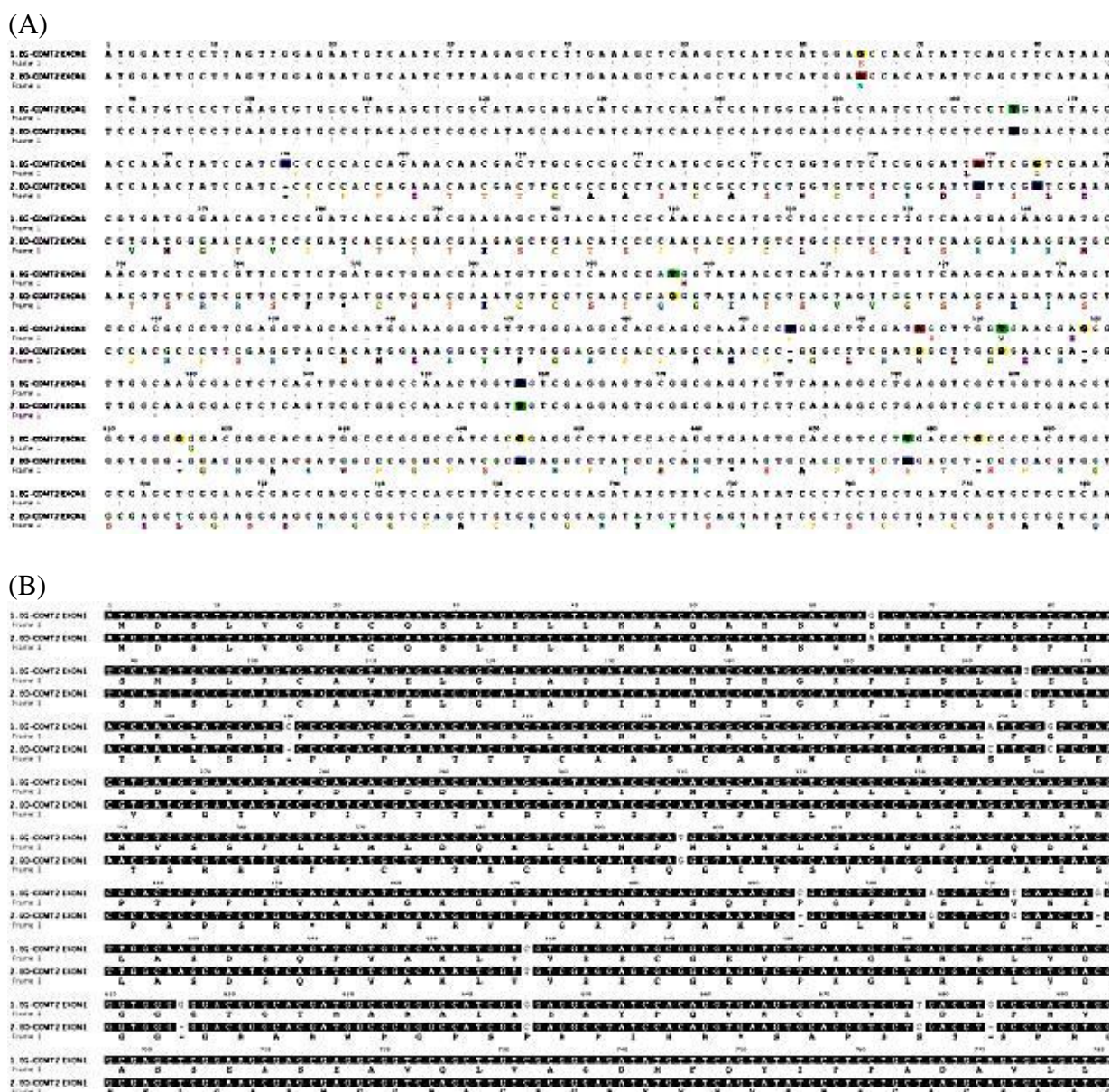


Figura 10: Alinhamento das seqüências do gene chavicol *O*- metiltransferase (*COMT2*) no éxon 1 das espécies Africana (EG) e Americana (EO). (A) Comparação entre as duas espécies, apresentando 15 polimorfismos, sendo cinco mutações *frame-shift*. (B) Diferença entre ambas as espécies. As seqüências idênticas encontram-se em preto e os *indels* estão indicados com um traço (-).

Alinhamentos com todas as seqüências homólogas de *E. oleifera* em comparação com *E. guineensis*, indica mutações *frame-shift*, para os genes ortólogos *COMT* e *EGS* (dados não mostrados).

Os resultados obtidos dos *indels* indicaram deslocamento no quadro de leitura (ORFs), o que aponta para mutações nos genes ortólogos, o que poderia explicar a deficiência da espécie Americana sintetizar o MeC (Figura 11).

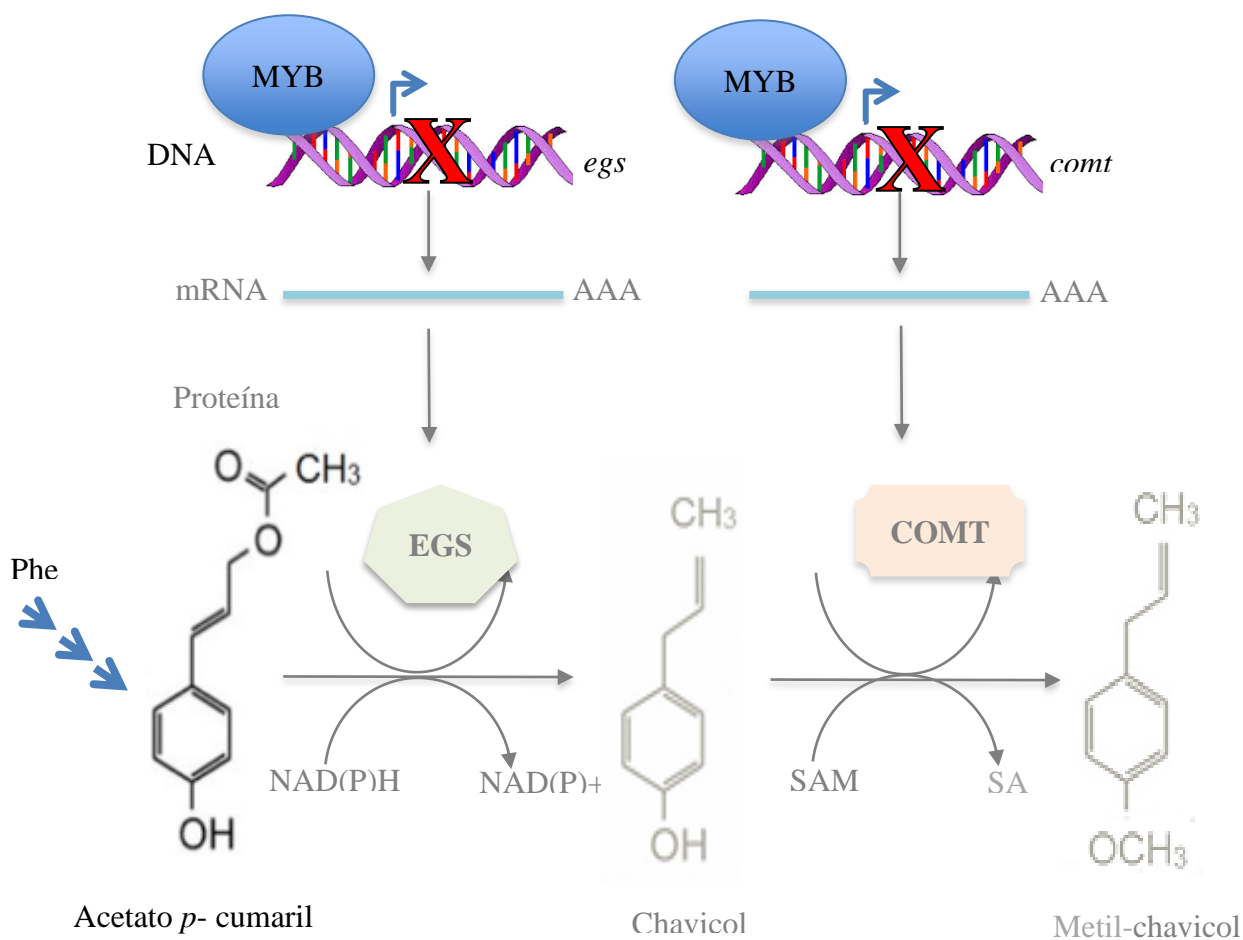


Figura 11: Deficiência na expressão do gene estrutural para a síntese de metilchavicol (MeC) em *E. oleifera*. A letra X em vermelho representa uma mutação *frame-shift*. Ambos os compostos chavicol e metil-chavicol são biossintetizados a partir do acetato *p*-cumaril, respectivamente. A espécie Americana apresenta uma mutação no gene funcional eugenol sintase (*EGS*) e chavicol O-metiltransferase (*COMT*), tornando-os inativos e incapazes de sintetizar o composto volátil metil-chavicol.

Para a verificação da mutação no gene estrutural na espécie Americana foram realizadas análises das regiões de pareamento dos *primers* AS16 e AS17, para amplificação do primeiro éxon do gene *COMT2* (*EO4N49844*) (Figura 12 e 14).



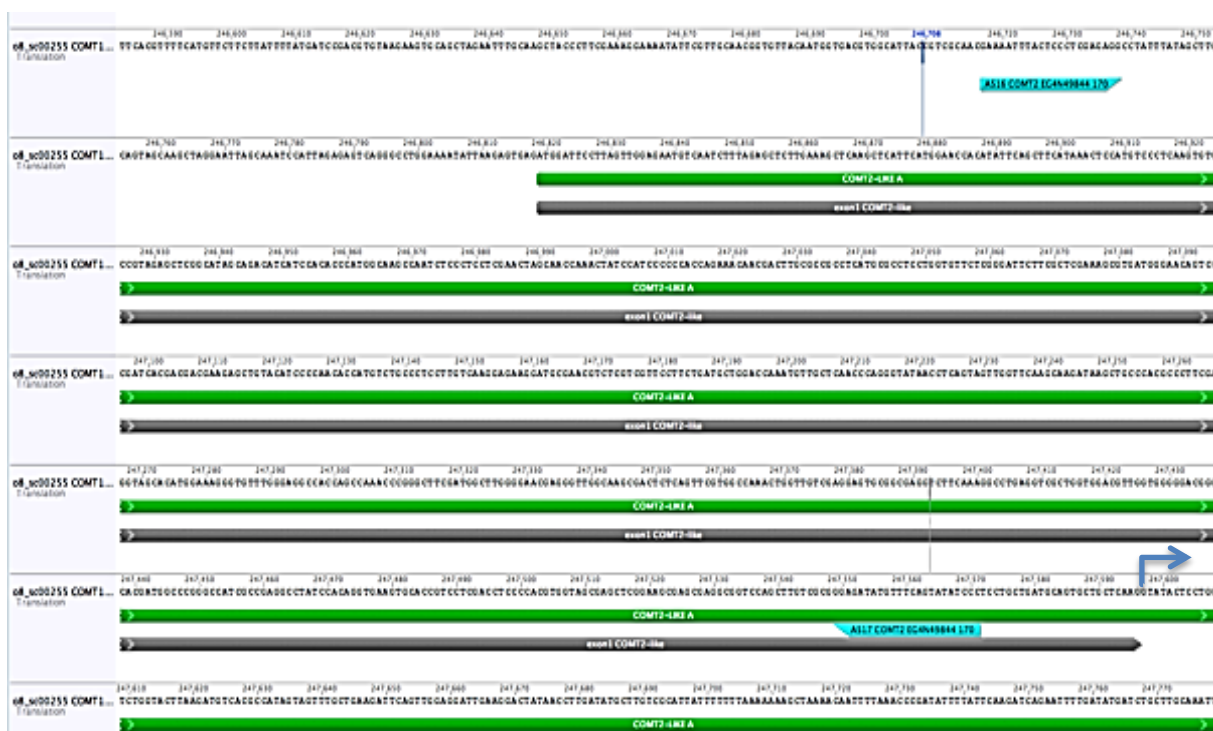


Figura 12: Desenho experimental dos *primers* AS16 e AS 17 no éxon 1 do gene *COMT2* em *E. oleifera* (EO). A seta (em azul) aponta para o início do segundo éxon, apresentando o *splincing* (AGGT).

Os *primers* AS16 e AS17 foram os sítios para a etapa de sequenciamento para comparar as espécies Africana e Americana. O *primer forward* AS16 está localizado a uma região não codificante do códon de iniciação ATG, já o *primer reverse* AS17 é complementar para as sequências do éxon 1. A região genômica de interesse foi amplificada utilizando o DNA genômico de *E. oleifera*, como modelo para os *primers* AS16 e AS17.

#### 4.1.3 Análise do DNA genômico

A extração do material genômico foi visualizada em gel de agarose 1% para estimar a concentração de DNA de cada amostra extraída (Figura 13). Foram observadas bandas definidas de DNA, fornecendo uma concentração aceitável para a análise da reação em cadeia da polimerase (PCR). Para as amostras de *E. oleifera*, apresentaram baixa concentração de material genético em relação a *E. guineensis*.

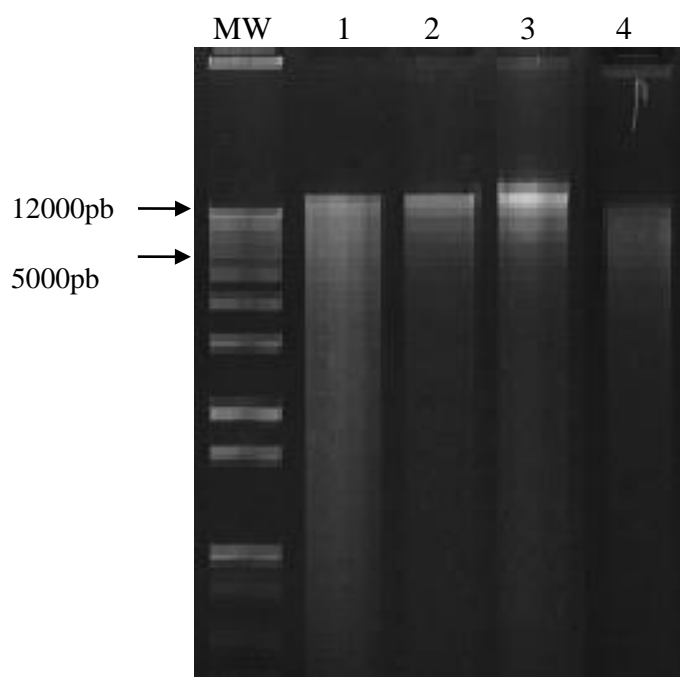


Figura 13: Eletroforese em gel de agarose 1% do DNA genômico de *E. guineensis*, *E. oleifera* e *O. basilicum*. Amostras no gel: (MW) Marcador de peso molecular 1Kb Plus DNA Ladder; (1) *O. basilicum*; (2 e 3) *E. guineensis*; (4) *E. oleifera*.

As amostras dos respectivos pares de *primers* identificados na reação de purificação do produto da PCR, corroboraram com o tamanho esperado de pares de bases (pb) (Figura 14). Os *primers* AS21/AS22 apresentam tamanho de 465pb e 900pb para os pares AS4/AS5, AS10/AS11, AS10/AS12, AS16/AS17, respectivamente.

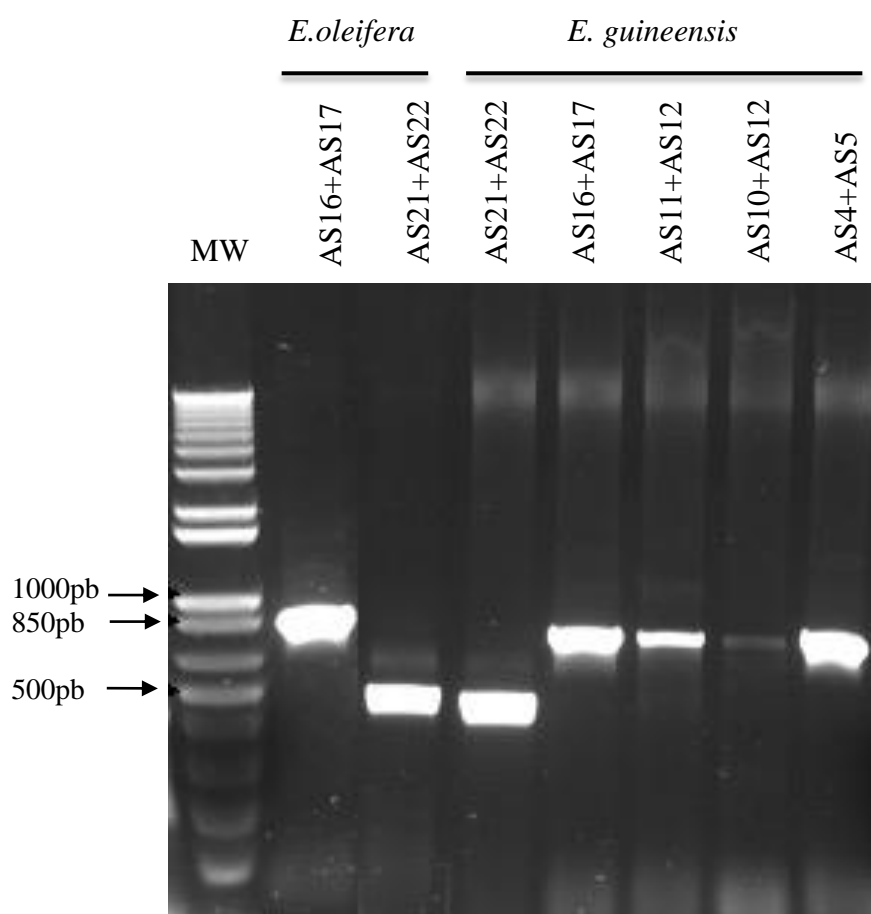


Figura 14: Visualização em gel de Agarose 2% do DNA genômico de *E. guineensis* e *E. oleifera*. MW: Marcador de peso molecular 1Kb Plus DNA Ladder (2 $\mu$ l).

Os pares de *primers* das amostras de *E. guineensis* (*tenera*) AS21/AS22; AS16/AS17; AS11/AS12; AS4/AS5 foram amplificados, conforme o esperado. As amostras de *E. oleifera*, apresentou baixa especificidade na etapa de extração do DNA, no entanto, observa-se, somente a amplificação de AS21/AS22 e AS 16/AS17. A não amplificação das amostras pode ser devido ao erro no desenho dos iniciadores, o que seria a mais provável causa da falha na reação de PCR para os homólogos de *E. oleifera*. Contudo, foram utilizados somente os *primers* AS16 e AS17 para a etapa de sequenciamento. Os fragmentos AS21/AS22, também foram sequenciados para o gene *COMT4* (EO4N49835), apresentou mutações, mas não foram confirmadas (dados não mostrados).

#### 4.1.4 Análises da reação de sequenciamento

A base de dados das sequências de *E. guineensis* e *E. oleifera* foram alinhadas e amplificadas pela PCR e, em seguida, sequenciadas pelo método de Sanger dideoxiterminal. Foram avaliados os três *indels* presentes no éxon 1 identificados na etapa de alinhamento para os *primers* AS16 e AS17 na espécie Americana.

A análise do indel 1, verificou a presença do nucleotídeo Citocina (C) (Figura 15a). Os alinhamentos dos indels 2 e 3 apresentam-se no meio da sequência, indicando a presença dos nucleotídeos Citocina (C) e Guanina (G), respectivamente (Figura 15b). O resultado do sequenciamento corrobora com a sequência genômica do banco de dados MPOB para a espécie Africana, indicando que ambas as espécies apresentam sequências idênticas para o gene avaliado *COMT2*.

As análises de alinhamentos das sequências gênicas na espécie Americana, mostraram a não ocorrência de mutações nos genes ortólogos. Tal deficiência poderia estar associada à mutações no gene estrutural ou no fluxo metabólico na biossíntese do MeC. Uma vez que as espécies Africana, Americana e o híbrido OxG, liberam quantidades semelhantes de MeC. A espécie Africana libera maiores quantidades de MeC que as demais, conseqüentemente, atrai maiores números de insetos polinizadores. A liberação do MeC na espécie Americana, pode não ser suficiente para atração de polinizadores (Gomes 2011).

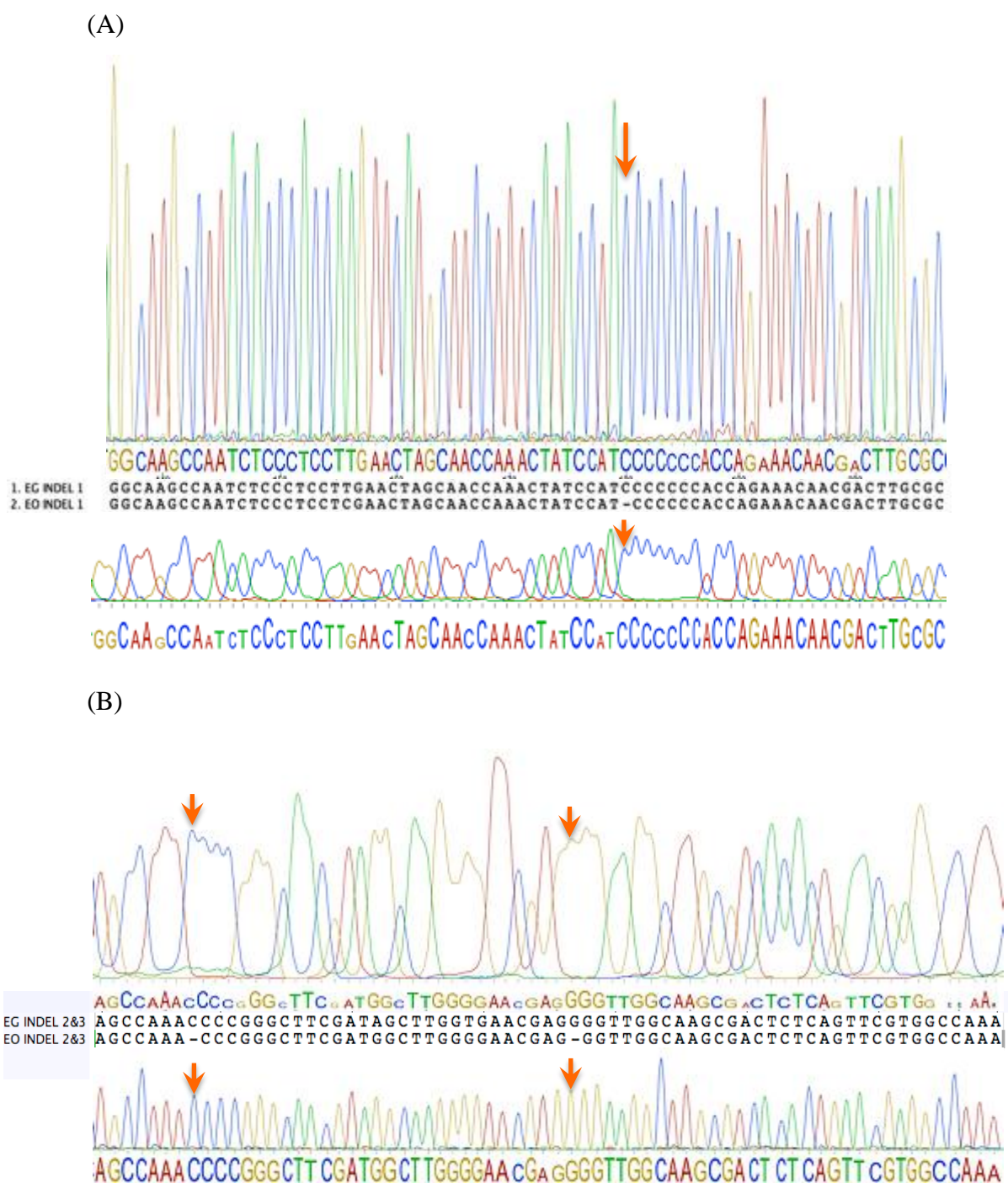


Figura 15: Imagem de cromatograma dos alinhamentos dos fragmentos gênicos amplificados pela PCR para os *primers* AS16 e AS17 do gene *COMT2* no éxon 1 em *E. oleifera* (EO). (A) Análise do *indel* 1, aponta (seta laranja) para a o nucleotídeo Citosina (C). (B) Alinhamento dos *indels* 2 e 3, apresentando os nucleotídeos Citosina (C) e Guanina (G).

Segundo Koeduka et al. (2009a) as flores de *Clarkia breweri*, emitem o composto volátil linalol, enquanto, que outra espécie *Clarkia concinnal* não é capaz de sintetizar esse composto. A razão para a falta da expressão do gene na síntese de linalol pode está relacionada a uma mutação no fator de transcrição do gene, o que poderia explicar a deficiência na produção de MeC na espécie Americana (Figura 16).

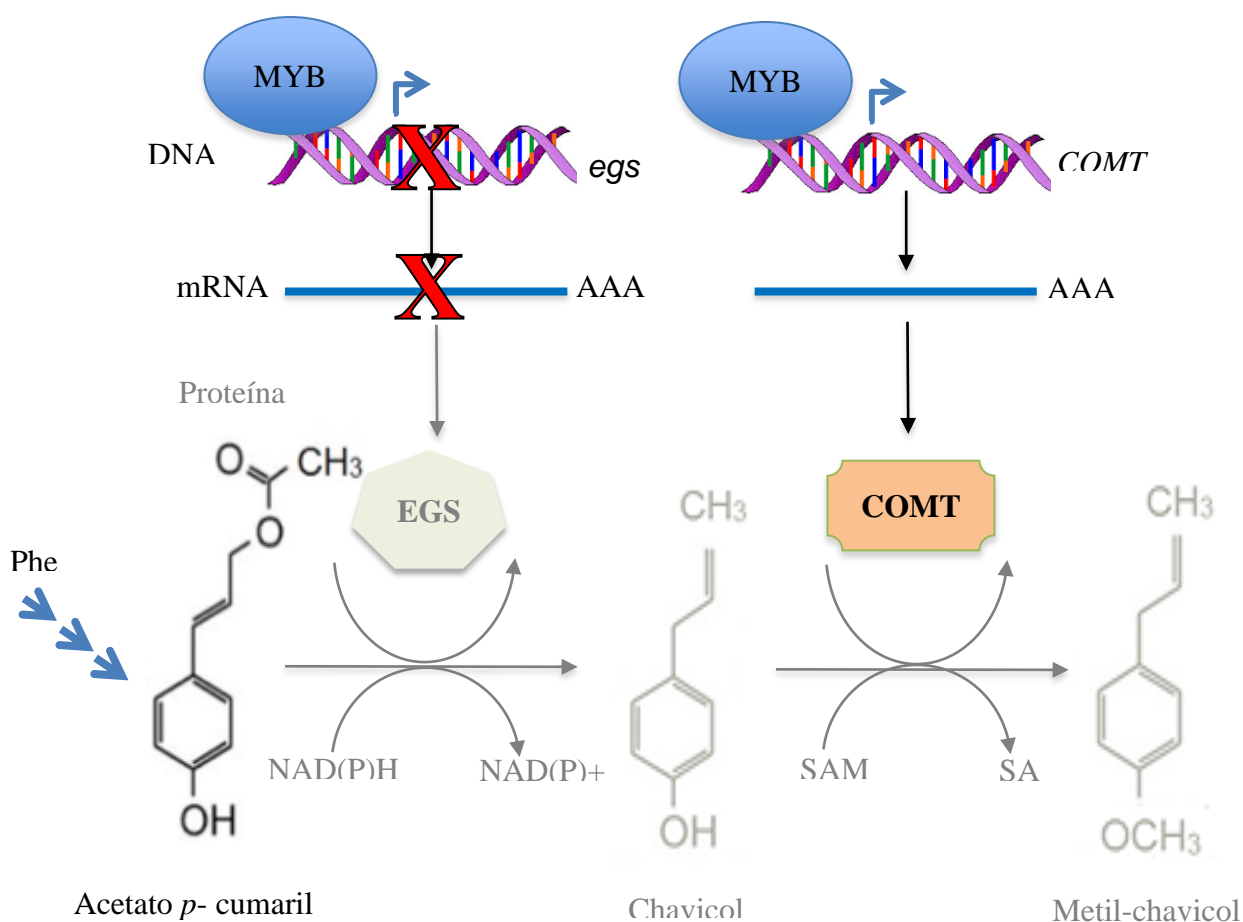


Figura 16: Deficiência de mutação no gene estrutural para a síntese de metilchavicol (MeC) em *E. oleifera*. O fator de transcrição MYB sintetiza o gene chavicol *O*-metiltransferase (*COMT*). A letra X em vermelho representa uma mutação *frame-shift* no gene eugenol sintase (*EGS*). A presença de alterações no gene *EGS*, impossibilita a formação do metil-chavicol.

Outro ponto de vista seria em relação à uma deficiência no fluxo metabólico (Figura 17). Estudos realizados por Kondo et al. (2007), verificaram na subespécie *Petunia axillaris paradii* a presença dos genes eugenol sintase funcional (*EGS1*) e eugenol *O*-

metiltransferase (*EOMTI*), no entanto, essas espécies não sintetizam o eugenol e o metil-eugenol, respectivamente. Os autores observaram uma competição metabólica de uma enzima desconhecida pelo acetato cumaril, e com isso o substrato é convertido para outro composto químico, assim as espécies sintetizam baixas concentrações ou são incapazes de produzir eugenol.

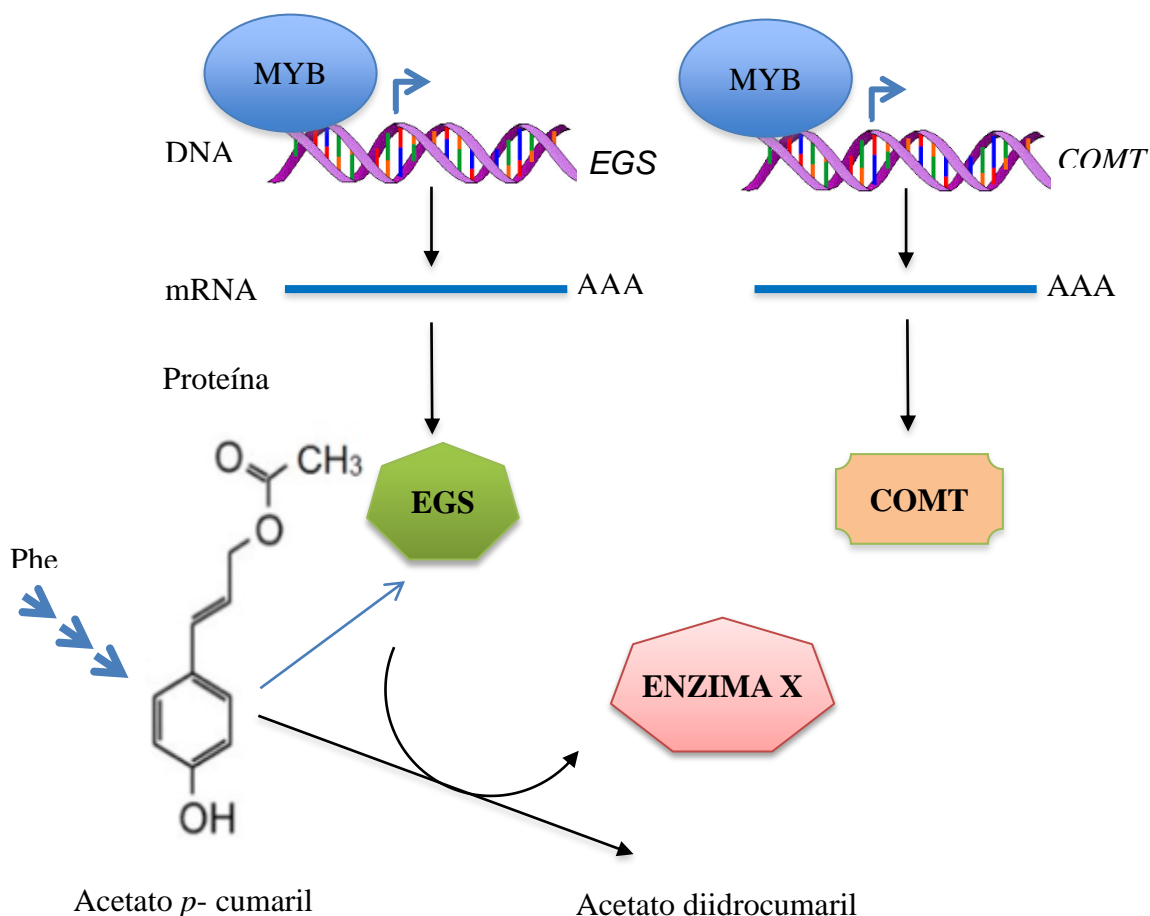


Figura 17: Deficiência no fluxo metabólico para a síntese de metilchavicol (MeC) em *E. oleifera*. O fator de transcrição MYB, inicia a síntese de metil-chavicol. A via metabólica sofre uma competição enzimática desconhecida (enzima X) com o eugenol sintase (EGS), pelo mesmo substrato acetato *p*-cumaril, sendo convertido para acetato diidrocumaril. A espessura das linhas do fluxo metabólico reflete a intensidade da reação.

Para confirmar os resultados do sequenciamento, foram realizadas análises de alinhamento obtidos do banco de dado genômico para as duas espécies de dendê no éxon 2 do gene *COMT2* (Figura 18). Foram incluídas sequências de cDNA, uma região não traduzida (UTR) na extremidade 3' e uma cauda poliA.

1. EG COMT2 EXON 2	1	10	20	30	40	50	60	70	80	90	100
Translation	W V L H D W S D E D C V N I L K Q C K K A V P S K E E G G K V I I I E M										
2. EO COMT2 cDNA	TGGGTGCTACATGATGGAGCGAGGACTCGCTAAAGATACAGAAACAAATGCAAGAGGCTGTCATCCAAAGAGAGGGGGGAAAGTGATCATTTGAAATG										
Translation	W V L H D W S D E D C V K I L K Q C K K A V P S K E E G G K V I I I E M										
3. EO COMT2 EXON 2	TGGGTGCTACATGATGGAGCGAGGACTCGCTAAAGATACAGAAACAAATGCAAGAGGCTGTCATCCAAAGAGAG-GGGGAAAGTGATCATTTGAAATG										
Translation	W V L H D W S D E D C V K I L K Q C K K A V P S K E E - G E R * S L L K W										
1. EG COMT2 EXON 2	110	120	130	140	150	160	170	180	190	200	210
Translation	V V D L D I G F P E L A E T Q L L F D M H M I V H T T G K Q R K E C E W										
2. EO COMT2 cDNA	TGGGTGCTACATGATGGAGCGAGGACTCGCTAAAGATACAGAAACAAATGCAAGAGGCTGTCATCCAAAGAGAGGGGGGAAAGTGATCATTTGAAATG										
Translation	V V D L D I G F P E L A E T Q L L F D M H M I V H T T G K Q R K E C E W										
3. EO COMT2 EXON 2	TGGGTGCTACATGATGGAGCGAGGACTCGCTAAAGATACAGAAACAAATGCAAGAGGCTGTCATCCAAAGAGAG-GGGGAAAGTGATCATTTGAAATG										
Translation	* L F W L L A F L S W L K H S Y S L T C T * * S T P Q G N K - E G H R H										
1. EG COMT2 EXON 2	220	230	240	250	260	270	280	290	300	310	320
Translation	K N F F T D A G F T G Y K I I P A L G V R S V I V A Y Y * TAGTGGATAGATAAGTGAAG										
2. EO COMT2 cDNA	AAAAAATTTTACTCATGAGGTTTACCGGCTACAAAATCATACCGGCATTAGCGGCGGCTCAGTTATCGTGGCTTATTTATGA										
Translation	K K I F T D A G F T D Y K I I P A L G V R S V I V A Y Y * TAGTGGATAGATAAGTGAAG										
3. EO COMT2 EXON 2	AAAAA--TTTACTCATGAGGTTTACCGGCTACAAAATCATACCGGCATTAGCGGCGGCTCAGTTATCGTGGCTTATTTATGA										
Translation	E K --I L L M Q V L P T T K S Y R H * A C V Q L S W L I I										
1. EG COMT2 EXON 2	330	340	350	360	370	380	390	400	410	420	430
Translation	TCACTACTAGTGGGCAATGATGACATTTGCTGCTTCTTCAAGTTTAACTTAGATCTTCATTATGTAACGCATCAATTATCCTAGTCTTATCGGTTACCTCAT										
2. EO COMT2 cDNA	TCACTACTAGTGGGCAATGATGACATTTGCTGCTTCTTCAAGTTTAACTTAGATCTTCATTATGTAACGCATCAATTATCCTAGTCTTATCGGTTACCTCAT										
Translation	TCACTACTAGTGGGCAATGATGACATTTGCTGCTTCTTCAAGTTTAACTTAGATCTTCATTATGTAACGCATCAATTATCCTAGTCTTATCGGTTACCTCAT										
1. EG COMT2 EXON 2	440	450	460	470	480	490	500	510	520	530	540
Translation	G TAGTAGAATCGGTGTGATGTTTGTATTGGTCTTCGTTGAGTTTGTGATCTATTGTGTAACGTATCAATTATCCATAGTCTTATCGGTTACCTCATG TAGTAGAATC										
2. EO COMT2 cDNA	G TAGTAGAATCGGTGTGATGTTTGTATTGGTCTTCGTTGAGTTTGTGATCTATTGTGTAACGTATCAATTATCCATAGTCTTATCGGTTACCTCATG TAGTAGAATC										
Translation	G TAGTAGAATCGGTGTGATGTTTGTATTGGTCTTCGTTGAGTTTGTGATCTATTGTGTAACGTATCAATTATCCATAGTCTTATCGGTTACCTCATG TAGTAGAATC										
1. EG COMT2 EXON 2	550	560	570	580	590	600	610	620	630	640	
Translation	A G T G T G A T G T T T G T A T T G G T C T T C G T T G A G T T T G T G A T C T A T C T G T C A T G C T A A G T T T G T T G G C C T T G T C A A G C T G A A A G A T A A A A A G T A C T C T T C A T T T G C C T C T										
2. EO COMT2 cDNA	A G T G T G A T G T T T G T A T T G G T C T T C G T T G A G T T T G T G A T C T A T C T G T C A T G C T A A G T T T G T T G G C C T T G T C A A G C T G A A A G A T A A A A A G T A C T C T T C A T T T G C C T C T										
Translation	A G T G T G A T G T T T G T A T T G G T C T T C G T T G A G T T T G T G A T C T A T C T G T C A T G C T A A G T T T G T T G G C C T T G T C A A G C T G A A A G A T A A A A A G T A C T C T T C A T T T G C C T C T										
3. EO COMT2 EXON 2	A G T G T G A T G T T T G T A T T G G T C T T C G T T G A G T T T G T G A T C T A T C T G T C A T G C T A A G T T T G T T G G C C T T G T C A A G C T G A A A G A T A A A A A G T A C T C T T C A T T T G C C T C T										
Translation	A G T G T G A T G T T T G T A T T G G T C T T C G T T G A G T T T G T G A T C T A T C T G T C A T G C T A A G T T T G T T G G C C T T G T C A A G C T G A A A G A T A A A A A G T A C T C T T C A T T T G C C T C T										
1. EG COMT2 EXON 2	650	660	667								
Translation	C T A G A A G A A A A A A A A A A										
2. EO COMT2 cDNA	C T A G A A G A A A A A A A A A A										
Translation	C T A G A A G A A A A A A A A A A										
3. EO COMT2 EXON 2	C T A G A A G A A A A A A A A A A										
Translation	C T A G A A G A A A A A A A A A A										

Figura 18: Diferença nas seqüências de DNA entre os acessos de bancos de dados genômicos. Os asteriscos (em laranja), indicam três *indels* na seqüência de *E. oleifera* (EO).

Os resultados indicam uma baixa qualidade da seqüência do genoma da espécie Americana, depositada no banco de dados do NCBI, entretanto, a seqüência da palma Africana, publicada no MPOB, apresenta uma alta qualidade, sendo considerada como um genoma de referência no respectivo estudo. Avaliações completas dos genes *COMT* e *EGS* são necessárias para um melhor entendimento da diferença de expressão entre as duas espécies na biossíntese do composto volátil MeC.



## 5. CONCLUSÃO

As análises de bioinformática sugerem que a espécie Africana sintetiza o metilchavicol (MeC) por uma via biossintética semelhante a do Manjeriço. O alinhamento das sequências publicadas entre as espécies Africana e Americana, indicam a ocorrência de mutações estruturais no gene *COMT*, o que poderia explicar a deficiência da síntese de MeC na espécie Americana. Entretanto, os resultados obtidos pelo sequenciamento em *E. oleífera*, indicam que esta espécie não apresenta mutações no fragmento do gene estudado que possam alterar a biossíntese de MeC.

Avaliações completas dos genes (*COMT* e *EGS*) são necessárias para uma melhor compreensão desta diferença de expressão em *E. oleífera*. As análises moleculares deste estudo estabelecem uma base experimental consistente para abordar questões futuras referentes à produção de frutos nas espécies de dendê, assim como o avanço e o conhecimento sobre as atividades enzimáticas relacionadas à biossíntese de MeC.

## 6. REFERÊNCIAS

ABDULLAH, N., YUSOP, M.R.; ITHNIN, M.; et al. Genetic variability of oil palm parental genotypes and performance of its' progenies as revealed by molecular markers and quantitative traits. **Comptes Rendus Biologies**, v. 334, n. 4, p. 290-299, 2011.

ADAM, H.J.,S.; JOUANNIE, S.; ESCOUTE, J.; et al. Reproductive developmental complexity in the African oil palm (*Elaeis guineensis*, Arecaceai). **American Journal of Botany**, v. 92, p. 1836-1852, 2005.

AKANI, G. C.; EBERE, N.; LUISELLI, L. et al.; Community structure and ecology of snakes in fields of oil palm trees (*Elaeis guineensis*) in the Niger Delta, southern Nigeria. **African Journal of Ecology**, v. 46, p. 500-506, 2007.

ALVES, S. M. A bioinformática e sua importância para a biologia molecular. **Revista Brasileira de Educação e Saúde**. 3: 18-25 p. 2013.

ARAÚJO, D. A. M. **Genômica e Bioinformática: Importância e perspectivas para o Nordeste**: Ciência e Cotidiano, v.1, p. 05-09, 2005.

ARIMURA, G.; OZAWA, R.; KUGIMIYA, S.; et al. Herbivore-induced defense response in a model legume. Two-spotted spider mites induce emission of (E)-beta-ocimene and transcript accumulation of (E)-beta-ocimene synthase in *Lotus japonicus*. **Plant Physiol**, v. 135, n. 4, p. 1976-83, 2004.

ARMBRUSTER, W.S. Floral specialization and angiosperm diversity: phenotypic divergence, fitness trade-offs and realized pollination accuracy. **AoB plants**, v.6, 2014.

ASEMOTA, O.; SHAH, F. H.; Detection of mesocarp oleoyl-thioesterase gene of the South American oil palm *Elaeis oleifera* by reverse transcriptase polymerase chain reaction. **African Journal of Biotechnology**, v. 3, p. 595-598, 2004.

ATKINSON, N.J.; URWIN, P.E.; The interaction of plant biotic and abiotic stresses: from genes to the field. **Journal of Experimental Botany**, v. 63, 2012.

AYALA, L. S. **Pesquisa sobre amarelecimento fatal do dendezeiro**. DENPASA. Belém-PA, p.319. 2001

BARCELOS, E.; AMBLARD, P.; BERTHAUD, J.; et al. Genetic diversity and relationship in American and African oil palm as revealed by RFLP and AFLP molecular markers. **Pesquisa Agropecuária Brasileira**, v. 37, p.1105-1114, 2002.

BARCELOS, E.; CUNHA, R.N.V.; NOUY, B. (2001) Recursos genéticos de dendê {*Elaeis guineensis* Jacq. e *E. oleifera* Kunth (Cortés)} disponíveis na Embrapa e sua utilização. In: MULLER, A.A. & FURLAN JUNIOR, J. (Eds.) *Agronegócio do dendê: uma alternativa social, econômica e ambiental para o desenvolvimento sustentável da Amazônia*. Belém-PA, Embrapa Amazônia Oriental.

BASTOS, T.X.; MULLER, A.A.; PACHECO, N.A.; et al. Zoneamento de riscos climáticos para a cultura do dendezeiro no estado do Pará. **Revista Brasileira de Agrometeorologia**, v. 9, p. 564-570, 2001.

BASIRON, Y.; Palm oil production through sustainable plantations. **European Journal of Lipid Science and Technology**, v.109, p.289-295, 2007.

BECKER, B. K. **Recuperação de áreas desflorestadas da Amazônia: será pertinente o cultivo da palma de óleo (Dendê)?** . **Revista Franco-Brasileira de Geografia**, v. 10, p. 01-20, 2010.

BERNARDES, M.S.R. Relatório de visitas à plantação de Paricatuba, na Denpasa, visando à identificação das causas do AF (1999). In: DENPASA. **Pesquisa sobre amarelecimento fatal**. Belém, PA, 2001.

BERTHAUD, A.; NUNES, C.D.M.; BARCELOS, E.; et al.; **Implantação e exploração da cultura do dendezeiro**. p.193-227, 2000.

BHORE, S.J.; SHAN, F.H.; Construction of efficient and effective transformation vectors for palmitoyl-acyl carrier protein thioesterase gene silencing in oil palm. **Bioinformation**, v. 6, n. 6, p. 212–220, 2011.

BOARI, A. J. **Estudos realizados sobre o Amarelecimento Fatal do dendezeiro (Elaeis guineensis Jacq.) no Brasil**. Embrapa Amazônia Oriental. Belém-PA, p.66, 2008

BIDRI, M., CONTI, M.; FRANETICH, J.F.; et al. Fresh aromatic herbs containing methylchavicol did not exhibit the pro-oxidative effects of pure methylchavicol on a human hepatoma cell line, HepG2. **Annales Pharmaceutiques Françaises**, v. 70, p. 256-263, 2012.

BRUCE, T.J.; WADHAMS, L.J.; WOODCOCK, C.M.; et al.; Insect host location: a volatile situation. **Trends Plant Sci**, v. 10, n. 6, p. 269-74, 2005.

CADENA, M.; DEVIS, M.; MALIKOV, I.; En Tumaco: fenómenos del niño y la niña. Eventos ENOS y su relación con el clima y la producción de palma de aceite. **Palmas**, v. 25, p. 46-62, 2004.

CANDWELL, R.W.; HUNT, D.; REID, A.; et al.; Insect pollination of oil palm - a comparison of the long term viability and sustainability of *Elaeidobius kamerunicus* in Papua New Guinea, Indonesia, Costa Rica, and Ghana, ASD Oil Palm Papers (Costa Rica), v.25, p.1-16, 2003.

CHANG, K. S.;AHN, Y. J. Fumigant activity of (E)-anethole identified in *Illicium verum* fruit against *Blattella germanica*. **Pest Manag Sci**, v. 58, n. 2, p. 161-6, 2002.

CHIA, G.S.; LOPES, R.; CUNHA, R.N.V.; et al. Repetibilidade da produção de cachos de híbridos interespecíficos entre o caiaué e o dendezeiro. **Acta Amazonica**, v. 39, p. 249-253, 2009.

CHINCHILLA, C. **Las pudriciones del cogollo en palma aceitera: la complejidad del desorden y una guía de convivencia**. Costa Rica: ASD-Palmatica: 22 p. 2010.

CONCEIÇÃO, H.E.O.; MULLER, A.A. **Botânica e morfologia do dendezeiro**. Embrapa Amazônia Oriental. Belém-PA, 2000.

CORLEY, R. H. V.; TINKER, P. **The oil palm**. Wiley. com, 2008.

CUNHA, R. N. V.; LOPES, R.; BARCELOS, E. **Domesticação e Melhoramento de caiaué**. p.275-296, 2009.

DOBSON, H.E.M. Relationship between floral fragrance composition and type of pollinator. In: Dudareva N, Pichersky E, eds. *Biology of floral scent*. Boca Raton, FL, USA: CRC Press, 147–198, 2006.

DRANSFIELD, J.; UHL, N.W.; ASMUSSEN, C.B.; et al.; *The Evolution and Classification of Palms*, Royal Botanic Gardens Kew, 2008.

DUDAREVA, N.; KLEMPIEN, A.; MUHLEMANN, J.K.; et al. Biosynthesis, function and metabolic engineering of plant volatile organic compounds. **New Phytologist**, v. 198, n. 1, p. 16-32, 2013.

DUDAREVA, N., MURFITT, L.M.; MANN, C.J.; et al. Developmental regulation of methyl benzoate biosynthesis and emission in snapdragon flowers. **Plant Cell**, v. 12, n. 6, p. 949-61, 2000.

E ESPINDOLA, F.S.; CALÁBRIA, L.K.; REZENDE, A.A.A.; et al. Recursos de bioinformática aplicados as ciências ômicas como genômica, transcriptômica, proteômica, interatômica e metabolômica. **Bioscience Journal**, v. 26, p. 463-477, 2010.

FAO (2002) Small scale palm oil processing in Africa. Agricultural services blulletin. Resources for Food and Agriculture, 148.

FAO. Conservation and management of pollinators for sustainable agriculture – the international response. In: FREITAS, B.M. & PEREIRA, J.O.P. (Ed.). Solitary bees: conservation, rearing and management for pollination. Fortaleza: Imprensa Universitária, p. 19-25, 2004.

FERREIRA, A.G.; AQUILA, M.E.A.; Alelopatia: uma área emergente da ecofisiologia. **Revista Brasileira de Fisiologia Vegetal**, v. 12, p. 175-204, 2000.

FERRER, J.L.; AUSTIN, M.B.; JÚNIOR STEWART, C.; et al.; Structure and function of enzymes involved in the biosynthesis of phenylpropanoids, **Plant Physiology and Biochemistry**, v. 46, p. 356-370, 2008.

FRANQUEVILLE, H. **La pudrición del cogollo de la palma aceitera em America Larina: revisión preliminar de hechos y fogros alcanzados**. CIRAD, p.41. 2001

GANG, D.R.; WANG, J.; DUDAREVA, N.; et al.; An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. **Plant Physiol**, v. 125, n. 2, p. 539-55, 2001.

GANG, D.R., LAVID, N.; ZUBIETA, C.; et al. Characterization of Phenylpropene O -Methyltransferases from Sweet Basil: Facile Change of Substrate Specificity and Convergent Evolution within a Plant O -Methyltransferase Family **The Plant Cell**, v. 14, p.505–519, 2002.

GOMES, S.M.S.; Polinizadores e semioquímicos do dendezeiro híbrido (*Elaeis oleifera* (H.B.K.) Cortés X *Elaeis guineensis* Jacq). Programa de Pós-Graduação em Entomologia, Universidade Federal de Viçosa, 65p, (2011).

GOINGUENÉ, S.P.; TURLINGS, T.C.J.; The Effects of abiotic factors on induced volatile emissions in corn plants1. **Plant Physiologists**, v. 129, p. 1296-130, 2002.

HAHN, W.J.; A phylogenetic analysis of the Arecoideae line of palms based on plastid DNA sequence data. **Molecular phylogenetics and evolution**, v. 23, n.2, p. 189-204, 2002.

HARDON, J.J.; CORLEY, R.H.V.; Pollination. In *Developments in Crop Science 1*. Corley, R.H.V., Hardon, J.J. and Wood, B.J (Eds). Elsevier, Amsterdam, p.532, 1976.

HOBALLAH, M.E.; GUBITZ, T.; STUURMAN, J.; et al. Single gene-mediated shift in pollinator attraction in *Petunia*. **The Plant cell**, v. 19, n.3, p.779-790, 2007.

HOMMA, A.K.O.; FURLAN-JÚNIOR, J.; CARVALHO, R.A.; et al. **Bases para uma política de desenvolvimento da cultura do dendezeiro na Amazônia**. Embrapa Amazônia Oriental / Manaus, p.11-30. 2000.

HORMAZA, P.; FUQUEN, E.M.; ROMERO, H.M.; Phenology of the oil palm interspecific hybrid *Elaeis oleifera* × *Elaeis guineensis*. **Scientia Agricola**, v. 69, p. 275-280, 2012.

HUSSEIN, M.Y.; LAJIS, N.H.; KINSON, A.; TEO, C.B.; Laboratory and field evaluation on the attractancy of *Elaeodorus kamerunicus* Faust to 4-allylanisole. *Porim Bulletin*, v. 18, p. 20-26, 1989.

KNUDSEN, J., ERIKSSON, R.; GERSHENZON, J.; et al. Diversity and distribution of floral scent. **The Botanical Review**, v. 72, n. 1, p. 1-120, 2006.

KOEDUKA, T.; ORLOVA, I.; BAIGA, T.J.; et al. The lack of floral synthesis and emission of isoeugenol in *Petunia axillaris* subsp. *parodii* is due to a mutation in the isoeugenol synthase gene. **The Plant journal: for cell and molecular biology**, v. 58, n. 6, p. 961-969, 2009a.

KOEDUKA, T.; BAIGA, T.J.; NOEL, J.P.; et al. Biosynthesis of t-anethole in anise: characterization of t-anol/isoeugenol synthase and an O-methyltransferase specific for a C7-C8 propenyl side chain. **Plant Physiol**, v. 149, n. 1, p. 384-94, 2009b.

KOEDUKA, T.; LOUIE, G.V.; ORLOVA, I.; et al. The multiple phenylpropene synthases in both *Clarkia breweri* and *Petunia hybrida* represent two distinct protein lineages. **Plant Journal**, v. 54, p.362–374, 2008.

KOEDUKA, T.; FRIDMAN, E.; GANG, D.R.; et al. Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. **Proc Natl Acad Sci USA**, v. 103, p. 10128–10133, 2006.

KONDO, M.; OYAMA-OKUBO, N.; SAGAE, M.; et al. Metabolic regulation of floral scent in *Petunia axillaris* lines: biosynthetic relationship between dihydroconiferyl acetate and iso-eugenol. **Bioscience, biotechnology, and biochemistry**, v. 71, n.2, p.458-463, 2007.

KRITCHEVSKY, D.; Impact of red palm oil an human nutrition and health. **Food Nutrition Bulletin**, n.21, p.182-188, 2000.

LAJIS, M. N.;HUSSEIN, M. Y.;TOIA, R. F. Extraction and identification of the main compound present in *Elaeis guineensis* flower volatiles. . **Pertanika**, v.. 8, p. 105-108, 1985.

LIJIMA, Y.; DAVIDOVICH-RIKANATI, R.; FRIDMAN, E.; et al. The biochemical and molecular basis for the divergent patterns in the biosynthesis of terpenes and phenylpropenes



in the peltate glands of three cultivars of basil. **Plant Physiology**, v. 136, p. 3724-3736, 2004.

LOUIE, G.V.; BAIGA, T.J.; BOWMAN, M.E.; et al. Structure and reaction mechanism of basil eugenol synthase. **PLoS One**, v. 2, p.993, 2007.

McCONKEY, M. F.; GERSHENZON, J.; CROTEAU, R. B. Developmental regulation of monoterpene biosynthesis in the glandular trichomes of peppermint. **Plant Physiology**, v.122, p. 215-224, 2000.

McDOWELL N.G.; BEERLING, D.J.; BRESHEARS, D.D.; et al. The interdependence of mechanisms underlying climate-driven vegetation mortality. **Trends Ecol Evol**, v. 26, n. 10, p. 523-32, 2011.

MOURA, J.I.L.; BENTO, J.M.S.; SOUZA, J.; et al. Captura de *Rhynchophorus palmarum* pelo uso de feromônio de agregação associado à árvore-armadilha mais inseticida. *Anais da Sociedade Entomológica da Brasil*, v. 26, p. 69-73, 1997.

MUHLEMANN, J. K.; KLEMPIEN, A.; DUDAREVA, N. Floral volatiles: from biosynthesis to function. **Plant, Cell and Environment**, v. 37, p. 1936–1949, 2014.

MULLER, A. A.; ALVES, R. M. **Avaliação de materiais genéticos de dendezeiro (*Elaeis guineensis* Jacq.) no Estado do Pará: uma alternativa de solução para o amarelecimento fatal**. Belém, PA, p.5. 1998.

MÜLLER, A. A.; ANDRADE, E. B. D. **Aspectos gerais sobre a fenologia da cultura da palma de óleo**. Embrapa Amazônia Oriental (CPATU), p.83-92, 2010.

NABHAN, G.P.; BUCHMANN, S.T.; Services provided by pollinator. In: G.C. Daily (ed.). *Nature's service: Societal dependence on natural ecosystems*. Washington, D.C., Island, 133-50, 1997.

NAGENDRAN, B.; UNNITHAN, U.R.; CHOO, Y.M.; et al.; Characteristics of red palm oil a carotene – and vitamin E – rich refined oil for food uses. **Food Nutrition Bulletin**, v.21, p.189-194, 2000.

NORIZZAH, A. R.; CHONG, C. L.; CHEOUW, C. S.; Effects of chemical interesterification on physicochemical properties of palm stearin and palm kernel olein blends. **Food Chemistry**, v.86, p.229-235, 2004.

PICHERSKY, E.; NOEL, J.P.; DUDAREVA, N. Biosynthesis of plant volatiles: nature's diversity and ingenuity. **Science**, v. 311, n. 5762, p. 808-11, 2006.

PINTO-ZEVALLOS, D.M.; MARTINS, C.B.C.; PELLEGRINO, A.C.; et al. Compostos orgânicos voláteis na defesa induzida das plantas contra insetos herbívoros. **Química Nova**, v. 36, p. 1395-1405, 2013.

PRÓCTOR, M. Y.; LACK., P.A.; The Natural History of pollination. London, 1996.

RAGUSO, R. A. **Advances in insect chemical ecology**. Cambridge University Press. 2004

REYES, R.; RIANO, N.; CHINCHILLA, C.; et al. **Physiological response of oil palm (*Elaeis guineensis*) to soil water availability under field conditions in Tumaco, Colombia**. Costa Rica: ASD OIL PALM PAPERS: 42 p. 2013.

RIOS, S. D. A.; CUNHA, R. N. V. D.; LOPES, R.; et al. **Recursos genéticos de palma de óleo (*Elaeis guineensis* Jacq.) e caiaué (*Elaeis oleifera* (H.B.K.) Cortês)**. Embrapa Amazônia Ocidental, p.39, 2012.

SANTOS, F. R. Sobre informática, genoma e ciência. **Ciência Hoje**, v.35, 2004.

SILVEIRA, N. J. F. **Bioinformática estrutural aplicada ao estudo de proteínas alvo do genoma do *Mycobacterium tuberculosis***. 2005. Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista.

SINGH, R.; ONG-ABDULLAH, M.; LOW, ENG-TI. L.; et al. Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. **Nature**, v. 500, n. 7462, p. 335-9, 2013.

SYED, R.A. Studies on oil palm pollination by insects. **Bulletin of Entomological Research**, v.69, p.213-224, 1979.

TAN, Y.; Weevil pollination in the *Elaeis oleifera* x *Elaeis guineensis* hybrid. Proceedings of the Symposium "Impact of the pollinating weevil on the Malaysian oil palm industry". Palm Oil Research Institute Malaysia, Kuala Lumpur, 1985.

TANDON, R.; MANOHARA, T.; NIJALINGAPPA, B.; et al. Pollination and Pollen-pistil Interaction in Oil Palm, *Elaeis guineensis*. **Annals of Botany**, v. 87, p. 831-838, 2001.

TORRES, G.A.; SARRIA, G.A.; VARON, E.; et al. First report of bud rot caused by *Phytophthora palmivora* on African oil palm in Colombia. **Plant Disease**, v.. 94, p.1163-1163, 2010.

TRINDADE, D. R. P., L.S.; FURLAN-JUNIOR, J. **Abordagem sobre o estado atual das pesquisas para a identificação do agente causal do amarelecimento fatal do dendezeiro**. Embrapa Amazônia Oriental. Belém-PA, p.439-450, 2005.

TURNER, P.D.; GILBANKS, R.A. Oil palm cultivation and management. Incorporated Society of Planters. Kuala Lumpur, 672p, 1974.

VAN SLOBBE, W.G.; SOUZA, R.L.R.; **Amarillamiento fatal o pudrición de cogollo en Denpasa - Brasil**. Palmas, v. 12, p.17-23, 1991.

VASSAO, D. G., GANG, D.R.; KOEDUKA, T.; et al. Chavicol formation in sweet basil (*Ocimum basilicum*): cleavage of an esterified C9 hydroxyl group with NAD(P)H-dependent reduction. **Org Biomol Chem**, v. 4, n. 14, p. 2733-44, 2006.

VEIGA, L.S.; FURIA, L.R.R.; **Avaliação do dendezeiro como opção para o sequestro de carbono na Amazônia**. Belém-PA, p.125-144, 2000.

WANG, J.; PICHERSKY, E.; Characterization of S-Adenosyl-l-Methionine:(Iso)eugenol O-Methyltransferase Involved in Floral Scent Production in *Clarkia breweri*. **Archives of Biochemistry and Biophysics**, v. 349, n. 1, p. 153-160, 1998.

WIECZOREK, E.M.; LEAL, E.; Caminhos e tendências do uso de banco de dados em bioinformática. <http://www.wieczorek.com.br/publicacoes>, 2008. Acesso em: 25 de fevereiro de 2014.

ZEVEN, A.C.; The origin of the oil palm, Institute of Oil Palm Research, 1965.

## **ANÉXOS**

## ANEXO A

1. Sequências consenso de COMT e EGS selecionadas e depositada no *GenBank* de *Ocimum basilicum* L.

---

>sp|Q93WU3|CVMT1\_OCIBA Chavicol O-methyltransferase  
 MALQNMDISLSTEQLLQAQAHVWNHMYAFANSMSLKCAIQLGIPDILHKHDHPMT  
 LSQLLKAIPINKEKSQSFQRLMRALVNSNFFIEENSNNQEVCYWLTASRLLLKGAP  
 LTVAPLVQVVLDPTFTNPWHYMSEWFKHENHATQFEAANGCTFWEKLANKPSMG  
 RFFDEAMSCDSRLVAHVLT KDYKHVIDGIRTLVDVGGNGTMAKAIVEAVPTMKC  
 TVLDLPHVVAGLESTDKLSYIGGDMFQSIPSADAILLKFIHDWDDEEGLKILKRCKD  
 AVGIGGKVIIDVVVGVNHDVDEVLEDQLHFDMMAMMSYFNAKERTMNEWKELISA  
 AGFTSYKLTPAFGVRSLEIAYP

---

>gi|87044868|gb|ABD17321.1| eugenol synthase 1 [*Ocimum basilicum*]

MEENGMKSKILIFGGTGYIGNHVMKGSCLKGHPTYVFTRPNSSKTTLLDEFQSLGAI  
 VKGELDEHEKLVLMKKVDVVISALAFPQILDQFKILEAIKVAGNIKRFLPSDFGVEE  
 DRINALPPFEALIERKRMIRRAIEEANIPYTYVSANCFASYFINYLRLPYDPKDEITVY  
 GTGEAKFAMNYEQDIGLYTIKVATDPRALNRVVIYRPSTNIITQLELISRWEKKIGKK  
 FKKIHPVEEEIVALTKEPPEPIAILHCLFIDGATMSYDFKENDVEASTLYPELKFT  
 TIDELLDIFVHDPPPPASAAF

---

Gang et al. (2002); Koeduka et al. (2006)

## ANEXO B

2. Sequências de proteínas de COMT de *E. guineensis* depositada no banco de dados MPOB.

---

>EG4P993 O-methyltransferase family 2 protein, putative  
 MDQLQDKAQLWNHALGFIRSMSFKCAIELGISDVLHNNHGKPITLSELATFLSIPPSKTPALGS  
 LMRLLVHSGVFASHPQQGGGEEGYFLTPTSELLVKEEKACVSPFALLILNWTMLVPGHQL  
 GEWFMAGAAAETPFDMAHGKGIFGETNVRADFNELFNEAMASDARLVTEVALLQRQWEM  
 LRGVRSVVDVGGGTGTVAKSIKALPEVRCMVLDLPHVVATMDEEQREGVEFIGGDMFEH  
 IPPADAVLLKWLHDWSDEHCILKHCKEAIPTKKDGGKYSPQLASIMELTNGQSFGKPFQ  
 AHTLLLNQSFSTAGLMSLNCVELGIADLIHSHGGPMPLSELARSIPPEKAPLSAASCVSS

---

>EG4P49835 O-methyltransferase family 2 protein, putative  
 MESLDGECQSLELLKAQAHAWNHIHFSFINSMSLKCAVELGIADIIHAHGKPISSLELATKLSIP  
 PTRNNDLRRLMRLLVFSGFFARERDGNSHDHDEELYIPTTMSALLVNQKDVNVSSFALVML  
 DQMFLHPWYHLSSWLKQDMLPATPFEEAHGKGVWEATSQSPAFNSLLNEGMASDARFVA  
 KVVVEECGEIFRGLSSLVDVGGGTGTMARAIADAYPHVKCTVLDLPQVVASSEAGETVQLV  
 AGDMFQYVPPADAVLLKWLHDWSDEECVKILKRCKEAVPSKEKGGKALEWWFT  
 SLVSVRIPTERFCS

---

>EG4P49884 O-methyltransferase family 2 protein, putative  
 MDSL VGECQSLELLKAQAHWSHIFSFINSMSLKCAVELGIADIIHGHGKPISSLELATKLSIPP  
 TRNNDLRRLMRLLVFSGFLFGRKRDNVSPDHDDEELYIPNTMSALLVKEKDANVSSFLMLD  
 QMLLNWPWYNLSSWFKQDKLPTPFVAHGKGVWEATSQTPGFDSLVEGLASDSQFVAKLV  
 VEECGEVFKGLRSLVDVGGGTGTMARAIADAYPHVKCTVLDLPQVVASSEASEAVQLVAG  
 DMFQYVPPADAVLLKWLHDWSDEDCVNILKQCKKAVPSKEEGGKKYFGKLSDE  
 SSLTVFGGRGSISGYRRIDFSLLIYSDGDRSSMIGLYVQA

---

>EG4P49859 O-methyltransferase family 2 protein, putative  
 MKERREPLEPSLWCYLPQLAIHKRLRAWKKSASVMESLDGECQSLELLKAQAHAWNHF  
 FINMSMLKCAVELGIADIIHAHGKPTSLELATKLSIPPTRNNDLRRLMRLLVFSGFFARERD  
 NSHDHDEELYIPTTMSALLVNQKDVNVSSFALAMLDQMFLHPWYHLSSWLKQDNLPTPL  
 EAHGKGVWEATSQSPAFNSLLNEGMASDARFVAKVVVEECGEIFRGLSSLVDVGGGTGT  
 MARAIADAYPHVKCTVLDLPQVVASSEAGETVQLVAGDMFQYVPPADAVLLKVIKRT  
 DCMLNQEWHHYPTK

---

>EG4P49850 O-methyltransferase family 2 protein, putative  
 MESLDGECQSLELLKAQAHAWNHIHFSFIKSMMLKCAVELGIADIIHGHGKPISSLELATKLSIP  
 PTKNNDLRRLMRLLVFSGFFARKRDGNSHDHDEELYIPTTMSALLVKENDANVSSFVLAILD  
 QTLFNPWYHLSSWFKQDKVPATPFVAHGRDVWEAISRSPAFNSLCYEAMASDSRFVAKL  
 VVEECGEVFRGLRSLVDVGGGKGTMARAIADAYPHVKCTVLDLPQVVASPEASEEVQLVA  
 GDMFQYVPPADAVLLKNQIGSTLSTSHATITITSFSAHYIYIGISGLCSVGSKLAITVASR  
 KASSSLVTAR

---

>EG4P99510 O-methyltransferase family 2 protein, putative  
 MEGRTMGEDILEELEAQGNVWNHTFSFVNSMSMLKCAVELGIPDAIHGHGSPVTLSELSAML  
 SLPPRRTAALRRLMRMLVHSGCFAKDVVGDGEEGAYVLT PFSRLLLKEKRTSVSAYVLGM  
 LDPLLIQTWHSLSAWFHGEERTPFRLAHGKGIFEAAAERPELASFLNAAMASDSQLVGEVM  
 VRECSEAFKGVKSLVDVGGCTGTMAMIIAEAFDPDIKCTVFDLPHIVAAAAAEAKTPAAVDF  
 VGGDMFERVPPADVALLKVSISLRLGLQTKVGMRSDSGMNRREPVRSLSGIGRVAVIG  
 WARDMRLPSNLL

---

3. Sequências de proteínas de EGS de *E. guineensis* no banco de dados MPOB.

---

>EG4P155924

MAGKMSKILIIGGTGYIGKFVVGASARSGHPTFALVRDTPSDPAKAQLLETFKNSG  
VTLLKGDLDNDHGSLVEAIKQVDVVISTVGLQQVMDQLKIIAAIKEVGTIKRFLPSEF  
GLDVDRIHKAEPASSVFALKAQIRRAIEAEGIAHTIVCCNCFAGYFLPTLAQPEPPKD  
KVTILGDGNQKVVFLAEEDVGTYTIKTVDDPRTLKILYVKPPANICSFNQLVSLWE  
KKTGKTLEKIYLPPEEEVLKKIQEAPFPFNMLMAVNYLVFVKGDTANIQIEPSFGVEA  
SKLYPDVKYTSVQEYLNRVRYGNM

---

>EG4P76665

MAGKMSKILIIGGTGYLGKFIVGASARSGHPTFVLVRDTAPSDQAKAQLLRAFKNSS  
VTLLKGDLDNDHGSLVEAMKQVDVVISTIGPRQVLDQFKIIAAIKEVGTIKREKIKREN  
NVKREKSHQGRLKGGKTTEVLALGPHPQEAPRSCASGIAVQPPPPASGVPVQASNLR  
LRRPDPPPASTDVDRIHTAEPASSVFALKAQIRRAIEAEGIPCTIVCCNCFAGYFLPTL  
AQPEPPKDKVTVLGDGNPKAVFLAEEDIGTYTIKTVDDPRTLKILYVKPPANICSF  
NDLVSLWEKKTGKTLEKIYLPPEEEVLKRIQEAPFPNILMAVNYLVFVKGDTANIEI  
DPSFGVEASELYPDDLPAHERKGLHSLLLDNLFAIPSVMEKFGGGFRVEAAELYPD  
VKHTIFDEYL NRCIQ

---

>EG4P76667

MAREMSKILIIGGTGYLGKFIVGASARSGHPTFAMVRDTAPSDPAKAQLLEAFKNSG  
VTLFKGDLDNDHGSLVEAIKQVDVVISTVGLQQVMDQLKIIAAIKEVGTIKRFLPSEFG  
LDVDRIHKAEPASSVFALKAQIRRAIEAEGIPHTIVCCNCFAGYFLPTLAQPEPPKDK  
VTILGDGNQKVVFLAEEDVGTYTIKTVDDPRTLKTLKPPANIYSFNDLVSLWE  
KKTGKTLEKIYLPPEEEALKKIEEAPFPFNMLMAVNYLVFVKGDTANIQIEPSFGVEA  
SELYPDVKYTSVQEYLNRVIQ

---



## ANEXO C

4. Alinhamentos das sequências proteicas de *E. guineensis* e *O. basilicum* L. do gene *COMT2*.

```

EG4P49835 -----MESLDGECQSLELLKA 16
EG4P49859 MKERREPLEPSLWCYLPQLAIHKRLRAWKKSASVMESLDGECQSLELLKA 50
EG4P49884 -----MDSLVGECQSLELLKA 16
EG4P49850 -----MESLDGECQSLELLKA 16
EG4P993 -----MDQLQD 6
EG4P99510 -----MEGRMTMGEDILEELEA 16
sp|Q93WU3|CVMT1_OCIBA -----MALQNMDISLSTEQLLQA 18
: *
:

EG4P49835 QAHAWNHI F S F I N S M S L K C A V E L G I A D I I H A H G K P I S L L E L A T K L S I P P T 66
EG4P49859 QAHAWNHI F S F I N S M S L K C A V E L G I A D I I H A H G K P T S L L E L A T K L S I P P T 100
EG4P49884 QAHWSH I F S F I N S M S L K C A V E L G I A D I I H T H G K P I S L L E L A T K L S I P P T 66
EG4P49850 QAHAWNHI F S F I K S M S L K C A V E L G I A D I I H T H D K P I S L T E L A A K L S I P P T 66
EG4P993 K A Q L W N H A L G F I R S M S F K C A I E L G I S D V L H N H G K P I T L S E L A T F L S I P P S 56
EG4P99510 Q G N V W N H T F S F V N S M S L K C A V E L G I P D A I H G H G S P V T L S E L S A M L S L P P R 66
sp|Q93WU3|CVMT1_OCIBA Q A H V W N H M Y A F A N S M S L K C A I Q L G I P D I L H K H D H P M T L S Q L L K A I P I N K E 68
: : * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
: :

EG4P49835 R N N D L R R L M R L L V F S G F F A R E R D G N S H D H D - E E L Y I P T T M S A L L V N Q K D V 115
EG4P49859 R N N D L R R L M R L L V F S G F F A R E R D G N S H D H D - E E L Y I P T T M S A L L V N Q K D V 149
EG4P49884 R N N D L R R L M R L L V F S G L F G R K R D G N S P D H D D E E L Y I P N T M S A L L V K E K D A 116
EG4P49850 K N N D L R R L M R L L V F S G F F A R K R D G N S H D H D - E E L Y I P T T M S A L L V K E N D A 115
EG4P993 K T P A L G S L M R L L V H S G V F A S H P Q --- Q G G E E E G Y F L T P T S E L L V K E E K A 103
EG4P99510 R T A A L R R L M R M L V H S G C F A K D V V G --- D G E - E G A Y V L T P F S R L L L K E K R T 112
sp|Q93WU3|CVMT1_OCIBA K S Q S F Q R L M R A L V N S N F F I E E N S --- N N Q E V C Y W L T P A S R L L L K G A P L 113
: . : * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
: . :

EG4P49835 N V S S F A L V M L D Q M F L H P W Y H L S S W L K Q D M L P A T P F E A A H G K G V W E A T S Q S 165
EG4P49859 N V S S F A L A M L D Q M F L H P W Y H L S S W L K Q D N L P A T P L E A A H G K G V W E A T S Q S 199
EG4P49884 N V S S F - L L M L D Q M L L N P W Y N L S S W F K D K L P - T P F E V A H G K G V W E A T S Q T 164
EG4P49850 N V S S F V L A I L D Q T L F N P W Y H L S S W F K D K V P A T P F E V A H G R D V W E A I S R S 165
EG4P993 C V S P F A L L I L N W T M L V P G H Q L G E W F M A G A A A E T P F D M A H G K G I F G E T N V R 153
EG4P99510 S V S A Y V L G M L D P L L I Q T W H S L S A W F H -- G E E R T P F R L A H G K G I F E A A A E R 160
sp|Q93WU3|CVMT1_OCIBA T V A P L V Q V V L D P T F T N P W H Y M S E W F K H E N H A - T Q F E A A N G C T F W E K L A N K 162
* : . : * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
: . :

EG4P49835 P A F N S L L N E G M A S D A R F V A K V V V E E C G - E I F R G L S S L V D V G G G T G T M A R A 214
EG4P49859 P A F N S L L N E G M A S D A R F V A K V V V E E C G - E I F R G L S S L V D V G G G T G T M A R A 248
EG4P49884 P G P D S L V N E G L A S D S Q F V A K L V V E E C G - E V F K G L R S L V D V G G G T G T M A R A 213
EG4P49850 P A F N S L C Y E A M A S D S R F V A K L V V E E C G - E V F R G L R S L V D V G G G K G T M A R A 214
EG4P993 A D F N E L F N E A M A S D A R L V T E V A L L Q R Q W E M L R G V R S V V D V G G G T G T V A K S 203
EG4P99510 P E L A S F L N A A M A S D S Q L V G E V M V R E C S - E A F K G V K S L V D V G G C T G T M A M I 209
sp|Q93WU3|CVMT1_OCIBA E S M G R F F D E A M S C D S R L V A H V L T K D Y K - H V I D G I R T L V D V G G N G T M A K A 211
. : : . : . : * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
: . :

EG4P49835 I A D A Y P H V K C T V L D L P Q V V A S S -- E A G E -- T V Q L V A G D M F Q Y V P P A D A V L 260
EG4P49859 I A D D Y P H V K C T V L D L P Q V V A S S -- E A G E -- T V Q L V A G D M F Q Y V P P A D A V L 294
EG4P49884 I A E A Y P Q V K C T V L D L P H V V A S S -- E A S E -- A V Q L V A G D M F Q Y I P P A D A V L 259
EG4P49850 I A V A Y P Q V K C T V L D L P H V V A S P -- E A S E -- E V Q L V A G D M F Q Y V P P A D A V L 260
EG4P993 I A K A L P E V R C M V L D L P H V V A T M D E E Q R E -- G V E F I G G D M F E H I P P A D A V L 251
EG4P99510 I A E A F P D I K C T V F D L P H I V A A A A A E A K T P A A A V D F V G G D M F E R V P P A D V A L 259
sp|Q93WU3|CVMT1_OCIBA I V E A V P T M K C T V L D L P H V V A G L -- E S T D -- K L S Y I G G D M F Q S I P S A D A I L 257
* . : * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
: . :

EG4P49835 L K W V L H D W S D E C V K I L K R C K E A V P S K E K G G K --- 292
EG4P49859 L --- K V I K R T D C M L N --- 306
EG4P49884 L K W V L H D W S D E D C V N I L K Q C K A V P S K E E G G K Y F G K L S D E S S L T --- 304
EG4P49850 L K N Q I G S T L S T S H A T I T I T S F S A H Y Y I G I S G --- 292
EG4P993 L K W I L H D W S D E H C I K I L K H C K E A I P T K K D G G K Y S P Q L A S I M E L T N G Q S F G 301
EG4P99510 L K V S I S L R L G - L Q T K V G M R S D S G M N R R E P V R --- 289
sp|Q93WU3|CVMT1_OCIBA L K P I I H D W D D E G L K I L K R C K D A V G I G G K V I I D V V V G V N --- 297
* . :
: . :

EG4P49835 -----ALEWVFTSLVSVRIPTERFCS----- 313
EG4P49859 -----QEWHHY-----PTK----- 315
EG4P49884 -----VFGGRGSI SGYRRIDFSLLIYSDGDRSSSMIGLYVQA- 341
EG4P49850 -----LCSVGSKLAI TVASRKASSSLVTAR--- 317
EG4P993 K P F Q A H T L L N Q S F S T A G L M S L N C A V E L G I A D L I H S H G G P M P L S E L A R S I 351
EG4P99510 -----LSGIGRVAVI GWARDMRLPSNLL--- 312
sp|Q93WU3|CVMT1_OCIBA -----HDVDEVLEDQLHFDMMAMSYFNAKERTMNEWELISA 334

EG4P49835 -----
EG4P49859 -----
EG4P49884 -----
EG4P49850 -----
EG4P993 P I P P E K A P L S A A S C V S S ----- 368
EG4P99510 -----
sp|Q93WU3|CVMT1_OCIBA A G F T S Y K L T P A F G V R S L I E A Y P 356

```

5. Alinhamento das sequências proteicas de *E. guineensis* e *O. basilicum* L. do gene *EGS*.

```

EG4P155924      --MAGKMSKILIIGGTGYIGKFVVGASARSGHPTFALVRDTPSPDPAKAQLLETFKNSGV 58
EG4P76667      --MAREMSKILIIGGTGYLGKFIVGASARSGHPTFAMVRDTAPSDPAKAQLLEAFKNSGV 58
EG4P76665      --MAGKMSKILIIGGTGYLGKFIVGASARSGHPTFVLVRDTAPSDQAKAQLLRAFKNSSV 58
gi|87044868|   MEENGMSKILIFGGTGYIGNHMVKGLKLGHPYVVFTRPNS----SKTTLLEDFQSLGA 56
                *****:*****:*.:. * . * : *****:..* .: .: * : * * * :. . .

EG4P155924      TLLKGDLDNDHGSLVEAIKQVDVVISTVGLQQVMDQLKIIAAIKEVGTIKR----- 108
EG4P76667      TLFKGDLDNDHGSLVEAIKQVDVVISTVGLQQVMDQLKIIAAIKEVGTIKR----- 108
EG4P76665      TLLKGDLDNDHGSLVEAMKQVDVVISTIGPRQVLDQFKIIAAIKEVGTIKREKIKRENNVK 118
gi|87044868|   IIVKGELDEHEKLVLEMKKVDVVISALAFPQLLDQFKILEAIKVAGNIKR----- 106
                :.:.:*.:. * . * * :*:*****:.. * :*:*.:. * * * . * . * *

EG4P155924      -----FLPSEFGL----- 116
EG4P76667      -----FLPSEFGL----- 116
EG4P76665      REKSHQGRRLKGTTEVLALGPHFQEAPRSCASGIAVQPPPPASGVPVQASNLRLRRPDP 178
gi|87044868|   -----FLPSDFGV----- 114
                . . * : :

EG4P155924      ----DVDRIHKAEPASSVFALKAQIRRAIEAEGIAHTIVCCNCFAGYFLPTLAQPEPPKD 172
EG4P76667      ----DVDRIHKAEPASSVFALKAQIRRAIEAEGIPHTIVCCNCFAGYFLPTLAQPEPPKD 172
EG4P76665      PASTDVDRIHTAEPASSVFALKAQIRRAIEAEGIPCTIVCCNCFAGYFLPTLAQPEPPKD 238
gi|87044868|   ----EEDRINALPPEALIERKRMIRRAIEANIPYTYVSANCFASYFINYLRLRPYDPKD 170
                : * * * : * . : : * * * * * . * . * * . * * * * * * * * * *

EG4P155924      KVTILGDGNQKVVFLAEEDVGYTIKTVDPRTLNKILYVKPPANICSFNQLVSLWEKKT 232
EG4P76667      KVTILGDGNQKVVFLAEEDVGYTIKTVDPRTLNKILYVKPPANICSFNQLVSLWEKKT 232
EG4P76665      KVTVLGDGNPKAVFLAEEDIGTYTIKTVDPRTLNKILYVKPPANICSFNQLVSLWEKKT 298
gi|87044868|   EITVYGTGEAKFAMNYEQDIGLYTIKVATDPRALNRVVIYRPNSTNIIITQLELISRWEKKI 230
                :.:.: * * : * . : * : * * * * . * * * * * : : * : * * * : : * * * *

EG4P155924      GKTLEKIYLPPEEVLKRIQEAPFPFNMLMAVNLYLVFKGDTANIQIEPSFGVEASELYP- 291
EG4P76667      GKTLEKIYLPPEEVLKRIQEAPFPFNMLMAVNLYLVFKGDTANIQIEPSFGVEASELYP- 291
EG4P76665      GKTLEKIYLPPEEVLKRIQEAPFPFNILMAVNLYLVFKGDTANIEIDPSFGVEASELYPD 358
gi|87044868|   GKKFKKIHVPEEIVALTKEPPEPENIPIAILHCLFIDGATMSYDFKEND-VEASTLYP- 288
                **.:*:*:*:* * : * * * * : * : : * : * * * . : . . * * * * *

EG4P155924      -----DVKYTSVQEYLNVRVYGN 309
EG4P76667      -----DVKYTSVQEYLNVRVIQ-- 307
EG4P76665      DLPASHERKGLHSLLDNLFAIPSVMEKFGGGFRVEAAELYPDVKHTIFDEYLNRCIQ-- 416
gi|87044868|   -----ELKFTTIDELLDIFVHDP 306
                :*.* .:* * :

EG4P155924      GM----- 311
EG4P76667      -----
EG4P76665      -----
gi|87044868|   PPPASAAF 314

```

## ANEXO D

6. Alinhamentos das sequências genômicas do gene *COMT2* de *E. guineensis* e *E. oleifera*.6.1. EG4P49835: Deleção no gene *E. oleifera* (1pb em vermelho).

EG4N49835E.guineensis	--ATGGAATCCTTGGATGGAGAATGTCAATCTTTGGAGCTCCTGAAAGCTCAAGCCCAT	57
EG4N49835E.oleifera	ATGGAGTACTCCTTGGATGGAGAATGTCAATCTTTGGAGCTCCTGAAAGCTCAAGCCCAT	60
	* * *	
EG4N49835E.guineensis	GCATGGAACCACATATTTAGCTTCATAAACTCCATGTCCCTCAAGTGTGCCGTGGAGCTC	117
EG4N49835E.oleifera	GCATGGAACCACATATTTAGCTTCATAAACTCCATGTCCCTCAAGTGTGCCGTGGAGCTC	120
	*****	
EG4N49835E.guineensis	GGCATAGCTGACATCATCCACGCCCATGGCAAGCCAATCTCCCTCCTTGAAGTACCAACC	177
EG4N49835E.oleifera	GGCATAGCTGACATCATCCACGCCCATGGCAAGCCAATCTCCCTCCTTGAAGTACCAACC	180
	*****	
EG4N49835E.guineensis	AAACTATCCATACCCCCACCAGAAACAACGACCTGCGCCGCCTCATGCGCCTCTGGTG	237
EG4N49835E.oleifera	AAaCTATCCATACCCCCACCAG-AAACAACGACCTGCGCCGCCTCATGCGCCTCTGGTG	239
	*****	
EG4N49835E.guineensis	TTCTCGGATTCCTCGCTCGGAGCGAGATGAAACAGTCACGACCACGACGAAGAGCTG	297
EG4N49835E.oleifera	TTCTCGGATTCCTCGCTCGGAGCGAGATGAAACAGTCACGACCACGACGAAGAGCTG	299
	*****	
EG4N49835E.guineensis	TACATCCCCACCACCATGTCTGCTCTCCTTGTCATCAGAAAGATGTGAACGTCTCGTCC	357
EG4N49835E.oleifera	TACATCCCCACCACCATGTCTGCTCTCCTTGTCATCAGAAAGATGTGAACGTCTCGTCC	359
	*****	
EG4N49835E.guineensis	TTGCTCTCGTAATGCTGGACCAAATGTTCTACACCCATGGTATCACCTCAGTAGTTGG	417
EG4N49835E.oleifera	TTGCTCTCGTAATGCTGGACCAAATGTTCTACACCCATGGTATCACCTCAGTAGTTGG	419
	*****	
EG4N49835E.guineensis	CTCAAGCAAGATATGCTGCCGGCCACTCCCTTTGAGGCGGCACATGGGAAGGGTGTGG	477
EG4N49835E.oleifera	CTCAAGCAAGaTAAGCTGCCGGCCActCCcTTtGAGGCGGCACATGGgAAGGgtGTTtGG	479
	*****	
EG4N49835E.guineensis	GAGGCCACCAGCCAAAGCCCGGCTTCAACAGTTTGTGTAACGAGGGAATGGCAAGCGAC	537
EG4N49835E.oleifera	gAGGCCACCAGCCAAAGCCcGGCCTTCAACAGTTtGTGTAACGgAGGGAATGGCAAGCGAC	539
	*****	
EG4N49835E.guineensis	GCTCGGTTGTTGGCCAAAGTGGTTGTGGAGGAGTGGCGGAGATCTCCGGGCTGAGC	597
EG4N49835E.oleifera	gCTCGGTTGTTGGCCAAaTGGTTGTGGAGGAGTGGCGGAGATCTCCGGGgCTGagC	599
	*****	
EG4N49835E.guineensis	TCGCTGGTGGATGTTGGTGGAGGGACGGGCACGATGGCAGTGCCATCGCCGATGCCTAT	657
EG4N49835E.oleifera	TCGCTGGTGGATGTTGGTGGAGGgACGGgCACGATGGCAGTGCCATCGCCGATGCCTAT	659
	*****	
EG4N49835E.guineensis	CCACACGTGAAGTGCACCGTACTTGACCTCCCTCAGTGGTAGCGAGCTCGGAAGCGGGC	717
EG4N49835E.oleifera	CCACACGTGAAGTGCACCGTACTTGACCTCCCTCAGTGGTAGCGAGCTCGGAAGCGGGC	719
	*****	
EG4N49835E.guineensis	GAGACGGTCCAGCTTGTGCGGGGAGACATGTTTCAGTATGTTCCCTGCTGATGCGGTG	777
EG4N49835E.oleifera	GCGAcGGTCCAGCTTgtGcGGgAGaCATGTTTCaGTATGTTCCCTGCTGATGCGGTG	779
	* *****	
EG4N49835E.guineensis	CTGCTCAAGTGGGTGCTGCATGATTGGAGTGACGAGGAATGCGTGAAGATACTGAAAAGA	837
EG4N49835E.oleifera	CTGCTCAAGTGGgTGctGCATGATTGGAGTGAcGAGGAATGCGTGAAGATACTGAAAaGA	839
	*****	
EG4N49835E.guineensis	TGTAAGGAGGCTGTTCCATCTAAAGAAAAGGGAGGGAAGGCTTTGGAGTGGTGGTTACA	897
EG4N49835E.oleifera	TGTAAGGAGGCTGTCCAtCTAAaGAAAAGGgAGGgAAGGCTTTGGAGTGGTGGTTACA	899
	*****	
EG4N49835E.guineensis	TCACTCGTCTCAGTGCATATTCCTACTGAGAGTTCTGCTCTTGA	942
EG4N49835E.oleifera	TCACTCGTCTCAGTGCATATTCCTACTGAGAGTTCTGCTCTTGA	944
	*****	

6.2. EG4P49884: Deleção no gene *E. oleifera* (1pb em vermelho).

EG4N49884guinennsis	ATGGATTCCTTAGTTGGAGAATGTCAATCTTTAGAGCTCTTGAAAGCTCAAGCTCATTCA	60
EG4N49884oleifera	ATGGATTCCTTAGTTGGAGAATGTCAATCTTTAGAGCTCTTGAAAGCTCAAGCTCATTCA *****	60
EG4N49884guinennsis	TGGAGCCACATATTTCAGCTTCATAAACTCCATGTCCCTCAAGTGTGCCGTAGAGCTCGGC	120
EG4N49884oleifera	TGGAACCACATATTTCAGCTTCATAAACTCCATGTCCCTCAAGTGTGCCGTAGAGCTCGGC *****	120
EG4N49884guinennsis	ATAGCAGACATCATCCACCCATGGCAAGCCAATCTCCCTCCTGAACTAGCAACCAAAA	180
EG4N49884oleifera	ATAGCAGACATCATCCAcACCCATGGCAAGCCAATCTCCCTCCTGAACTAGCAACCAAAA *****	180
EG4N49884guinennsis	CTATCCATcCCCCACCAGAAACACGACTTGGCCGCCTCATGCGCCTCCTGGTGTTC	240
EG4N49884oleifera	CTATCCATCCCC-CCACCAGAAaCAACgACTTGCgCCGCCTCATGCGCCTCCTGGTGTTC *****	239
EG4N49884guinennsis	TCGGGATTATTCGGTTCGAAAGCGTGATGGGAACAGTCCCGATCAGCAGCAGCAAGAGCTG	300
EG4N49884oleifera	TCGGGATTCTTCGCTCGAAaCGTGATGGGAACAGTCCCGATCAGCAGCAGCAAGAGCTG *****	299
EG4N49884guinennsis	TACATCCCCAACCCATGTCTGCCCTCCTTGTCAAGGAGAAGGATGCGAACGCTCTCGTCG	360
EG4N49884oleifera	TACATCCCCAACCCATGTcTGCcTCTTGTCAAGGagAAGGATGCGAACGCTCTCGTCG *****	359
EG4N49884guinennsis	TTCTTCTGATGCTGGACCAAATGTTGCTCAACCCATGGTATAACCTCAGTAGTTGGTTC	420
EG4N49884oleifera	TTCTTCTGATGCTGGACCAAaTgTTGCTCAACCCAGGTATAACCTCAGTAGTTGGTTC *****	419
EG4N49884guinennsis	AAGCAAGATAAGCTGCCACGCCCTTCGAGGTAGCACATGGAAAGGGTGTGGGAGGCC	480
EG4N49884oleifera	AAGCAAGaTAAGTGCcAcGCCcTTCGAGGTAGCACATGGAAaGGgTGTTGGgAGGCC *****	479
EG4N49884guinennsis	ACCAGCCAAACCCGGGCTTCGATAGCTTGGTGAACGAGGGGTGGCAAGCAGCTCTCAG	540
EG4N49884oleifera	ACCAGCCAAaCCCGGgCT-TCGATGGCTTG-GGgAACGAGGGTGGCAAGCAGCTCTCAG *****	537
EG4N49884guinennsis	TTCGTGGCCAAACTGGTTCGTCGAGGAGTGCAGGAGGCTTCAAAGGCCTGAGGTGCTG	600
EG4N49884oleifera	TTCGTGGCCAAACTGGTTCGTCGAGGAGTGCAGGAGGCTTCAAaGGCCTGAGGTGCTG *****	597
EG4N49884guinennsis	GTGGACGTTGGTgGGGGACGGGCACGATGGCCGGGCCATCGCGGAGGCCTATCCACAG	660
EG4N49884oleifera	GTGGACGTTGGTGGGg-GACGGgCACGaTGGCCcGGgCCATCGCCGAGGCCTATCCACAG *****	656
EG4N49884guinennsis	GTGAAGTGCACCGTCTTTCGACCTGCCACCGTGGTAGCGAGCTCGGAAGCGAGCGAGGCC	720
EG4N49884oleifera	GTgAAGTGCACCGTCTcCGACCT-CCCCAcgTGGTAgCGAGCTCGGAAGCGAGCGAGGCC *****	715
EG4N49884guinennsis	GTCACGCTTGTCGCGGGAGATATGTTTCAGTATATCCCTCCTGCTGATGCAGTGTGCTC	780
EG4N49884oleifera	GTCACGCTTGTCGCGGGAGATATGTTtCAGTATATCCCTCCTGCTGATGCAGTGTGCTC *****	775
EG4N49884guinennsis	AAGTGGTGTACATGATTTGGAGCAGCAGGACTGCGTAAACATACTGAAACAATGCAAG	840
EG4N49884oleifera	AAGTGGTGTACATGATTTGGAGCAGCAGGACTGCGTAAAGaTACTGAAACAATGCAAG *****	835
EG4N49884guinennsis	AAGGCTGTCCATCCAAAGAAGAgGGGGAAAGAAATACTTTGAAAGCTCTCCGACGAG	900
EG4N49884oleifera	AAGGCTGTCCATCCAAAGAAGAGGGG--GAAaAaAtACTTTGAAAGCTcTCCGACGAG *****	893
EG4N49884guinennsis	AGTAGTCTTACGGTCTTTGGTGTAGGGTTTCGATATCGGGCTACCGTCCGATCGATTTT	960
EG4N49884oleifera	AGTAGTCTTACGGTCTTtGGTGTAGGGgTTtGATATCGGACTATCgTCGGATCGaTTTT *****	953
EG4N49884guinennsis	AGCCTGCTCATCTACTCCGATGGCGATCGGAGCAGCTCGATGATTGGACTGTACGTCCAG	1020
EG4N49884oleifera	AGCCTGCTCATCTAcTCCGATGGCGATCGGAGCAGCTcGACGATTGGACTgTAGGTCCCTA *****	1013
EG4N49884guinennsis	GCC_TGA 1026	
EG4N49884oleifera	GGCTGA 1020 * **	

6.2.1. A sequência EG4P49859 é mesma EGN49835 para *E. oleifera*.

G4N49859E.guinnessis	ATGAAGGAAAGAAGGGAACCCCTGGAGCCGTCCTTTGGTGTCTATTGCCGCAATTAGCA	60
EG4N49859Eoleifera	-----	0
EG4N49859E.guinnessis	ATCCATAAGAGACTCAGGCGCTGGAAAAGAGTGCAGTGTGATGGAATCCTTGGATGGA	120
EG4N49859Eoleifera	-----ATGGAGTACTCCTTGGATGGA	21
	* * * * *	
EG4N49859E.guinnessis	GAATGTCAATCTTTGGAGCTCCTGAAAGCTCAAGCCCATGCATGGAACCACAAATTTAGC	180
EG4N49859Eoleifera	GAATGTCAATCTTTGGAGCTCCTGAAAGCTCAAGCCCATGCATGGAACCACATATTTAGC	81
	*****	
EG4N49859E.guinnessis	TTCATAAACTCCATGTCCCTCAAGTGTGCCGTGGAGCTCGGCATAGCTGACATCATCCAC	240
EG4N49859Eoleifera	TTCATAAACTCCATGTCCCTCAAGTGTGCCGTGGAGCTCGGCATAGCTGACATCATCCAC	141
	*****	
EG4N49859E.guinnessis	GCCCATGGCAAGCCAACCTCCCTCCTTGAAGTAGCAACCAACTATCCATACCCCCCACC	300
EG4N49859Eoleifera	GCCCATGGCAAGCCAATCtCCCTCCTTGAAGTAGCAACCAaCTATCCATACCCCCcACC	201
	*****	
EG4N49859E.guinnessis	AGAAACAACGACCTGCGCCGCTCATGCGCCTCTGGTGTCTCGGGATTCTTCGCTCGG	360
EG4N49859Eoleifera	aG-AAACACGACCTGCGCCGCTCATGcGCTCCTGGTGTCTCGGgATTCTTCGCTCGG	260
	* * * * *	
EG4N49859E.guinnessis	GAGCGAGATGGAACAGTACAGCACCAGCAGAGAGCTGTACATCCCCACCACCATGTCT	420
EG4N49859Eoleifera	gAgCGAGATGGAaACAGtCACGACCAGCAGAGAGCTGTACATCCCCACCACCATGTCT	320
	*****	
EG4N49859E.guinnessis	GCTCtCCTTGTCAAtcAGAAAGATGTGAACGCTCGTCTTCGCTCTCGCAATGCTGGAC	480
EG4N49859Eoleifera	GCTCtCCTTGTCAATCaGAAAGATGTGAACGCTCGTCTTCGCTCTCGCAATGCTGGAC	380
	*****	
EG4N49859E.guinnessis	CAATGTTCCTACCCCATGGTATCACCTCAGTAGTTGGCTCAAGCAAGaTAATCTGCCG	540
EG4N49859Eoleifera	CAATGTTCCTAcACCCAtGGTATCACCTCAGTAGTTGGCTCAAGCAAGaTAAGCTGCCG	440
	*****	
EG4N49859E.guinnessis	GcCACTCCCCTTGAgGCAGCACATGGGAAGGGTgtTTGGGAGGCCACCAGCCAAAGCCCG	600
EG4N49859Eoleifera	GCCActCCcTtGAGGCGGCACATGgAAGGgtGTTtGGgAGGCCACCAGCCAAAGCCCG	500
	*****	
EG4N49859E.guinnessis	GCCTTCAACAGCTTGCTGAACAGGGGATGGCAAGCGACGCTCGGTTCGTGGCCAAAGTG	660
EG4N49859Eoleifera	GCCTTCAACAGTtTGCTGAACgAGGGATGGCAAGCGACgCTCGGTTCGTGGCCAAAcTG	560
	*****	
EG4N49859E.guinnessis	GTTGTGGAGGAGTGCAGGAGATCTTCCGGGCGCTGAGCTCGCTGGTGGATGTTGGTGGGA	720
EG4N49859Eoleifera	GTTGTGGAGGAGTGCAGGAGATCTTCCGGGgCCTGagCTCGCTGGTGGATGTTGGTGGGA	620
	*****	
EG4N49859E.guinnessis	GGGACGGGCACGATGGCACGTGCCATCGCCGATGACTATCCACACGTGAAGTGCACCGTA	780
EG4N49859Eoleifera	GgGACGGgCACGATGGCACGTGCCATCGCCGATGCCTATCCACACGTGAAGTGCACCGTA	680
	*****	
EG4N49859E.guinnessis	CTTGACCTCCCTCAGGTGGTAGCGAGCTCGGAAGCGGGCGAGACGGTCCAGCTTGTCCG	840
EG4N49859Eoleifera	CTTGACCTCCcTCAGGTGGTAGCGAGCTCGGAAGCGGGCGGAcGGTCCAGCTTGTcCG	740
	*****	
EG4N49859E.guinnessis	GGAGACATGTTTCAGTATGTTCCCTGCTGATGCGGTGCTGCTCAAGGTGATTAAGA	900
EG4N49859Eoleifera	GgAGaCATGTTTCaGTATGTTCCCTGCTGATGCGGTGCTGCTCAAGTGGgTGctGCAT	800
	*****	
EG4N49859E.guinnessis	ACCGATTGCATGCTAAATCAAGAGTGGCATCATACCCCAAGTG-----	947
EG4N49859Eoleifera	GATTGGAGTGAcGAGGAATGCGTGAAGATActGAAAaGATGTAAAGAGGCTGTTCCAtCT	860
	* * * * *	
EG4N49859E.guinnessis	-A-----	948
EG4N49859Eoleifera	AAaGAAAaGGgaGGaAAGGCTTTGGAGTGGTGGTTcACATCACTCGTCTCAGTGCgTATT	920
	*	
EG4N49859E.guinnessis	-----	948
EG4N49859Eoleifera	CCTACTGAGAGGTTCTGCTCTTGA	944

6.2.2. EG4P993: Deleção no gene *E. oleifera* (2pb em vermelho).

EG4N993E.guinnessis	-----ATGGACCAACTCCAAGACAAGCCAGCTGT	31
EG4N993E.oleifera	ATGGGACTCATCCAAGagAAGCAGCCAGCCAAGGATACACTCCAAAaCAAGCCAGCTGT	60
	* * * * *	
EG4N993E.guinnessis	GGAACCACGCTTAGGCTTCATTAGATCCATGTCTTCAAGTGGCCATCGAGCTCGGCA	91
EG4N993E.oleifera	GGAACCacGccCtTAAGCTTCATTAGATCCATGTCTCTCAAGTGGCCATCGAGCTCGGCA	120
	*****	
EG4N993E.guinnessis	TTTCCGAGTTCCTGCACAACCACGGCAAGCCCATCACTCTCTCCGAGCTTGCCACTTCC	151
EG4N993E.oleifera	TCCcGGAcGTTctGCACAACCACGGCAAGCCCTCACTCTCTCCGAGCTTGCCACTTCC	180
	* * *****	
EG4N993E.guinnessis	TCFCCATCCCTCCGTCCAAAACCCGGCCCTCGGAAGCCTCATGCGCTTGCTCGTCCACT	211
EG4N993E.oleifera	tCTCCATCCctCCGTCCAAAaCCcC--GCTCGGAAGCCTCATGCGCTTGCTCGTCCACT	238
	*****	
EG4N993E.guinnessis	CGGGCGTTTCGCGAGTCACCCACAACAAGGAGGGGAGAGAAGAGGGGTATTTCCTCA	271
EG4N993E.oleifera	CGGGCATTTtCGCGAGTCACCCGAcCCAAGTAGCGGgAgAAGAAGAGAGTTATTtCCCTCA	298
	*****	
EG4N993E.guinnessis	CGCCGACCTCTGAACTCCTGGTGAA---GGAGGAGAAGCGGTGTCTCGCCGTTCGCGC	328
EG4N993E.oleifera	CGCCGACCTCTGAACTTctGGTGATCAAGGAGGAGAAGGTGTGTATCTCGCCGTTCGCGC	358
	*****	
EG4N993E.guinnessis	TGCTGATATTGAACTGGACCATGTTGGTGCCGGGACCCAGCTGGGGAGTGGTTCATGG	388
EG4N993E.oleifera	TGATGATaTTGAACTGGACcctGCTGGGCGGcG-CACCACCTGGGgAgTGGTTCAGGG	417
	* * *****	
EG4N993E.guinnessis	CGGGGCGGCGGAGACGCGCTTCGACATGGCGCACGGCAAGGGGATATTGGGGAGA	448
EG4N993E.oleifera	CTG--GGcCGTGGCCGACCGCTTCGACATGGTGACCGCAAGGgATtCTTCGAGGAGA	476
	* * * * *	
EG4N993E.guinnessis	CGAACGTGAGGGCAGACTTCAACGAGCTGTTCAACGAGCGATGGCGAGCGCGCGGC	508
EG4N993E.oleifera	CGAACCGAGGgCAGACTTCAACGAGCTGTTCCACAAGCGATGGCGAGCGACGCTCGGC	536
	*****	
EG4N993E.guinnessis	TGGTGACGGAGTGGCGTTCCTGCACGGCAGTGGGAGATGCTCCGCGGGTGGCGTCGG	568
EG4N993E.oleifera	TGGTGACGGaGTTGGTGTCTGCGGcGGCAGTGGgAgGTTGCTCCGtGGGgTGGCGTCCG	596
	*****	
EG4N993E.guinnessis	TGGTGGACGTGGGCGGGAACGGGCACGGTGGCGAAATCCATCGCCAAGGCGTTGCCCG	628
EG4N993E.oleifera	TGGTGGACGTGGGCGGGGgCACGGGCACGCTGGCGAAGGCCATCGCCAAGGCGTTtCCcG	656
	*****	
EG4N993E.guinnessis	AGGTGCGGTGCATGGTGTGGACCTGCCACACGTGGTGGCGAGATGGACGAGGAGCAGC	688
EG4N993E.oleifera	aGGTGGCGGTGCATGGTGTGGACCTGCCcAcCGTGGTGGACACGGTGCACGAGGGGgagA	716
	*****	
EG4N993E.guinnessis	GGGAGGGCGTGGAGTTCATCGAGGGGACATGTTGAGCAGATCCCTCCGCGGATGCCG	748
EG4N993E.oleifera	AGGaGGCGTGGAGTTCGtCGGAGGGgAcATGTTGAGCAGATCCCTCCGCGGATGCCG	776
	*****	
EG4N993E.guinnessis	TCTTGCTCAAaTGGATATTACATGACTGGAGTGTGAACACTGCATCAAGATACTAAAGC	808
EG4N993E.oleifera	TCTTGCTCAaTGGATATTAcATGATTGGAgTGTGAGCATTCATCAGGATACTAAaGC	836
	*****	
EG4N993E.guinnessis	ATTGCAAGAAGCTATTCTACTAAGAAAGATGGTGGAAAGTACTCGCCACAACCTAGCTA	868
EG4N993E.oleifera	GTTGCAAGAAGCTATTCTcCTTAAGAAaGATGGCGGAAA-----	876
	*****	
EG4N993E.guinnessis	GTATCATGGAGCTCACCAATGGCCAGAGCTTTGGCAAGCCTTTCCAAGCCATAACCTTC	928
EG4N993E.oleifera	-----	876
EG4N993E.guinnessis	TTTTGAACCAATCTTTTCAGCACCCCGGGCTGATGTCCCTCAACTGTGCGGTTGAGCTCG	988
EG4N993E.oleifera	-----	876
EG4N993E.guinnessis	GCATCGCCGACCTCATCCACAGCCACGGCGGCCCTATGCCCTCTCCGAGCTGGCCAGAT	1048
EG4N993E.oleifera	-----	876
EG4N993E.guinnessis	CCATCCCCATCCCTCCCGAGAAGGCTCCCTCTCCCGCCCTCATGCGTCTCTCGTGA	1107
EG4N993E.oleifera	-----	876

6.2.3. EG4P49850: Deleção no gene *E. oleifera* (2pb em vermelho).

EG4N49850Eguinennesis	ATGGAATCCTTGGACGGAGAATGCCAATCTTTGGAGCTCCTGAAAGCTCAAGCCCATGCA	60
EG4N49850Eoleifera	ATGGAATCCTTGGACGGAGAATGCCAATCTTTGGAGCTCCTGAAAGCTCAAGCCCATgCA	60
	*****	
EG4N49850Eguinennesis	TGGAACCaCATATTTAGCTTCATAAAGTCCATGTcCCTcAAGTGTGCCGTGGAGCTCGGC	120
EG4N49850Eoleifera	TGGAACCaCATATTTAGCTTCATAAAGTCCAcGTCCcTCAAGTGTGCCgTGGAGCTCGGC	120
	*****	
EG4N49850Eguinennesis	ATAGCAGACATCATCCATACCCATGACAAGCCAATCTCCCTCACTGAAGTCCGCTAAA	180
EG4N49850Eoleifera	ATAGCAGaCATCATCCATACCCATGACAAGCCATcTCCCTCACTGAAGTCCGCTAAA	180
	*****	
EG4N49850Eguinennesis	CTATCCATTCCCCCACCaaaAACACGACCTGCGCCGCCTCATGCGCTCTTGGTGTTC	240
EG4N49850Eoleifera	CtATCCATTCCCCaC--CAAAACaCaGACCTGCGCCGCTCATGCGCTTtGGTGTTC	238
	*****	
EG4N49850Eguinennesis	TCGGGATTCCTCGCTCGGAAGCGTGATGGAAACAGTCACGACCACGACGAAGAGTTGTAC	300
EG4N49850Eoleifera	tcGGGATTCCTCGCTCGGAAGCGTGATGGAAaCaGTCACGACCACGACGAAGaGCTGTAC	298
	*****	
EG4N49850Eguinennesis	ATCCCCACCACCATGTCTGCTCTCCTTGTCAAGGAGAAGATGCGAAGCTCTCCTCCTTC	360
EG4N49850Eoleifera	ATCCcCaCCACCATGTCTGCTCTCCTTGTCAAGGaAAGATGCGAAGCTCTCCTCCTTC	358
	*****	
EG4N49850Eguinennesis	GTTCTCGCAATTCTCGACCAACGTTGTTCAACCCATGGTATCACCTCAGTAGTTGGTTC	420
EG4N49850Eoleifera	GTTCTCGCAGCGCTGGaCCAAATGTTGcACCCATGGTATCACCTCAGtAgTTGGTTC	418
	*****	
EG4N49850Eguinennesis	AAGCAAGATAAGGTACCCGCCACTCCCTTTGAGTAGCACATGGAAGGGATGTTTGGGAG	480
EG4N49850Eoleifera	AAGCAAGaTAAGGTGCCCGCTActCCCTTTGAGTAGCACATGGAAGGGtGTTtGGGAG	478
	*****	
EG4N49850Eguinennesis	GCCATCAGCCGAAGCCCGCCTTCAACAGCTTGTGCTACGAGGCGATGGCAAGCGACTCT	540
EG4N49850Eoleifera	GCCTTCAACCGAAGCCcGGCCtTCAACaGCTTGTTCACGAGGCGATGGCAAGCGACTct	538
	*** **	
EG4N49850Eguinennesis	CGGTTGCGTGGCCAAACtGGTTGTCGAGGAGTGGCGGAGGCTCTCCGAGGCTGAGGTCTG	600
EG4N49850Eoleifera	CGGTTGCGTGGCCAAaCtGGTTGTCGAGGAGTGGCGGAGGCTCTCCGAGGgCCTGAGGTCTG	598
	*****	
EG4N49850Eguinennesis	CTGGTGGATGTCGGTGGGGAAAAGGCACGATGGCCCGGCCATCGCGGTGGCCTATCCA	660
EG4N49850Eoleifera	CTGGTGGATGTCGGTGGGGGA--AaGGCACGATGGCCCGGCCATCgCGGAGGCCTATCCA	657
	*****	
EG4N49850Eguinennesis	CAGGTGAAGTGCACCGTCCTTGACCTGCCTCAGTGGTAGCAAGCCGGAAGCGAGCGAG	720
EG4N49850Eoleifera	CAGGTGAAGTGCACCGtCCTTGACCTGCCCcAcGTGGTAGCAAGCTCGGAAGCGAGCgAG	717
	*****	
EG4N49850Eguinennesis	GAGGTCCAGCTTGTGCGGGAGACATGTTTCAGTATGTCCCTCCTGCGGATGCAGTGTCTG	780
EG4N49850Eoleifera	GCGGTCCAGCTAGTCGCGGgAgAcATGTTtCAGTATGTCCcTCCTGCGGATGCAGTGTCTG	777
	* *****	
EG4N49850Eguinennesis	CTCAAGAACCAGATTGGCTCCACCCTATCCACCAGCCATGCCACCATCACCATCACCTCT	840
EG4N49850Eoleifera	CTCAAGAACCAGATTGGCTCCACCCTTgCCATCGGCCAtGCCACCATCACCATCACCTCT	837
	*****	
EG4N49850Eguinennesis	TTCAGCGCTCACTATTACATCTATGGCATCAGCGGCCTCTGCTCCGTCGGCTCGAAGCTG	900
EG4N49850Eoleifera	T-----	838
	*	
EG4N49850Eguinennesis	GCCATCACTGTGCTTCAAGAAAGGCATCGTCGTCGTTGGTACAGCGAGGTAG	954
EG4N49850Eoleifera	-----	838

## 7. Alinhamentos das sequências genômicas do gene *EGS* de *E. guineensis* e *E. oleifera*.

### 7.1 EG4N155924: Deleção no gene *E. oleifera* (1pb em vermelho).

EG4N155924	ATGGCGGGAAAGATGAGCAAGATCCTGATCATCGGCGGGACGGGGTACATCGGTAAGTTT	60
EG4N155924.E.oleifera	ATGGCGGGA <b>AA-AT</b> GAGCAAGATCCTGATCATCGGCGGGACGGGGTACATCGGTAAGTTT *****	59
EG4N155924	GTTGTGGGAGCTAGCGCCCGGTCCGGTCATCCACCTTCGCCCTGGTGAGAGATACCACG	120
EG4N155924.E.oleifera	GTTGTGGGAGCTAGCGCCCGGTCCGGTCATCCACCTTCGCCCTGGTGAGAGATACCACG *****	119
EG4N155924	CCCTCCGACCCAGCCAAGGCCAGCTCCTGGAGACTTTCAGAAGCTCCGGCGTTACCCCTC	180
EG4N155924.E.oleifera	CCCTCCGACCCAGCCAAGGCCAGCTCCTGGAGACTTTCAGAAGCTCCGGCGTTACCCCTC *****	179
EG4N155924	CTCAAAGGGGATTTAAATGACCACGGGAGCTTGGTAGAGGCCATTAACAAGTGGACGTT	240
EG4N155924.E.oleifera	CTCATAGGGGATTTAAATGACCACGGGAGCTTGGTAGAGGCCATTAACAAGTGGACGTT **** *****	239
EG4N155924	GTAATCTCAACGGTCGGCTTACAGCAGGTGATGGACCAGCTTAAGATCATTCGCTGCTATT	300
EG4N155924.E.oleifera	GTAATCTCAACGGTCGGCTTACAGCAGGTGATGGACCAGCTTAAGATCATTCGCTGCTATT *****	299
EG4N155924	AAAGAAGTCGGAACGATTAAGAGGTTCTTGCCCTCCGAGTTTGGTTTAGATGTTGACCGC	360
EG4N155924.E.oleifera	AAAGAAGTCGGAACGATTAAGAGGTTCTTGCCCTCCGAGTTTGGTTTAGATGTTGACCGC ***** **	359
EG4N155924	ATACACAAGCAGAGCCTGCATCGTCGGTGTTCGCACTCAAGGCTCAGATTCGGCGTGCC	420
EG4N155924.E.oleifera	ATGCATAAAGCCGAGCCTGCATCGTCGGTGTTCGCACTCAAGGCTCAGATTCGGCGTGCC ** * *****	419
EG4N155924	ATTGAGGCAGAAGGGATCGCTCACACCATTGTGTGCTGCAACTGCTTCGCTGGATACTTT	480
EG4N155924.E.oleifera	ATTGAGGCAGAAGGGATCGCTCACACCATTGTGTGCTGCAACTGCTTCGCTGGATACTTT ***** ****	479
EG4N155924	CTTCCAACACTAGCACAGCCAGAACCTCCAAGAACAAGTTACAATATTAGGCGACGGA	540
EG4N155924.E.oleifera	CTTCCAACACTAGCACAGCCAGAACCTCCAAGAACAAGTTACAATATTAGGCGACGGA *****	539
EG4N155924	AATCAGAAAGTTGTCTCCTTGCTGAAGAAGATGTCGGCACGTATACCATCAAACAGTG	600
EG4N155924.E.oleifera	AATCAGAAAGTTGTCTCCTTGCTGAAGAAGATGTCGGCACGTATACCATCAAACAGTG ***** * * ** *****	599
EG4N155924	GATGACCCAAGGACTTTGAACAAGATTTTGTACGTGAAACCTCCTGCCAACATCTGTCTCC	660
EG4N155924.E.oleifera	GATGACCCAAGGACTTTGAACAAGATTTTGTACGTGAAACCTCCTGCCAACATCTGTCTCC *****	659
EG4N155924	TTCACCAGCTTGTCTCCCTCTGGGAGAAGAAGACCGGCAAGACATTGGAGAAGATCTAC	720
EG4N155924.E.oleifera	TTCACCAGCTTGTCTCCCTCTGGGAGAAGAAGACCGGCAAGACATTGGAGAAGATCTAC *****	719
EG4N155924	CTTCCGGAAGAAGAGGCTTTGAAGAAGATTCAGGAGGCTCCATTCATCAATATGCTG	780
EG4N155924.E.oleifera	CTTCCGGAAGAAGAGGCTTTGAAGAAGATTCAGGAGGCTCCATTCATCAATATGCTG CTTCCGGAAGAAGAGGCTTTGAAGAAGATTCAGG----- *****	753
EG4N155924	ATGGCAGTCAACTACTTGGTGTGTTGTAAGGAGACACTGCTAACATCCAAATCGAACCA	840
EG4N155924.E.oleifera	----- -----	753
EG4N155924	TCATTTGGTGTGGAGGCTTCTAAGCTCTATCCTGATGTCAAATACACCAGTGTCAAGAG	900
EG4N155924.E.oleifera	----- -----	753
EG4N155924	TATCTCAACCGTGTGCGCTATGGGAATGGCATGTAA	936
EG4N155924.E.oleifera	-----	753



7.2. EG4N76667: Deleção no gene *E. oleifera* (1pb em vermelho).

EG4N76667Eguinennesis	ATGGCGCGGGAGATGAGCAAAATCCTGATCATCGGTGGGACGGGGTACCTCGGTAAGTTT	60
EG4N76667Eoleifera	ATGGCGCGGGAGATGAGCAAAATCCTGATCATCGGTGGGACGGGGTACCTCGGTAAGTTT *****	60
EG4N76667Eguinennesis	ATTGTGGGAGCTAGCGCGCGGTCCGGTCATCCCACCTTCGCCATGGTGAGAGATACCGCC	120
EG4N76667Eoleifera	ATTGTGGGAGCTAGCGCGCGGTCCGGTCATCCCACCTTCGCCATGGTGAGAGATACCGCC *****	120
EG4N76667Eguinennesis	CCCTCCGACCCAGCGAAGGCCAGCTCTGGAGGCTTCAAGAACTCCGGCGTTACCCCTT	180
EG4N76667Eoleifera	CCCTCCGACCCAGCGAAGGCCAGCTCTGGAGGCTTCAAGAACTCCGGCGTTACCCCTC *****	180
EG4N76667Eguinennesis	TTCAAAGGGGATTTAAATGACCACGGGAGCTTGGTGGAGGCCATTAACAAGTGGACGTG	240
EG4N76667Eoleifera	CTCAAAGGGGATTTAAATGACCACGGGAGCTTGGTGGAGTCCATTAACAAGTGGACGTG *****	240
EG4N76667Eguinennesis	GTAATCTCAACGGTCGGCTACAGCAGGTGATGGACCAGCTTAAGATCATAGCTGCTATT	300
EG4N76667Eoleifera	GTAATCTCAACGGTCGGCTACAGCAGGTGATGGACCAGCTTAAGATCATAGCTGCTATT *****	300
EG4N76667Eguinennesis	AAAGAAGTCGGAACGATTAAGAGGTTCTTGCCCTCCGAGTTTGGTTAGATGTTGACCGC	360
EG4N76667Eoleifera	AAAGAATCGGAACGATTAAGAGGTTCTTGCCCTCCGAGATGGTTTGGATGTTGACCGC *****	360
EG4N76667Eguinennesis	ATACACAAGCAGAGCCTGCATCGTCGGTGTTCGCACTCAAGGCTCAGATTCGGCGTGCC	420
EG4N76667Eoleifera	CTACACAAGCAGAGCCTGCATCGTCGGTGTTCGCACTCAAGGCTCAGATTCGGCGTGCC *****	420
EG4N76667Eguinennesis	ATTGAGGCAGAAGGGATCCCTCACACCATTTGTGTGCTGCAACTGCTTCGCTGGATACTTT	480
EG4N76667Eoleifera	ATTGAGGCAGAAGGGATCCCTCACACCATTTGTGTGCTGCAACTGCTTCGCTGGATACTTT *****	480
EG4N76667Eguinennesis	CTTCCAACACTAGCACAGCCAGAACCTCCAAAAGCAAAGTTACAATATTAGGCGACGGA	540
EG4N76667Eoleifera	CTTCCAACACTAGCACAGCCAGAACCTCCAAAANNNNNNNNNNNNNNNNNNNNNNNNNN *****	540
EG4N76667Eguinennesis	AATCAGAAAGTTGTCTTCCCTTGCTGAAGAAGATGTCGGCACGTATACCATCAAACGGTG	600
EG4N76667Eoleifera	NNNNNNNNGTAGTGTCTTCCCTTGCTGAAGAAAATGTCGGCACGTATACCATCAAACGGTG ** * *****	600
EG4N76667Eguinennesis	GATGATCCAAGGACTTTGAACAAGACTTTGTACATGAAACCTCCTGCCAACATCTACTCC	660
EG4N76667Eoleifera	GATGATCCAAGGACTTTGAACAAGACTTTGTACATGAGACCTCCTGCCAACATCTACTCC **** *****	660
EG4N76667Eguinennesis	TTTAACGACCTTGTCTCCCTCTGGGAGAAGAAGACCGGCAAGACATTGGAGAAGATCTAC	720
EG4N76667Eoleifera	TTTAACGACCTTGTCTCCCTCTGGGAAAAGAAGACCTGCAAGATATTGGAGAAGATCTAC *****	720
EG4N76667Eguinennesis	CTTCCGGAAGAAGAGGCTTGAAGAAGATCGAGGAGGCTCCATTCCCCTTCAATATGCTG	780
EG4N76667Eoleifera	CTTCCGGAAGAAGAGGCTTGAAGAAGATCGAGGAGGCTCCATTCCCATTCAATATGCTG *****	780
EG4N76667Eguinennesis	ATGGCAGTCAACTACTTGGTGTGGTGGAGGAGACTGCTAACATCCAAATCGAACC	839
EG4N76667Eoleifera	ATGGCAGTCAACTACTTGGTGTGGTGGAGGAGACTGCTAACATCCAAATCGAACC *****	840
EG4N76667Eguinennesis	ATCATTGGTGTGGAGGCTTCTGAGCTCTATCCTGATGTCAAATATACTAGTGTCAAGA	899
EG4N76667Eoleifera	ATCATTGGTGTGGAGGCTTCTGAGCTCTATCCTGATGTCAAATACACCAGTGTCAAGA *****	900
EG4N76667Eguinennesis	GTATCTCAACCGTGTATCCAATGA	924
EG4N76667Eoleifera	GTATCTCAACCGTGTATCCAATGA	925

7.3. EG4N76665: Deleção no gene *E. oleifera* (1pb em vermelho).

EG4N76665Eguinennesis	ATGGCGGGAAAGATGAGCAAGATCCTGATCATCGGCGGGACGGGGTACCTCGGTAAGTTT	60
EG4N76665E.oleifera	ATGGCGGGAAA-ATGAGCAAGATCCTGATCATCGGCGGGACGGGGTACATCGGTAAGTTT	59
	*****	
EG4N76665Eguinennesis	ATTGTGGGAGCTAGCGCGGGTCTGGTCATCCCACCTTCGTCTGGTGGAGAGATACCGCC	120
EG4N76665E.oleifera	GTTGTGGGAGCTAGCGCCCGTCCGGTCATCCCACCTTCGCCCTGGTGGAGAGATACCA	119
	*****	
EG4N76665Eguinennesis	CCCTCCGACCAAGCCAAGGCCAGCTCTTGAGGGCATTCAAGAACTCCAGCGTCACCTC	180
EG4N76665E.oleifera	CCCTCCGACCCAGCCAAGGCCAGCTCCTGGAGACTTCAAGAACTCCGGCGTTACCTC	179
	*****	
EG4N76665Eguinennesis	CTCAAAGGGGATTTAAATGACCACGGGAGCTTGGTGGAGGCCATGAAACAAGTGGACGTG	240
EG4N76665E.oleifera	CTCAAAGGGGATTTAAATGACCACGGGAGCTTGGTGGAGGCCATTAACAAGTGGACGT	239
	*****	
EG4N76665Eguinennesis	GTAATCTCAACGATCGGGCCACGGCAGGTGTGGACCAGTTTAAAGATCATCGCTGCTATT	300
EG4N76665E.oleifera	GTAATCTCAACGATCGGCTTACAGCAGGTGATGGACCAGCTTAAGATCATTCGCTGCTATT	299
	*****	
EG4N76665Eguinennesis	AAAGAAGTCGGAACGATTAAGAGAGAGAAAAATAAAAAGAGAAAAATAATGTGAAGAGAGAG	360
EG4N76665E.oleifera	AAAGAAGTCGGAACGATTAAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	359
	*****	
EG4N76665Eguinennesis	AAAAGTCATCAGGGCGGCTGAAGGGCAAGACCACCGAAGTCTTCGCCTTAGGcCCCCAc	420
EG4N76665E.oleifera	NN	419
EG4N76665Eguinennesis	CCCCAAGAGGcCCCCCGCTCGTGCGCATCCGGCATCGCGGTCCAGCCTCCGCCTCCCGCC	480
EG4N76665E.oleifera	NNNNNNNN--NNN--NN	473
EG4N76665Eguinennesis	TCTGGCGTCCCGGTCCAAGCGTCCAACCTCCGCCTCCGGCGTCCCGATCCAcCCCCGGCC	540
EG4N76665E.oleifera	NN	533
EG4N76665Eguinennesis	TCCACCGATGTTGATCGCATCCACACAGCGGAGCCTGCATCGTCGGTGTTCGCTCTCAAG	600
EG4N76665E.oleifera	NNNNNNNGATGTTGATCGCATAAAGCCGAGCCTGCATCGTCGGTGTTCGCTCTCAAG	593
	*****	
EG4N76665Eguinennesis	GCTCAAATTCGTGGGCCATTGAGGCAGAAGGGATCCCTTGCAACCATTTGTGTGCTGCAAC	660
EG4N76665E.oleifera	GCTCAGATTCGGCGTGCCATTGAGGCAGAAGGGATCCCTCACACCATTGTGTGCTGCAAC	653
	*****	
EG4N76665Eguinennesis	TGCTTCGCGGATATTTTCTTCCAACGCTAGCACAGCCAGAACCTCCAAAAGaCAAAGTT	720
EG4N76665E.oleifera	TGCTTCGCTGGATTCTTCTTCCAACACTAGCACAGCCAGAACCTCCAAAAGaCAAAGTT	713
	*****	
EG4N76665Eguinennesis	ACAGTATTAGCGATGGAATCCGAAGGCTGTCTTCTTGCCGAAGAAGATATCGGCACG	780
EG4N76665E.oleifera	ACAATATTAGCGACGGAATCAGAAAGGTGCTCTCTGCTGAAGAAGATGTCGGCACG	773
	*****	
EG4N76665Eguinennesis	TATACCATCAAGACTGTAGATGACCCGAGGACTTTGAACAAGATTTGTACGTGAAACCT	840
EG4N76665E.oleifera	TATACCATCAAAACGGTGGATGACCCAAGGACTTTGAACAAGATTTGTACGTGAAACCT	833
	*****	
EG4N76665Eguinennesis	CCTGCTAACATCTGCTCCTTTAATGATCTGTCTCCCTCTGGGAAAAGAAGACCGGCAAG	900
EG4N76665E.oleifera	CCTGCCAACATCTGCTCCTTCAACCAGCTTGTCTCCCTCTGGGAGAAGAAGACCGGCAAG	893
	*****	
EG4N76665Eguinennesis	ACGTTGGAGAAGATTTACCTTCCGAAGAAGAGGTCTTGAAGAGGATCCAGGAGGCTCCA	960
EG4N76665E.oleifera	ACATGGAGAAGATCTACCTTCCGAAGAAGAGGTCTTGAAGAAGATTCAGNNNNNNNN	953
	*****	
EG4N76665Eguinennesis	TTCCCATGAACATTCTGATGGCAGTCAACTACTTGGTGTGTTGTAAGGGAGACTGCT	1020
EG4N76665E.oleifera	NN	1013
EG4N76665Eguinennesis	AATATTGAGATCGATCCATCATTTGGTGTGGAGGCTTCTGAGCTCTATCCTGACGACCTT	1080
EG4N76665E.oleifera	NN--NNNN	1070
EG4N76665Eguinennesis	CCAGCCTCACACGAGAGAAAGGCCCTCCACTCCTTGTGTGCTAGACAACCTCTTTGCTATT	1140
EG4N76665E.oleifera	NNNNNNNN--NNNNNNNN--NNNNNNNN--NNNNNNNNNNNNNN--NNNNNNNNNNNN	1126
EG4N76665Eguinennesis	CCCTCTGTTATGGA AAAAATTGGA GCGGCTTTAGAGTAGAGGCTGCTGAGCTTTATCCA	1200
EG4N76665E.oleifera	N--NNNNNNNNNN--NNNNNNNNNN--NNNNNNNNNNNNNNNNNNNNNNNNNNNN--	1179
EG4N76665Eguinennesis	GATGTCAAACACACCATCTTGGATGAGTACCTTAATCGTGTATCCAATGA	1251
EG4N76665E.oleifera	-----	1179

