

Traditional Medicinal Plants into New Traditions:  
A Botanical Study of Tulsi (*Ocimum tenuiflorum*)

Zur Erlangung des akademischen Grades einer

DOKTORIN DER NATURWISSENSCHAFTEN  
(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften  
des Karlsruher Instituts für Technologie (KIT)  
genehmigte

DISSERTATION

von

Daniela Ríos Rodríguez

aus

Chile

Referent: Prof. Dr. Peter Nick  
Korreferent: Prof. Dr. Tilman Lamparter  
Tag der mündlichen Prüfung: 20. Oktober 2021

Die vorliegende Dissertation wurde am Botanischen Institut des Karlsruher Instituts für Technologie (KIT), Lehrstuhl I für molekulare Zellbiologie, im Zeitraum von Oktober 2017 bis September 2021 angefertigt.

This dissertation was printed with the support of the German Academic Exchange Service.



## I. Statement

Hiermit erkläre ich, dass ich die vorliegende Dissertation, abgesehen von der Benutzung der angegebenen Hilfsmittel, selbstständig verfasst habe.

Alle Stellen, die gemäß Wortlaut oder Inhalt aus anderen Arbeiten entnommen sind, wurden durch Angabe der Quelle als Entlehnungen kenntlich gemacht.

Diese Dissertation liegt in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vor.

Zudem erkläre ich, dass ich mich beim Anfertigen dieser Arbeit an die Regeln zur Sicherung guter wissenschaftlicher Praxis des KIT gehalten habe, einschließlich der Abgabe und Archivierung der Primärdaten, und dass die elektronische Version mit der schriftlichen übereinstimmt.

Relevante Teile dieser Dissertation wurden veröffentlicht in:

- Jürges, G., Sahi, V., Ríos Rodríguez, D., Reich, E., Bhamra, S., Howard, C., Slater, A. and Nick, P. (2018) Product authenticity versus globalisation—The Tulsi case. PLoS One, 13, 1–22.
- Ríos-Rodríguez, D., Sahi, V.P. and Nick, P. (2021) Authentication of holy basil using markers relating to a toxicology - relevant compound. European Food Research and Technology. Available at: <https://doi.org/10.1007/s00217-021-03812-z>.

Karlsruhe, den 8. September 2021

Daniela Ríos Rodríguez



## II. Acknowledgements

Words cannot accurately describe gratefulness, but I will try.

First and foremost, I would like to express my gratitude to my supervisor Prof. Dr. Peter Nick for accepting me as a student. His passion about science, positive and open minded attitude, and his endless curiosity was a source of inspiration during my research years. A true mentor to look up.

Thanks to Prof. Dr. Tilman Lamparter for being the second examiner of my thesis and his expertise.

I would like to thank Dr. Vaidurya Pratap Sahi for his support and guidance during the thesis. Guidance that often went beyond scientific matters, with an extraordinary understanding capability, and from there, finding a word of motivation. Thank you Vaidurya.

Thanks to Dr. Michael Riemann for his insights on plant defence, qPCR, and most of all his kind predisposition to help with different difficulties.

Many thanks to Dr. Adnan Kanbar and Dr. Gabriele Jürges for their support and help in several topics on the science field.

At the Botanical Institute, the Nick Lab group members who were part of my everyday life for about four years, thank you. Thanks for the work, technical and emotional support, wisdom and laughs too. Particularly, the ones from office 517: Sascha, Karwan, Xin and Pallavi. But also others members of the institute, including the Lamparter Lab group, namely Afaf, Nasim, Eman, Toranj, Islam, Chris, Wenjing, Kunxi, Gero, Manuel, Nathalie, Sarheed, Lena B., Lena S., Isabel, and Victoria. I would also like to thank Sarah Gliemann and Simon Friedmann for their contribution to my work.

For technical support, I would like to thank Ernst Heene, who not constructed the UV-B box and supported me on Gas Chromatography; Sabine Purpur, who takes care of the cell cultures, and the Azubis and Nadja Wunsch, for their help on the lab maintenance. For administrative support, I would like to thank Renate Herberger-Biester and Jochen Krüger. And for growing and taking care of all the plant material used in this thesis, particular thanks to the gardeners at KIT Botanical Garden: Dutsareeya Kusterer, Anna-Louise Kuppinger, and Joachim Daumann.

On a personal note, I would like to thank my parents, for their unconditional love and support, and all my family and friends that light up my life.

Particularly, a very special thanks to my dearest husband *Dietrich Wins*, for his love, friendship, patience, support, motivation, enthusiasm and for always believing in me.

Last but not least, thanks to the German Academic Exchange Service (DAAD) for the doctoral scholarship granted to me. The DAAD support was more than financial, it allowed me to do my PhD in Germany, learn German and meet great scholars (and friends), from different fields and corners of the world. Therefore, thanks to DAAD and all their members that made this experience possible.

Daniela Ríos Rodríguez



### III. Abstract

Traditional medicinal plants (TMP) have been long used by humans for health purposes. Nowadays, several TMP are advertised as ‘superfoods’. Yet, it is questionable whether TMP should be straightforwardly considered as food products because some TMP have toxicologically relevant compounds, such as Tulsi or Holy Basil (*Ocimum tenuiflorum*). This TMP also produces the genotoxic phenylpropene methyleugenol (ME), synthesised by the enzyme eugenol *O*-methyltransferase (EOMT). However, it is currently advertised as a ‘superfood’. Moreover, Tulsi is susceptible to food fraud. Therefore, Tulsi was selected as the case study for this research. The aim is on one hand, to better understand the mechanisms under which the synthesis of ME is regulated, and on the other hand, to address the issue of food fraud through developing a species identification technique with DNA barcoding.

To achieve the first aim, the genetic variation of the EOMT enzyme, the phenylpropanoid pathway, and the subcellular localization of EOMT were studied. The results disclosed differences on the nucleotide sequence of the EOMT within the genus *Ocimum*. EOMT-based phylogeny revealed two distinctive groups, the first one called the Tulsi clade (*O. tenuiflorum* and *O. gratissimum*), and the second one called the Basil clade (*O. basilicum*, *O. americanum*, *O. africanum* and *O. kilimandscharicum*). However, similarities were found between the sequences of the Basil clade and only *O. gratissimum* sequences from the Tulsi clade. Furthermore, *O. gratissimum* was the only accession from the Tulsi clade that did not produce ME in the given growing conditions. Also none of the Basil clade accessions produced ME. Thus, it is hypothesised that the amino acid similarities between *O. gratissimum* and the Basil clade sequences, are relevant for the synthesis of ME in the genus *Ocimum*. These results suggest the strong genetic component playing a role on the ME synthesis.

Then, the phenylpropanoid pathway in *O. tenuiflorum* was researched by using high doses of UV-B radiation as stress factor. Two *O. tenuiflorum* chemotypes were used, Krishna and Rama. The gene expression of several enzymes showed different regulation patterns on the phenylpropanoid pathway for the synthesis of ME when comparing the two chemotypes. The upstream reactions to ME in Rama seemed to be highly regulated by the enzyme caffeic acid *O*-methyltransferase. Whereas in Krishna, the enzyme caffeoyl-CoA *O*-methyltransferase seemed to play an important role as well. In addition, it seems that besides phenylalanine, tyrosine might also be a relevant first substrate of the phenylpropanoid pathway in *Ocimum* sp.

In relation to the subcellular localization studies, EOMT was localized in the nucleus and the cytosol of transformed BY-2 cell. The cytosolic localization of EOMT is in line with the localization of other phenylpropanoid pathway enzymes such as PAL and eugenol synthase. However, the nuclear localization of EOMT was unprecedented and further research is needed to understand this outcome.

The second aim was to address the issue of food fraud via DNA barcoding. DNA barcoding for *O. tenuiflorum* discrimination in commercial samples was achieved through a one-reaction assay using a trait-independent marker (*psbA-trnH* *igs*) together with a trait-related marker (EOMT<sup>®</sup>). The latter was based on the enzyme EOMT. This assay allowed for discrimination of *O. tenuiflorum* from other *Ocimum* sp. in reference plants and commercial samples.

Overall, the Tulsi case study highlights the complexity of considering a particular TMP as a food or ‘superfood’ because of the possible presence of toxicologically relevant compounds. Further, this study exposed the need of research regarding the regulation of metabolites synthesis in plants and DNA-Barcoding techniques. Therefore, it is suggested to carry out more precautionary research prior considering a TMP as a food or even so-called ‘superfood’.

## Zusammenfassung

Traditionelle medizinische Pflanzen (TMP) wurden schon lange von Menschen für Gesundheitsanwendungen eingesetzt. Heutzutage werden einige TMP als „Superfood“ beworben. Allerdings ist es fraglich, ob TMP einfach als Lebensmittel angesehen werden können, weil manche TMP giftige Verbindungen beinhalten, z.B. Tulsi bzw. Heiliges Basilikum (*Ocimum tenuiflorum*). Diese TMP produziert das genotoxische Phenylpropanoid Methyleugenol (ME), synthetisiert aus dem Enzym Eugenol *O*-Methyltransferase (EOMT). Trotzdem wird es aktuell als „Superfood“ beworben. Zusätzlich besteht bei Tulsi die Gefahr Lebensmittelbetrug. Deswegen wurde Tulsi als Fallstudie ausgewählt. Die Ziele dieser Forschungsarbeit sind einerseits den Mechanismus besser zu verstehen, der die Biosynthese von ME reguliert und andererseits für das Problem des Lebensmittelbetrugs eine Technik zur Identifikation der richtigen Tulsi Spezies mittels DNA-Barcoding zu entwickeln.

Zur Erreichung des ersten Ziels wurden die genetische Variation des EOMT Enzyms, der Phenylpropanoid-Weg und die subzelluläre Lokalisierung von EOMT untersucht. Die Ergebnisse zeigten Sequenzunterschiede der Nukleotidsequenz des EOMT Enzyms verschiedener *Ocimum* Arten innerhalb der Gattung. EOMT-basierte Phylogenie erlaubte die Unterscheidung von zwei Gruppen, die erste benannt als Tulsi-Klade (*O. tenuiflorum* und *O. gratissimum*), die zweite benannt als Basilikum-Klade (*O. basilicum*, *O. americanum*, *O. africanum* und *O. kilimandscharicum*). Jedoch wurden Ähnlichkeiten gefunden zwischen den Sequenzen der Basilikum-Klade und ausschließlich *O. gratissimum* aus der Tulsi-Klade. Zusätzlich war *O. gratissimum* die einzige Art der Tulsi-Klade, die unter den verwendeten Wachstumsbedingungen kein ME produzierte, sowie ME auch von keiner der Arten der Basilikum-Klade produziert wurde. Deswegen wird die Hypothese aufgestellt, dass Ähnlichkeiten der Aminosäuresequenz in der Basilikum-Klade und *O. gratissimum* aus der Tulsi-Klade relevant sind für die Synthese von ME in der Gattung *Ocimum*. Diese Ergebnisse deuten auf eine starke genetische Komponente, welche bei der Synthese von ME eine Rolle spielt.

Dann wurde der Phenylpropanoid-Weg bei *O. tenuiflorum* durch hohe Dosierung von UV-B Strahlung als Stressfaktor untersucht. Zwei Chemotypen von *O. tenuiflorum* wurden verwendet, Krishna und Rama. Die genetische Ausprägung mehrerer Enzyme zeigten unterschiedliche Regulierungsmuster des Phenylpropanoid-Wegs zur Synthese von ME im Vergleich zwischen den zwei Chemotypen. Die vorgelagerten Reaktionen zu ME in Rama schienen stark von dem Enzym Kaffeesäure *O*-Methyltransferase reguliert zu sein. Wohingegen bei Krishna, das Enzym Caffeoyl-CoA *O*-Methyltransferase auch eine wichtige Rolle zu spielen schien. Zusätzlich scheint es, dass neben Phenylalanin auch Tyrosin ein relevanter erster Ausgangsstoff des Phenylpropanoid-Wegs sein kann in *Ocimum* Arten.

Die subzelluläre Lokalisierung von EOMT wurde im Nukleus und dem Cytosol von transformierten BY-2 Zellen gezeigt. Die cytosolische Lokalisierung von EOMT entspricht der Lokalisierung anderer Komponenten des Phenylpropanoid-Wegs, mit Enzymen wie PAL und Eugenol-Synthase. Jedoch wurde die nukleäre Lokalisierung von EOMT in dieser Arbeit erstmalig gezeigt, sodass weitere Forschung dazu notwendig ist, um dieses Ergebnis zu verstehen.

Das zweite Ziel war das Problem des Lebensmittelbetrugs durch DNA-Barcoding zu adressieren. DNA-Barcoding zur Identifikation von *O. tenuiflorum* in kommerziellen Proben wurde erreicht durch eine einstufige Technik mit einem merkmalsunabhängigen Marker (*psbA-trnH* igs) und mit einem merkmalsabhängigen Marker erreicht (EOMT<sup>+</sup>). Der letztgenannte basierte auf dem Enzym EOMT. Diese Untersuchungsmethode erlaubt die Unterscheidung von *O. tenuiflorum* von anderen *Ocimum* Arten in Referenzpflanzen und kommerziellen Proben.

Insgesamt zeigte die Tulsi Fallstudie die Komplexität dabei eine spezifische TMP als Lebensmittel oder „Superfood“ zu erwägen. Darüber hinaus stellte diese Fallstudie den Forschungsbedarf heraus bei der Regulierung der Stoffwechselproduktsynthese in Pflanzen und DNA-Barcoding Techniken. Deswegen wird empfohlen mehr vorsorgliche Forschung durchzuführen, bevor eine TMP als Lebensmittel oder sogar als sogenanntes „Superfood“ betrachtet wird.



# Contents

<b>I. Statement</b> .....	<b>5</b>
<b>II. Acknowledgement</b> .....	<b>6</b>
<b>III. Abstract</b> .....	<b>7</b>
<b>Contents</b> .....	<b>11</b>
<b>List of abbreviations</b> .....	<b>14</b>
<b>1. Introduction</b> .....	<b>15</b>
<b>1.1 Traditional Medicinal Plants</b> .....	<b>15</b>
<b>1.2 ‘Superfoods’</b> .....	<b>16</b>
1.2.1 Traditional Medicinal Plants as Superfoods.....	16
1.2.2 Food Fraud and DNA-based Authentication Methods .....	18
<b>1.3 Case Study: <i>Ocimum tenuiflorum</i> - Tulsi</b> .....	<b>20</b>
1.3.1 Tulsi: The Indian Tradition.....	20
1.3.2 <i>Ocimum tenuiflorum</i> and the genus <i>Ocimum</i> .....	21
1.3.3 Methyleugenol .....	22
1.3.4 Eugenol <i>O</i> -methyltransferase .....	23
1.3.5 Phenylpropanoid pathway.....	24
1.3.6 Tulsi and DNA barcoding targeting Food fraud .....	26
1.3.7 The ‘Superfood’ Tulsi as Case Study .....	26
<b>1.4 Scope of the study</b> .....	<b>27</b>
<b>2. Materials and Methods</b> .....	<b>28</b>
<b>2.1 Reference plant material</b> .....	<b>28</b>
<b>2.2 Analysis of the Nucleotide and Amino Acidic Sequences of the Enzyme Eugenol <i>O</i>-methyltransferase in the genus <i>Ocimum</i></b> .....	<b>28</b>
2.2.1 Sampling, nucleic acid extraction, cDNA synthesis .....	28
2.2.2 PCR and PCR purification .....	29
2.2.3 A-tailing, cloning and plasmid extraction .....	30
2.2.4 Sequencing and sequences analysis .....	31
2.2.5 Chemical profile analysis .....	32
<b>2.3 Phenylpropanoid Pathway</b> .....	<b>33</b>
2.3.1 Phenylpropanoid pathway enzymes and primers design .....	33
2.3.2 Experiment design and UV-B treatment .....	34

2.3.3 Gene expression analysis (RNA extraction, cDNA synthesis, qPCR analysis).....	36
2.3.4 Chemical profile analysis .....	37
2.3.5 Trichomes analysis .....	37
<b>2.4 Eugenol <i>O</i>-methyltransferase Subcellular localization .....</b>	<b>38</b>
2.4.1 Eugenol <i>O</i> -methyltransferase full sequence .....	38
2.4.2 Localization peptides search .....	39
2.4.3 Plasmid construction for <i>Agrobacterium</i> transformation .....	39
2.4.4 <i>Agrobacterium</i> transformation .....	40
2.4.5 BY2 cells transformation and EOMT localization .....	40
2.4.6 Protein analysis .....	41
<b>2.5 DNA barcoding .....</b>	<b>42</b>
2.5.1 Plant material .....	42
2.5.2 DNA extraction .....	42
2.5.3 DNA barcodes and PCR conditions .....	43
2.5.4 Chemical profile analysis .....	44
<b>3. Results .....</b>	<b>45</b>
<b>3.1 Eugenol <i>O</i>-methyltransferase in the genus <i>Ocimum</i>.....</b>	<b>45</b>
3.1.1 EOMT phylogeny discloses two distinctive groups in <i>Ocimum</i> sp. ....	45
3.1.2 Protein analysis reveals <i>aa</i> residues and motifs in EOMT which are potentially relevant for the synthesis of methyleugenol in the genus <i>Ocimum</i> .....	50
3.1.3 Chemical profile suggest <i>O. tenuiflorum</i> is a methyleugenol producer species compared to other <i>Ocimum</i> sp. ....	52
3.1.4 OMT-based phylogeny with other species than <i>Ocimum</i> shows cluster of order Lamiales .....	54
<b>3.2 Phenylpropanoid Pathway in <i>O. tenuiflorum</i> .....</b>	<b>56</b>
3.2.1 Steady-state relative gene expression of enzymes from the phenylpropanoid pathway reveals similarities and differences in two <i>O. tenuiflorum</i> chemotypes .....	56
3.2.2 Enzymes relative gene expression shows differential regulation mechanisms for methyleugenol synthesis as response to UV-B stress in two <i>O. tenuiflorum</i> chemotypes .....	60
3.2.3 Chemical profile analysis supports findings on differential gene expression and response to UV-B stress for two <i>O. tenuiflorum</i> chemotypes .....	63
3.2.4 Model of the phenylpropanoid pathway for the synthesis of methyleugenol in two <i>O. tenuiflorum</i> chemotypes .....	64

3.3	The subcellular localization of <i>O. tenuiflorum</i> EOMT is in the cytosol and the nucleus .....	66
3.4	DNA barcoding for <i>O. tenuiflorum</i> discrimination .....	69
3.4.1	Trait-independent marker <i>psbA-trnH</i> coupled with RFLP allows to discriminate <i>O. tenuiflorum</i> from other species in a two-step protocol.....	69
3.4.2	The trait-related marker EOMT has potential for discriminating <i>O. tenuiflorum</i> from other species in a one-step protocol .....	72
3.4.3	Multiplexing trait-independent marker ( <i>psbA-trnH</i> ) and trait-related marker (EOMT) allows for <i>O. tenuiflorum</i> discrimination in a one-step protocol .....	75
3.4.4	Commercial samples chemical profile highlights the specificity of the trait-related marker .....	77
3.5	Results summary .....	79
4.	<b>Discussion</b> .....	80
4.1	EOMT variability in the genus <i>Ocimum</i> .....	80
4.1.1	EOMT gene variability among and within <i>Ocimum</i> species. ....	80
4.1.2	EOMT across the genus <i>Ocimum</i> : amino acid residues variability and their potential influence on the chemical profile .....	83
4.1.3	EOMT beyond the genus <i>Ocimum</i> : OMT-based phylogeny .....	84
4.2	Understandings on the Phenylpropanoid Pathway for the synthesis of Methyleugenol in <i>O. tenuiflorum</i> .....	85
4.2.1	Insights on UV-B stress and the Regulation of the Phenylpropanoid Pathway ....	89
4.3	Comprehension of the subcellular localization of the phenylpropanoid pathway and EOMT.....	92
4.4	DNA barcoding for discrimination of <i>O. tenuiflorum</i> in commercial samples .....	93
4.4.1	DNA-based authentication using trait-independent marker needs optimization	93
4.4.2	DNA-based authentication using a trait-related marker and its potential .....	94
4.4.3	DNA-based authentication by combined trait-related and -independent makers is a feasible option in commercial products .....	95
4.4.4	The EOMT DNA marker is not related to methyleugenol content .....	96
5.	<b>Conclusion</b> .....	97
6.	<b>Outlook</b> .....	99
7.	<b>References</b> .....	100
	<b>Appendix</b> .....	111

## List of abbreviations

<b>4CL</b>	4-coumarate-CoA ligase (EC 6.2.1.12)
<b>Amino acid</b>	<i>aa</i>
<b>BSA</b>	Bovine serum albumin
<b>C3H</b>	4-Coumarate 3-hydroxylase (EC 1.14.13.)
<b>C4H</b>	Cinnamate 4-hydroxylase (EC 1.14.14.91)
<b>CCOMT</b>	Caffeoyl-CoA <i>O</i> -methyltransferase (EC 2.1.1.104)
<b>COMT</b>	Caffeic acid <i>O</i> -methyltransferase (EC 2.1.1.68)
<b>cDNA</b>	Complementary DNA
<b>EGS</b>	Eugenol (and Chavicol) synthase (EC 1.1.1.318)
<b>EOMT</b>	Eugenol <i>O</i> -methyltransferase (EC 2.1.1.146)
<b>EtAc</b>	Ethyl acetate
<b>EU</b>	Eugenol
<b>FID</b>	Flame ionization detector
<b>GC</b>	Gas Chromatography
<b>LB</b>	Luria Broth
<b>ME</b>	Methyleugenol
<b>MS</b>	Murashige and Skoog
<b>OMT</b>	<i>O</i> -methyltransferase
<b>PAL</b>	Phenylalanine ammonia lyase (EC 4.3.1.24)
<b>PCR</b>	Polymerase chain reaction
<b>TMP</b>	Traditional medicinal plants

## 1 Introduction

### 1.1 Traditional Medicinal Plants

The term “Traditional medicinal plants” has many definitions. The World Health Organization (WHO) defines *herbal medicine* as “herbs, herbal materials, herbal preparations and finished herbal products that contain, as active ingredients, parts of plants, other plant materials or combinations thereof” (WHO, 2019). While *traditional medicine* is defined as “medicine (that) has a long history. It is the sum total of the knowledge, skill and practices based on the theories, beliefs and experiences indigenous to different cultures, whether explicable or not, used in the maintenance of health as well as in the prevention, diagnosis, improvement or treatment of physical and mental illness” (WHO, 2019). In this research, the term “Traditional medicinal plants” (TMP) is defined as the plants that can be used as herbal medicine and are, at the same time, part of traditional medicine.

Traditional medicinal plants have accompanied human history since ancient times and continue so until today. When the usage of medicinal plants by humans started is unclear. Authors have different suggestions of a starting point that accommodates to the region, culture and archaeological findings. For instance, it is suggested that the Neanderthal might have used herbal medicine given the discovery of medicinal plants in burial sites (Solecki, 1975). Besides archaeological suppositions, ancient societies from countries such as China and India, have written the legacy of traditional medicinal plants usage. Eventually pharmacopoeias were written, replacing traditional knowledge of medicinal plants by detailed specification on the compounds, from natural and synthetic sources, that can be used in pharmaceuticals. But the traditional understanding still survives through sciences such as ethnobotany, which collects the knowledge of how people through history have used plants, specifically native plants of a given region.

In this sense, the usage of plants for health purposes has been historically intertwined with other uses such as worship of deities given their ‘healing powers’, foods and cosmetics. For example, Indian traditional medicine in Ayurveda texts, several types of basil are described for its healing purposes, worship and cooking (Brahmavarchas, 2007). However, basil is extensively used as a spice all over the world, although its healing properties and worship tradition is not wide-spread and most of its consumers are not familiar with them. The previous case is not an isolated one because nowadays several traditional medicinal plants are experiencing globalization through the food industry. This trend started decades ago with dietary supplements or nutraceuticals (Phua *et al.*, 2009; Ćwieląg-Drabek *et al.*, 2020), which now is shifting into the new tendency of the so called ‘superfoods’.

## 1.2 ‘Superfoods’

The term ‘superfoods’ is vastly used for marketing by the food industry to advertise products with a supposed capacity to improve people’s health. However, a proper definition has not been given by neither scientific nor governmental institutions. Rather scholars from several corners of the world have tried to explain the concept. For example, Loyer (2016) discusses that it is not accurate to grant the creation of the word ‘superfood’ to marketing for the food industry, but rather the term ‘superfood’ most likely appeared back in 1915 in allusion to wine. Then, the concept developed over several years under the involvement of different actors. For instance, authors of cooking books used the word ‘superfood’ to probably enhance certain properties of common foods, rich in vitamins, antioxidants, fiber and omega-3. As a result, several common foods and beverages are now called ‘superfoods’ such as broccoli, berries, fish, grains, enriched milk, and green tea (Loyer, 2016; Breeze, 2017).

Eventually, due to the extensive use of the word ‘superfood’ in common foods, the undefined term started to settle around the 1980s (Loyer, 2016; Butterworth *et al.*, 2020). Nowadays, the term ‘superfood’ is being used as an advertising term to promote consumption of various common and exotic foods and dietary supplements associated with a claim to ‘improve health and beauty’. In this research we will consider the ‘superfood’ definition given by the Cambridge Dictionary (online, last visited on June 2021): “a food that is considered to be very good for your health”.

### 1.2.1 Traditional Medicinal Plants as Superfoods

Historically, certain medicinal plants have also been traditionally used as foods. However, the global market of traditional medicinal plants has been shifting more and more to the food industry. First, through dietary supplements or nutraceuticals and nowadays as ‘superfoods’. This raises an issue related to the regulatory frame, which is different for foods and medicinal products. Consequently, the health and scientific community have highlighted the problem of potential toxicity of TMP. Several studies have discussed the importance to consider toxicological relevant compounds, present in dietary supplements that are based on TMP, suggesting to be cautious on their intake (Phua *et al.*, 2009; Charen and Harbord, 2020; Ćwieląg-Drabek *et al.*, 2020; Goswami and Ram, 2017). Considering that even though TMP have a positive effect in treating diseases, as any medicine, there might be side effects. Therefore, unless a risk assessment has approved certain TMP-based products, these should probably not be consumed as everyday meal as suggested by some advertising of ‘superfoods’ (see **Table 1.1**). These includes certain dietary supplements, beverages, and powders.

**Table 1.1** Selected traditional medicinal plants (TMP) that are advertised as ‘superfoods’<sup>1</sup> and contain toxicological relevant compounds.

<b>Botanical name (common name)</b>	<b>Known potentially harmful compound(s)</b>	<b>Potentially harmful effect</b>	<b>References</b>
<i>Curcuma longa</i> (Turmeric)	Cadalene, diphenylheptanoids, diphenylpentanoids, germacrane, guaiane, labdane, monoterpenoids, sesquiterpinoids	Hepatotoxicity	Balaji and Chempakam (2010)
<i>Dioscorea quartiniana</i> (Yam)	Dioscorine, dioscin	Acute tubular necrosis	Charen and Harbord (2020)
<i>Ginkgo biloba</i> (Gingko)	Quercetin	Mutagenic	Corazza <i>et al.</i> (2014)
<i>Lepidium meyenii</i> (Maca)	1R,3S-1-methyl-1,2,3,4-tetrahydro- $\beta$ - carboline- 3-carboxylic acid (MTCA)		Corazza <i>et al.</i> (2014)
<i>Ocimum basilicum</i> (Sweet basil)	Estragole, methyleugenol, safrol	Genotoxic	Jeurissen, (2007) Al-Malahmeh <i>et al.</i> (2017)
<i>Ocimum tenuiflorum</i> (Tulsi or Holy Basil)	Methyleugenol	Genotoxic	Jeurissen <i>et al.</i> (2006)
<i>Tribulus terrestris</i> (Tribulus)	Saponines	Carcinogenesis, hepatotoxicity	Abudayyak <i>et al.</i> (2015) Ryan <i>et al.</i> (2015)
<i>Trigonella foenum- graecum</i> (Fenugreek)	Tannic acid, diosgenin, trigonelline, trigocoumarin, trigomethyl coumarin, gitogenin	Abortive, teratogenic effect	Khalki <i>et al.</i> (2010)
<i>Withania somnifera</i> (Ashwagandha)	Withanone	Genotoxic	Siddiqui <i>et al.</i> (2021)

1. TMP sold as ‘Superfoods’ products in local and online stores in Germany and the European Union.

In order to face this challenging situation research is needed. First, from a botanical perspective, research on TMP is required to understand the accumulation of toxicological relevant compounds. From a practical perspective, these studies help to elucidate the amount and types of metabolites produced by TMP. Second, from a product perspective, products containing toxicologically relevant compounds need to be evaluated through a risk assessment. For example, different suggestions are made for three commercial types of products based on *Ocimum basilicum* (Table 1.1). First, a risk assessment done in pesto for its content of alkenylbenzenes concluded that a periodic consumption of pesto during long time could be of concern (Al-Malahmeh *et al.*, 2017). Second, a risk assessment on *O. basilicum* nutraceuticals suggests that even though basil contains the genotoxic compounds, other metabolites on the basil extract might inhibit the bioactivation of these genotoxic compounds, hence potentially blocking toxic effects (Jeurissen, 2007). Third, *O. basilicum* seeds, sold as ‘basil like chia seeds’ contain significant amounts of estragole (Doerr, 2018), and thus their use in beverages is limited to 3g/200ml (EC, 2017).

The previous examples communicate the need to study medicinal plants and assess the risk of TMP-based products, ideally before their availability to consumers, and not to create panic over all TMP-based products already existing in the market. Having this understanding can help to safe guard consumers' safety. However, there is yet another potential threat in this very fast growing industry of 'superfoods', which is food fraud.

### 1.2.2 Food Fraud and DNA-based Authentication Methods

#### *Food fraud*

Food fraud is the adulteration of food products to obtain an economical advantage (Spink and Moyer, 2011). The dietary supplement industry is a high value one, worth \$40 billion dollars in the U.S. (U.S. Food and Drug Administration, 2019). In this frame, dietary supplements advertised as 'superfoods' given their 'good attributes', are highly valued in the market. And the fast growing trend of 'superfoods', that brings new plant species into the food market, is an attractive scenario for food fraudsters. Similarly, with several traditional plants there is a problem of unintentional surrogation, this happens when the same common name is used for different species (Jürges *et al.*, 2018). Though this adulteration is unintentional the result of deception is the same. Therefore, to fight food fraud, detection methods are important in order to identify the false labelled products. Within the tactics employed for doing so, authentication techniques based on DNA are among the preferred ones for authenticating 'superfoods' and plant-based food products (Jürges *et al.*, 2018; Wetters *et al.*, 2018; Tichy *et al.*, 2020).

#### *DNA-based authentication methods*

There are a number of DNA-based authentication techniques, examples of these are listed in **Table 1.2**. Authentication techniques were developed in the frame of species identification for taxonomical purposes (Hebert *et al.*, 2003). Eventually, they serve commercial purposes and have been used for the identification of TMP for drugs formulations or already as part of a medicinal product. Shaw *et al.* (2002) explains that the pharmacological properties of several Traditional Chinese Medicine derived from plant material, have been affected by lack of quality control in the material used. This problem can be attribute to fraud but also to unintentional substitutions in cases where different species share a common name, or when different common names are given to a single species (Shaw *et al.*, 2002; Jürges *et al.*, 2018). Therefore, molecular methods have been helpful in identifying plant material in plant pharmaceuticals, expanding their use in food products (Galimberti *et al.*, 2013; Haynes *et al.*, 2019).

Given the problem of TMP being sold as 'superfood' this research will explore more on the topic using *Ocimum tenuiflorum* (Tulsi – see **Table 1.1**) as a case study.

**Table 1.2.** Selected DNA-based methods used in the identification of traditional medicinal plants (TMP), plant-based foods and beverages (PFB) and plant-based nutraceuticals (PN).

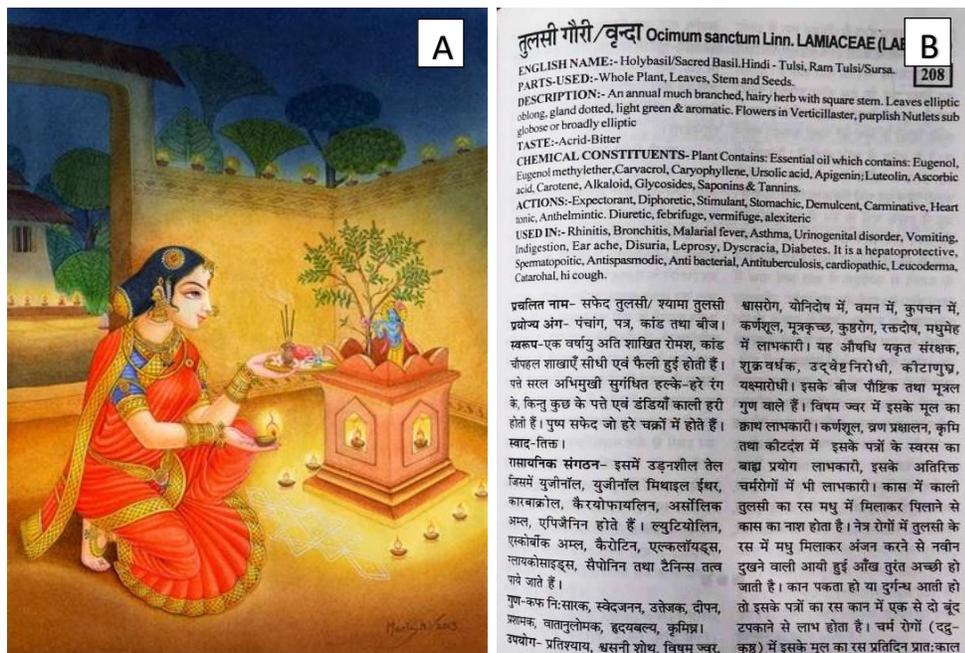
DNA-based method	Sample	References
<b>Polymerase chain reaction (PCR).</b> A DNA section is amplified by a polymerase enzyme using primers that are specific for a certain region. Further sequencing is used to obtain the nucleotide sequence of the amplified fragment. DNA sequences that are specific for a species can be used as an identification tool.	TMP PFB	Shaw <i>et al.</i> (2002), de Boer <i>et al.</i> (2015), Coutinho Moraes <i>et al.</i> (2015), Gong <i>et al.</i> (2018) Galimberti <i>et al.</i> (2013), Ramakrishnan <i>et al.</i> (2015)
<b>Simple sequence repeats (SSRs).</b> Also called microsatellites, is a PCR technique based on the amplification of short tandem repeats resulting in different patterns when amplified in different species.	TMP PFB	Shaw <i>et al.</i> (2002), de Boer <i>et al.</i> (2015) Lopez-Vizcón and Ortega (2012), Fang <i>et al.</i> (2012), Wu <i>et al.</i> (2014), Ramakrishnan <i>et al.</i> (2015)
<b>Amplified fragment length polymorphism (AFLP).</b> In this method, DNA is digested using a restriction enzyme. An adaptor is then ligated to the resulting DNA fragments and a PCR is done using primers matching the adaptors. A band pattern is then obtained based on the restriction enzyme digestion.	TMP PFB	Shaw <i>et al.</i> (2002), de Boer <i>et al.</i> (2015) Torricelli <i>et al.</i> (2012), Ramakrishnan <i>et al.</i> (2015)
<b>Restriction fragment length polymorphism (RFLP).</b> DNA is amplified by PCR. The resulting fragment is digested using a restriction enzyme. A fragment pattern is obtained based on the digestion.	TMP PFB	Shaw <i>et al.</i> (2002), de Boer <i>et al.</i> (2015) Horn <i>et al.</i> (2014), Ramakrishnan <i>et al.</i> (2015), Jürges <i>et al.</i> (2018)
<b>Amplification refractory mutation system (ARMS).</b> The method is based on the amplification by PCR of one section of the DNA using a pair of primers plus a third primer designed on a single mutations present on target samples, resulting in a double band pattern for the samples having the mutation.	TMP PFB	Shaw <i>et al.</i> (2002), de Boer <i>et al.</i> (2015) Horn <i>et al.</i> (2014), Wetters <i>et al.</i> (2018)
<b>Random amplified polymorphic DNA (RAPD).</b> Arbitrary primers are used in a PCR reaction for DNA amplification resulting in a banding fingerprint. A RAPD derived method is Sequence characteristic amplified region (SCAR). Here, a species-specific band obtained from RAPD is sequenced and primers for this specific section are designed in order to obtain a species-specific amplicon.	TMP PFB	Shaw <i>et al.</i> (2002), de Boer <i>et al.</i> (2015), Chowdhury <i>et al.</i> (2017), Ravi <i>et al.</i> (2021) Wang <i>et al.</i> (2012), Ramakrishnan <i>et al.</i> (2015)
<b>Real-time quantitative PCR (RT-qPCR).</b> The DNA amplification during the PCR is measured by collecting the fluorescence emitted by added dyes that intercalate with double-stranded DNA. The signals obtained can vary according to the DNA amount and DNA fragment length, among others.	PFB	An <i>et al.</i> (2019), Tichy <i>et al.</i> (2020)
<b>Metabarcoding.</b> Combines DNA barcoding and high-throughput sequencing, allowing the identification of a number of species at one time. Here, next generation sequencing (NGS) are methods capable of sequence mixtures of DNA in a sample. The sensitivity of NGS has been proved to be high, however, the costs exceed those of the previous techniques.	TMP PFB PN	de Boer <i>et al.</i> (2015), Kreuzer <i>et al.</i> (2019) Raclariu <i>et al.</i> (2017), Barbosa <i>et al.</i> (2019), Haynes <i>et al.</i> (2019) Coutinho Moraes <i>et al.</i> (2015), Ivanova <i>et al.</i> (2016), Raclariu <i>et al.</i> (2017), Barbosa <i>et al.</i> (2019)

## 1.3 Case Study: *Ocimum tenuiflorum* - Tulsi

### 1.3.1 Tulsi: The Indian Tradition

The ethnobotany of Tulsi is rooted in India, where it apparently originated, and where its usage is quite versatile. The accepted botanical name for Tulsi is *Ocimum tenuiflorum* ([www.theplantlist.org](http://www.theplantlist.org)). However, in India, given the usage of Tulsi for worshipping, the synonym *Ocimum sanctum* is preferred. The Latin name '*sanctum*' means "that which is holy". This concept of holiness also extends to the common given names of Tulsi: Holy basil or Sacred basil. The origins of using Tulsi as worship plant, partially originated in its medicinal properties. It is a common practice of the Hindus to offer Tulsi leaves to the gods as a food offering (**Figure 1.1 a**). In particular, Tulsi is employed for worshipping the God Vishnu. Currently, it is very common for Hindu households in India to keep a plant of Tulsi in their homes.

In Indian traditional medicine, Tulsi is described in Ayurveda as a medicinal plant (see **Figure 1.1 b**), and it is used to treat rhinitis, bronchitis, malaria fever, asthma, vomiting, indigestion, leprosy and diabetes among others (Brahmavarchas, 2007; Pullaiah *et al.*, 2017). It is also described as hepatoprotective, antispasmodic and antibacterial (Brahmavarchas, 2007; Pullaiah *et al.*, 2017). Besides *O. tenuiflorum*, also *O. basilicum*, *O. gratissimum* and *O. americanum* are described in Ayurveda as medicinal plants (Brahmavarchas, 2007; Pullaiah *et al.*, 2017).



**Figure 1.1** Tulsi, the plant in India's tradition. **a** Tulsi as a plant for worship -image copyrights: Manish Vyas (web-page: [www.manishvyas.ch](http://www.manishvyas.ch)) **b** Tulsi described as a medicinal plant in Ayurveda by Brahmavarchas (2007) -image taken by: Dr. Vaidurya Pratap Sahi.

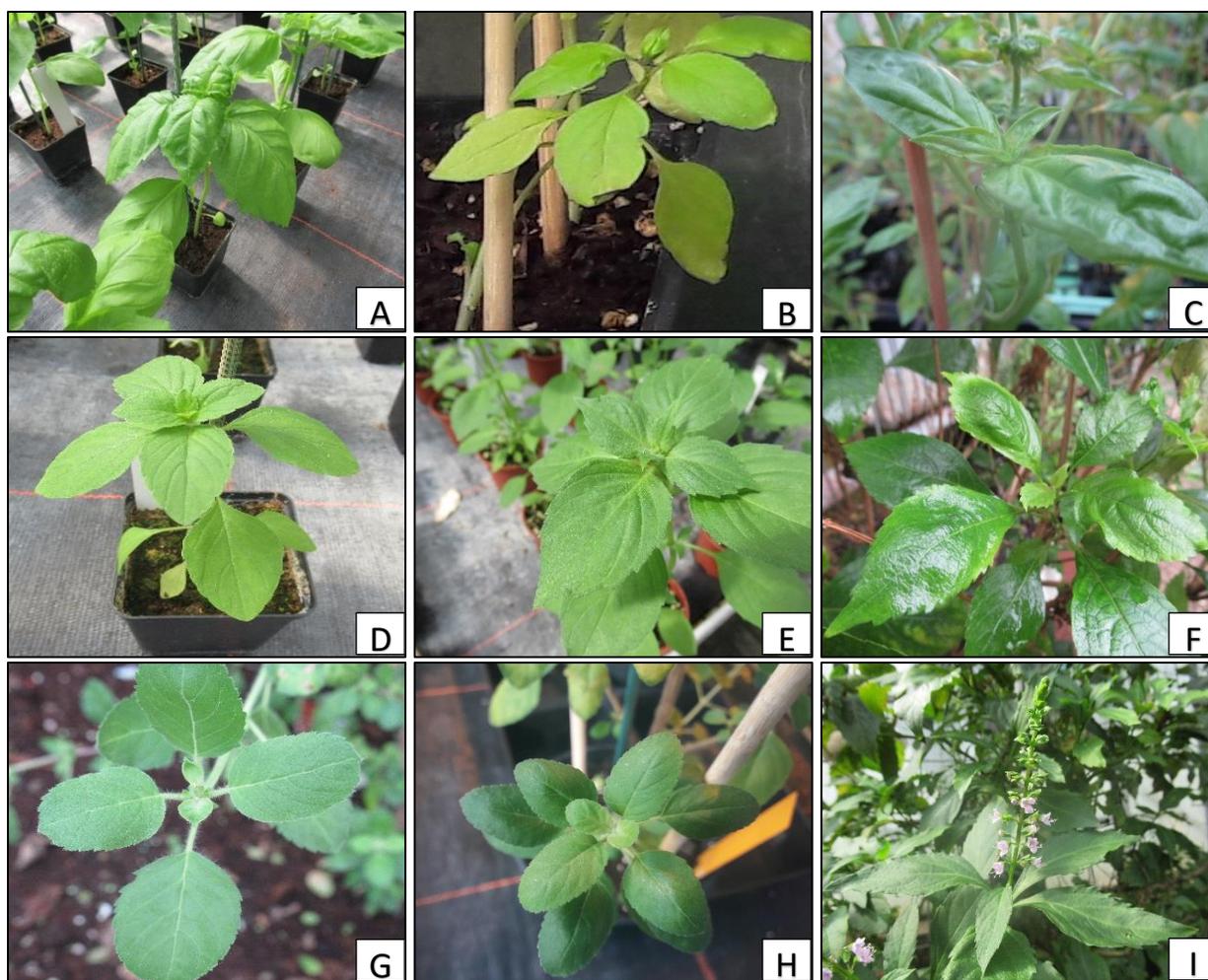
### 1.3.2 *Ocimum tenuiflorum* and the genus *Ocimum*

*O. tenuiflorum* is one of the 66 *Ocimum* species described within the genus *Ocimum* from the Lamiaceae family ([www.theplantlist.org](http://www.theplantlist.org)). It is unclear how all these 66 *Ocimum* species are clustered, however, work done by Jürges *et al.* (2018) has been able to distinguish at least four haplotypes within the genus by using the barcoding marker *rbcL*. The resulting haplotypes are haplotype I which involves *O. basilicum*, *O. americanum*, *O. x africanum* and *O. kilimandscharicum*; haplotype II comprising *O. tenuiflorum* and *O. campechianum*; haplotype III consisting of *O. gratissimum* and haplotype IV consisting of *O. filamentosum*. Examples of haplotypes I, II and III can be seen in **Figure 1.2**. Hence, *O. tenuiflorum* belongs to a very distinctive group within the genus *Ocimum*. Moreover, different chemotypes can be found within the species.

Particularly, there are two distinctive chemotypes in *O. tenuiflorum* commonly known as Krishna and Rama Tulsi (**Figure 1.2 g and h**). Krishna Tulsi has purple leaves and stems due to the anthocyanin content, which is highly produced when plants are exposed to UV radiation. Whereas Rama Tulsi does not produce as much anthocyanins as Krishna. Therefore, the colour of the leaves and stems is green even under exposure to UV radiation. Such dissimilarities emphasise the different chemical profiles that can be found among *O. tenuiflorum* chemotypes.

Research shows that different compounds can be found in Tulsi, such as monoterpenes (i.e. linalool, eucalyptol), sesquiterpenes (i.e.  $\beta$ -caryophyllene, ursolic acid), and phenylpropenes (i.e. eugenol, isoeugenol, methyl(iso)eugenol) among others (Raina and Misra, 2018; Upadhyay *et al.*, 2015; Rastogi *et al.*, 2014; Singh *et al.*, 2015; Jürges *et al.*, 2018). The production and accumulation of the different metabolites depends on the *O. tenuiflorum* genotype and chemotype, physiological aspects and environmental conditions (Jürges *et al.*, 2009; Kumari and Agrawal, 2011; Vyas *et al.*, 2014; Malav *et al.*, 2015; Raina and Misra, 2018).

Compounds accumulation, from a commercial perspective, becomes relevant when a desired (add value to products) or an undesired (rest value to products) compound accumulates significantly. In the particular case of *O. tenuiflorum*, an undesired metabolite that has a tendency to accumulate in the species is methyleugenol (Tan and Nishida, 2012; Raina and Misra, 2018; Rajeswara Rao *et al.*, 2011; Joshi, 2013).



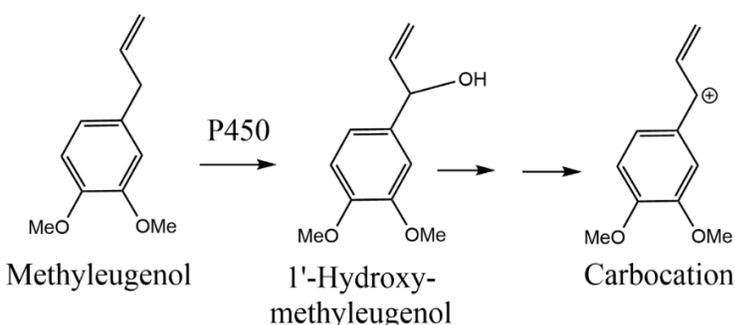
**Figure 1.2.** *Ocimum* sp. from the plant collection at the Botanical Garden from the Karlsruhe Institute of Technology. **a** *O. basilicum* (ID 9056). **b** *O. americanum* (ID 7811). **c** *O. x africanum* (ID 7537). **d** *O. kilimandscharicum* (ID 7809). **e** *O. kilimandscharicum* (7810). **f** *O. gratissimum* (ID 5749). **g** *O. tenuiflorum* “Krishna” (ID 5751). **h** *O. tenuiflorum* “Rama” (ID 8256). **i** *O. campechianum* (ID 7564).

### 1.3.3 Methyleugenol

As mentioned above, methyleugenol is a phenylpropene, and it can be found in within species of the genus *Ocimum* and several other genera such as *Heterotropa* (Asteraceae), *Artemisia* (Asteraceae), *Cinnamomum* (Lauraceae), *Melaleuca* (Myrtaceae) and *Croton* (Euphorbiaceae) (Tan and Nishida, 2012). The role of ME in plants is divergent, it has been described as defence against microbes, insects and herbivores, as well as, an attractant for certain insect species to deal with a particular pest or to attract pollinators (Tan and Nishida, 2012).

Humans have been using ME as flavouring agent, scent agent, and insecticide (Tan and Nishida, 2012). This last use of ME as to control pests is based on its capability

of being toxic. Moreover, according to an *Opinion of the Scientific Committee on Food on Methyleugenol* from the European Commission, it was concluded that ME is genotoxic and carcinogenic (EC-SCF, 2001). The biological activation of methyleugenol as genotoxic compound has been suggested by Jeurissen *et al.* (2006) as shown in **Figure 1.3**. Basically, ME is transformed into 1'-Hydroxy-methyleugenol by the cytochrome P450 following by sulfotransferase enzymes mediated reactions resulting in the formation of a methyleugenol molecule with a positive charge, the carbocation. This charged molecule can origin DNA adducts (DNA segment bond to a chemical causing damage) which could eventually develop into liver tumours (Jeurissen *et al.*, 2006; Rajalakshmi *et al.*, 2015). Subsequently, maximum levels of ME were established for food products (EC, 2008). Further, because ME is a toxicological relevant compound and it naturally occurs in *O. tenuiflorum*, ME and the enzyme responsible for its synthesis, the eugenol *O*-methyltransferase, will be a central focus of this study.

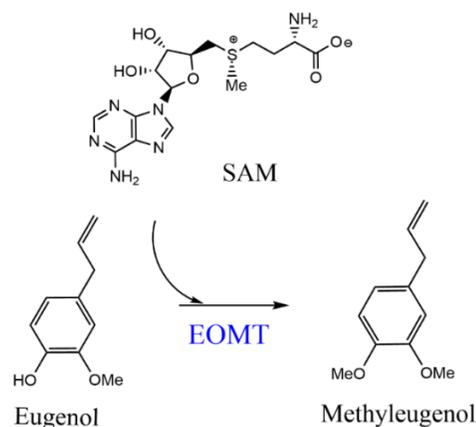


**Figure 1.3** Reaction model proposed for the bioactivation of methyleugenol starting with the hydroxylation by the cytochrome P450 (P450) and finally resulting in the carbocation. Reaction model modified from the suggested model by Jeurissen *et al.* (2006).

### 1.3.4 Eugenol *O*-methyltransferase

Methyleugenol is synthesized by the enzyme eugenol *O*-methyltransferase (EOMT, EC 2.1.1.146). The reaction is described in **Figure 1.4**. EOMT is responsible for the methylation of eugenol (substrate) forming methyleugenol (product) by the addition of a methyl group obtained from S-adenosyl methionine as donor molecule (<https://enzyme.expasy.org/EC/2.1.1.146>).

EOMT belongs to the *O*-methyltransferase (OMT) superfamily which has been described to participate in the methylation of flavonoids and phenylpropanes in plants (Ibrahim *et al.*, 1998). EOMT has also been described as a promiscuous enzyme in different plant species. For example, studies on the recombinant enzyme from *O. basilicum*, have shown that EOMT is capable of synthesising the conversion of (iso)eugenol to methyl(iso)eugenol, and chavicol to methylchavicol (Gang, Lavid, *et al.*, 2002). Differences on the specificity of the EOMT could be attributed to variability of the *aa* sequences residues, and triggered by substrate availability from the phenylpropanoid pathway.



**Figure 1.4** Reaction catalysed by eugenol *O*-methyltransferase (EOMT) using eugenol as a substrate and S-adenosyl methionine (SAM) as a methyl group donor, synthesising methyleugenol.

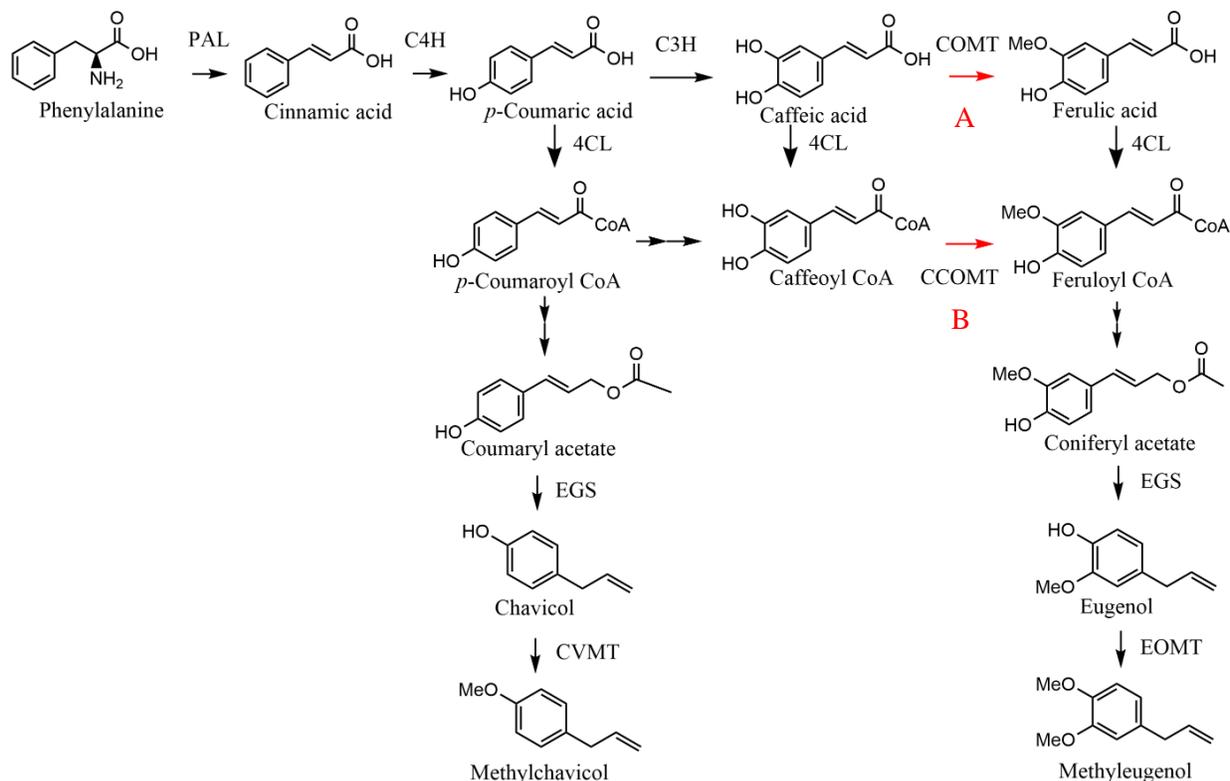
### 1.3.5 Phenylpropanoid pathway

The reaction shown in **Figure 1.4** is part of the phenylpropanoid pathway (**Figure 1.5**). The phenylpropanoid pathway starts with the conversion of phenylalanine to cinnamic acid by the enzyme phenylalanine ammonia lyase (PAL, EC 4.3.1.24), followed by the conversion of cinnamic acid into coumaric acid by the enzyme cinnamate 4-hydroxylase (C4H, EC 1.14.14.91) (Gang, Lavid, *et al.*, 2002). The precise reactions leading to the synthesis of ME are unclear. Moreover, a review of literature reveals that there is no consensus on the phenylpropanoid pathway upstream reactions towards the synthesis of ME.

Three main possible streams of reactions are described for the synthesis of ME in *Ocimum* sp. The first one goes through caffeoyl-CoA, potentially originated from coumaroyl-CoA, which is converted directly to feruloyl-CoA by the enzyme caffeoyl-CoA *O*-methyltransferase (CCOMT, EC 2.1.1.104) (**Figure 1.5b**). The pathway via coumaroyl-CoA has been widely studied and suggested as the principal pathway in evolution, when plant conquer solid land. From coumaroyl-CoA, the first suggested compounds to be synthesised are lignin and flavonoids, later anthocyanins were derived from flavonoids (Emiliani *et al.*, 2009; Buchanan *et al.*, 2015). The second one is via ferulic acid which is synthesised by the enzyme caffeic acid *O*-methyltransferase (COMT, EC 2.1.1.68) using caffeic acid as substrate (**Figure 1.5a**). The third possibility, that several researchers propose is that these streams of reactions are connected, thus both paths seemed to provide substrates for the synthesis of methyleugenol (Gang *et al.*, 2001; Gang, Lavid, *et al.*, 2002; Khakdan *et al.*, 2017; Rastogi *et al.*, 2013; Rastogi *et al.*, 2014; Kapteyn *et al.*, 2007; Xie *et al.*, 2008).

This scenario leaves several open questions on the regulation processes under which ME synthesized in *O. tenuiflorum*. For instance, is there any path that plays a more important role on the synthesis of ME? Or, which enzymes are acting as the limiting

steps on these pathways? Or, how many different mechanisms of regulation exist? Therefore, studies on the phenylpropanoid pathway enzymes are crucial to better understand path regulation.



**Figure 1.5.** Suggested schematic representation of methyleugenol and methylchavicol synthesis through the Phenylpropanoid pathway in *Ocimum* sp. **a** Suggested phenylpropanoid pathway via ferulic acid for the synthesis of methyleugenol. **b** Suggested phenylpropanoid pathway via caffeoyl-CoA for the synthesis of methyleugenol. Double arrows indicate intermediated suggested reactions not shown. The enzymes shown in the pathway are: Phenylalanine ammonia lyase (PAL), Cinnamate 4-hydroxylase (C4H), 4-coumarate 3-hydroxylase (C3H), 4-Coumarate-CoA ligase (4CL), Caffeic acid *O*-methyltransferase (COMT), Caffeoyl-CoA *O*-methyltransferase (CCOMT), Eugenol (and Chavicol) synthase (EGS), Eugenol *O*-methyltransferase (EOMT) and Chavicol *O*-methyltransferase (CVOMT).

Another important aspect of the phenylpropanoid pathway -for the synthesis of methyleugenol- that remains unclear, is where exactly in the cell are these reactions taking place. The enzyme PAL has been localized in the cytoplasm, which is colocalized with C4H -anchored in the endoplasmic reticulum-; while EGS has been localized in the cytoplasm (Achnine *et al.*, 2004; Reddy *et al.*, 2021). Therefore, one

could hypothesise that EOMT is also located in the cytosol. It is important to understand the subcellular localization of an enzyme because it can give insights on the complexity of a metabolic pathway regulation.

### 1.3.6 Tulsi and DNA barcoding targeting Food fraud

In order to detect food fraud there are several DNA authentication methods discussed in section 1.2.2 (see **Table 1.2**). These authentication methods are based on different DNA markers. DNA or molecular markers are basically a section of the DNA used to identify an organism at a taxonomical level (Hollingsworth *et al.*, 2011). Molecular markers are desirable to be universal, have power of discrimination and a quality sequence (Hollingsworth *et al.*, 2011). For instance, RAPD and ISSR markers have been used to differentiate and identify *Ocimum* sp. (Patel *et al.*, 2015), however, these type of markers lack practicability on commercial samples, where the DNA can be highly degraded. These are limitations to consider when choosing a fitting marker. In the case of *O. tenuiflorum*, chloroplastidic markers have been found to be better for identifying Tulsi in commercial samples (Jürges *et al.*, 2018; Sgamma *et al.*, 2017). Particularly the *psbA-trnH intergenic spacer* which was added to the British Pharmacopeia for the identification of *O. tenuiflorum*, has shown to be a good marker for identification of *O. tenuiflorum* (Christina and Annamalai, 2014; Jürges *et al.*, 2018; Sgamma *et al.*, 2017). Nonetheless, when handling large numbers of samples, the process of PCR followed by gel electrophoresis, sequencing and sequence analysis, can be time consuming. Thus, optimization is desirable to shorten the process time. Under this frame, *O. tenuiflorum* is used as a case study in this research, where the chemical profile of commercial samples will be studied, and optimization of current authentication methods of the herb via the enzyme eugenol *O*-methyltransferase will be aimed.

### 1.3.7 The ‘Superfood’ Tulsi as Case Study

Tulsi has been highly marketed as a ‘superfood’ that contains antioxidants, has anti-stress properties and helps to control anxiety. The main types of Tulsi products found in the market are teas and powders with a wide range of prices. However, considering Tulsi as a ‘superfood’ possesses challenges. Firstly, the presence of methyleugenol, whose mechanisms of its accumulation are poorly understood, might be of concern for consumers’ safety. Secondly, Tulsi is a common name for several species (Jürges *et al.*, 2018), which can lead to substitution of the plant in commercial products, besides the intentional adulteration, both situations leading to food fraud.

## 1.4 Scope of the study

Traditional medicinal plants (TMP) that contain toxicological relevant compounds, are increasingly being sold as food products. This transition is developing faster than the capability of the scientific community to individually evaluate the safety and botanical aspects of each TMP that becomes a food-trend. A botanical evaluation might include to determine inaccuracies between plant common names vs. botanical name, the presence and concentration of toxic compounds vs. parts of the plant, and the existence of species chemotypes that vary in their content of toxicological relevant compounds. Therefore, there is a need to botanically evaluate TMP -containing toxicological relevant compounds- that are transitioning to food products. For doing so, *Ocimum tenuiflorum* was used as a case study in this research. *O. tenuiflorum*, commonly known as Tulsi or Holy Basil, is a TMP that is now part of the ‘superfoods’ trend. It produces the genotoxic compound methyleugenol (ME) by the enzyme eugenol *O*-methyltransferase (EOMT), and surrogation of Tulsi in commercial products exists due to confusions of its common name. Under this frame the following research questions were formulated:

- Which genetic variation of the enzyme eugenol *O*-methyltransferase exists across the genus *Ocimum*, and could this influence the methyleugenol synthesis?

ME accumulation varies among *Ocimum* sp., therefore, variations on the EOMT enzyme are expected. However, it is unknown how strong these variations are and how these might be related to the chemical profile in the genus *Ocimum*.

- How is the synthesis of methyleugenol regulated in *O. tenuiflorum*?

Previous studies suggest different ways on the phenylpropanoid pathway for the synthesis of ME in *Ocimum* sp., as elaborated in section 1.3.3.2. Knowledge on the metabolic pathway and the key enzymes playing a role in the metabolism will help to understand the accumulation of ME.

- Where is the enzyme eugenol *O*-methyltransferase located inside the cell?

The subcellular localization of the enzyme EOMT is believed to be in the cytosol as discussed in section 1.3.5. However, it has not yet been investigated. Knowledge on the subcellular localization of enzymes helps to understand metabolites synthesis and accumulation processes.

- What DNA authentication method serves best to discriminate *O. tenuiflorum* accurately and efficiently to detect food fraud?

Food fraud is a current problem in the food industry and DNA-based authentication tools have helped to fight the issue. The authentication work on this thesis aims to explore different approaches in order to design solutions for a potential transfer in industrial application.

## 2 Materials and Methods

### 2.1 Reference plant material

Reference plants were grown in the Botanical Garden of the Karlsruhe Institute of Technology (**Table 2.1**). Plants were cultivated in soil substrate (Floraton 3, Floragard Vertriebs-GmbH, Oldenburg, Germany), with a light/darkness ratio of 16/8 hours and the temperature was set at 25°C. Plants were grown for 6 weeks before sampling.

**Table 2.1.** Reference plant material used in this study; ID of the voucher specimen cultivated in the Botanical Garden of the KIT; scientific name; common name for each species.

Accession number	Species	Common name
5192	<i>Ocimum basilicum</i> L.	Sweet Basil
7811	<i>Ocimum americanum</i>	Hoary Basil
7537	<i>Ocimum x africanum</i> Lour.	Lemon Basil
5748	<i>Ocimum x africanum</i> Lour.	Lemon Basil
5751	<i>Ocimum tenuiflorum</i> L.	Krishna Tulsi – Holy Basil
8097	<i>Ocimum tenuiflorum</i> L.	Krishna Tulsi – Holy Basil
8099	<i>Ocimum tenuiflorum</i> L.	Krishna Tulsi – Holy Basil
8256	<i>Ocimum tenuiflorum</i> L.	Rama Tulsi – Holy Basil
8257	<i>Ocimum tenuiflorum</i> L.	Krishna Tulsi – Holy Basil
8258	<i>Ocimum</i> sp.	Vana Tulsi
9056	<i>Ocimum basilicum</i> L.	Sweet Basil
5749	<i>Ocimum gratissimum</i>	Clove Basil
7564	<i>Ocimum campechianum</i> Mill.	Wild Sweet Basil
7809	<i>Ocimum kilimandscharicum</i> Gürke	Camphor Basil
7810	<i>Ocimum kilimandscharicum</i> Gürke	Camphor Basil
5391	<i>Mentha spicata</i> Crispa	Spearmint
5393	<i>Mentha x piperita</i>	Pepper Mint

### 2.2 Analysis of the Nucleotide and Amino Acidic Sequences of the Enzyme Eugenol *O*-methyltransferase in the genus *Ocimum*

#### 2.2.1 Sampling, nucleic acid extraction, cDNA synthesis

Leaves from reference plants (see **Table 2.1**) were sampled and immediately frozen in liquid nitrogen to be either directly processed or stored at -80°C for further analysis.

#### DNA extraction

The plant material was ground in liquid nitrogen using a mortar and pestle. Between 30 and 60 mg of the powdered leaves was used for extraction. Genomic DNA was isolated using the Invisorb Spin Plant Mini Kit (Stratec molecular, Birkenfeld) following the company's protocol. DNA was obtained using the kit's elution buffer. The

concentration and purity of the genomic DNA (gDNA) was evaluated using a NanoDrop ND-1000 spectrophotometer (Peqlab Biotechnologie GmbH, Erlangen). Samples were stored at -20°C for later experiments.

### RNA extraction and cDNA synthesis

The plant material was ground in liquid nitrogen using a mortar and pestle. RNA was extracted with the Spectrum Plant Total RNA Kit (Sigma-Aldrich, Darmstadt) following the manufacturer instructions for protocol A. To maximize the RNA concentration, the elution step was repeated using the same supernatant. The concentration and RNA purity was measured by spectrophotometry. A volume for 1 µg of RNA was calculated to synthesise Complementary DNA (cDNA). The volume was adjusted to 14.6 µl with nuclease free water, 1 µl of 10 mM dNTPs and 0.4 µl of 100 µM Oligo dT were added. The reactions were incubated in a Primus 96 advanced® thermal cycler (PEQLAB, Erlangen, Germany) at 65°C for 5 minutes and put immediately in ice. Subsequently, 4 µl of the following master mix was added to each sample: 2 µl of Reverse transcriptase buffer (New England Biolabs, Frankfurt), 0.25 µl of M-MuLV Reverse Transcriptase (New England Biolabs, Frankfurt), 0.5 µl of RNase inhibitor (New England Biolabs, Frankfurt) and 1.25 µl nuclease free water. The reaction was carried for 60 min at 42°C, following 10 min at 90°C and lowered to 12 °C. The produced cDNA was stored at -20°C for later experiments.

### **2.2.2 PCR and PCR purification**

In order to get the sequence of the EOMT enzyme, the amplification of it was done using gDNA (coding and non-coding regions) and cDNA (coding region) as templates. A proofreading enzyme Q5® *High-Fidelity* DNA polymerase (New England Biolabs, Frankfurt) was used in a semi-qPCR. The reaction set up and components, were as per the enzyme provider instructions, adjusted for a 30 µl reaction. Betaine 5M and/or BSA 10mg/ml was used as enhancers. The reactions were carried in a FlexCycler thermocycler (Analytik Jena AG, Jena). The thermocycling conditions were an initial denaturation 98°C for 30 sec, followed by 35 cycles of 30 sec denaturation at 98°C, 30 sec annealing at 56°C and 30 sec elongation at 72°C. A final 2 min elongation at 72°C was added. The primers used are listed in **Table 2.2**. The result of the PCR was evaluated by agarose gel electrophoresis. Then, 2 µl of 30 µl reaction was loaded in a 1.5% agarose gel prepared in 0.5 X TAE buffer. The samples were run for 30 min at 100 V in an electrophoresis chamber (MupidOne, Advance, Mupid CO., Tokio, Japan). DNA amplification was visualized using SYBRsafe (Invitrogen, Thermo Fisher Scientific, Germany) or Midori green Xtra (Nippon Genetics Europe GmbH), and a blue light gel scanner (Safe Imager, Invitrogen GmbH, Karlsruhe). The size of amplicons was determined by running along the gel a 100 bp standard ladder (New England Biolabs, Frankfurt). The rest of the reaction was purified from the PCR components using the PCR purification MSB Spin PCRapace kit (Strattec molecular, Birkenfeld)

following manufacturer instructions and resuspending the DNA in 20  $\mu$ l nuclease free water (Lonza, Biozym). Optionally, samples were purified by gel excision and gel clean-up with Invisorb® Fragment CleanUp kit (Stratag molecular, Birkenfeld), when unspecific amplifications were observed (below 1000 bp or above 1200bp).

**Table 2.2.** Primers used for obtaining the Eugenol *O*-methyltransferase enzyme sequence for phylogeny studies.

Name	Purpose	Sequence 5'-3'	Reference/provider
EOMT <sub>fwR</sub>	EOMT	TGTCGACAGAGCAACTTCTT	Renu <i>et al.</i> (2014)
EOMT <sub>revR</sub>	amplification	GGATAAGCCTCTATGAGAGACC	
M13 fw	Plasmid	CGCCAGGGTTTTCCAGTCACGAC	Promega
M13 rv	Sequencing	TCACACAGGAAACAGCTATGAC	M13 primers

### 2.2.3 A-tailing, cloning and plasmid extraction

An A-tail is needed in the PCR product for a proper insertion into a plasmid. The reaction consisted of 7  $\mu$ l PCR product, 1  $\mu$ l of 10x Thermopol Buffer (New England Biolabs), 0.2  $\mu$ l of 10 mM dNTPs (New England Biolabs), one unit of Taq polymerase (New England Biolabs) and nuclease free water to a total volume of 10  $\mu$ l. The reaction was immediately incubated for 60 min at 68°C in a thermocycler. After the incubation step followed the ligation in a pGEM-T Easy Vector System (Promega GmbH, Mannheim). The ligation was completed following the manufacturer instructions using 3  $\mu$ l of the A-tailing product and an overnight incubation at 4°C for PCR product-plasmid ligation. The 10  $\mu$ l ligation reaction was then added to 45  $\mu$ l of competent cells *E. coli* strain DH5- $\alpha$ . The cells and the ligation were gently mixed and incubated in ice for 30 min. Cells were heat shocked at 42°C for 45 sec. The samples were immediately returned to ice for 2 min. 950  $\mu$ l of sterile LB media was added to each cell sample and incubated at 37°C, shaking (150 rpm) for 2 hours. After this time the samples were centrifuged at 3000 rpm for 2 min. 850  $\mu$ l of the LB media was removed and 150  $\mu$ l were plated in LB-agar plates with ampicillin (0.1%) (Carl Roth GmbH + Co. KG, Karlsruhe), IPTG (0.2%) (Carl Roth GmbH + Co. KG, Karlsruhe) and X-gal (0.8%) (Biomol GmbH, Hamburg). Plates were incubated overnight at 37°C. Blue-white (negative-positive) colonies screening was made on the next day. White colonies were picked and transferred to single eppis containing 1 ml liquid LB medium and 0.1% ampicillin. Selected colonies were incubated at 37°C for 2 hours after which 1  $\mu$ l of the incubated cells was taken to perform a standard semi-qPCR reaction which had 1  $\mu$ l of thermos polymerase buffer (New England Biolabs), 0.2  $\mu$ l 10 mM dNTPs (New England Biolabs), 0.2  $\mu$ l 10  $\mu$ M of each EOMT<sub>fwR</sub> and EOMT<sub>revR</sub> primers (see **Table 2.2**), 1  $\mu$ l of 5M Betaine and nuclease free water to a final volume of 10  $\mu$ l. The thermocycling parameters used were as described by Renu *et al.* (2014). The result of the PCR was evaluated by agarose gel electrophoresis as

described in section 2.2.2. the colonies giving a positive amplification were transferred to a 4 ml LB media tube and incubated overnight at 37°C and shaking at 150 rpm. The next day the plasmids were isolated from the bacteria using the Roti-Prep Plasmid Mini kit (Carl Roth GmbH + Co. KG, Karlsruhe) according to the manufacturer's instructions.

#### 2.2.4 Sequencing and sequences analysis

Sanger sequencing was done by MacroGen or Eurofines Genomics using the universal M13 primers (Table 2.2), forward and reverse, to obtain the two direction of the gene. The sequences were first analysed with the software FinchTV version 1.4.0 (Geospiza Inc., Seattle, WA; Windows) in order to do a quality check. Individual sequences were later transferred to BioEdit version 7.0.4.1 (Hall, 1999) to align the two directions of each gene and to obtain the complete consensus region with the overlapping fragments. The sequences were then aligned by MEGA X version 10.0.4 (Kumar *et al.*, 2018) using the MUSCLE algorithm (Edgar, 2004) integrated in MEGA X. Comparisons between the sequences from gDNA and cDNA were used to determine the exons and introns regions of the EOMT. The results were also compared to the full mRNA sequence of EOMT obtained by Gang *et al.* (2002), who performed experiments with the recombinant EOMT enzyme from *Ocimum basilicum*. This model enzyme was used to determine the codon reading frame and to see whether a sequence was to be considered as EOMT based on a serine-residue motif. The sequences were later used for phylogenetic analysis.

#### Phylogeny

The sequences obtained were analysed with MEGA X using the Neighbour-joining analysis tool for creating phylogeny tree. Bootstrap analysis with 1000 repetitions was used to determine the significance of the branching in a tree. Evolutionary distances were computed by MEGA X using the Tamura-Nei method (Tamura and Nei, 1993). The analysis was repeated separately for different groups such as gDNA, exons together with cDNA and introns. In exons and cDNA analysis, the sequences were verified by translating the nucleotide sequences into amino acid sequences to detect stop codons (potential pseudogenes) within a sequence and discard these for further analysis.

#### Phylogeny with other species than *Ocimum* sp.

The sequence of EOMT obtained from *O. tenuiflorum* ID 5751 was used with a template in the NCBI website nucleotide blast tool (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, visited on June 2021) in order to find similar sequences. Similar sequences belonging to *Ocimum* sp. and other species were selected. Sequences including “hypothetical” or “predicted” in the title were avoided. Selected OMT sequences were analysed using the program MEGA X. These were

aligned using the MUSCLE algorithm (Edgar, 2004) integrated in MEGA X. The sequences were then trimmed and analysed using the Neighbour-joining analysis tool for creating phylogeny tree. Bootstrap analysis with 1000 repetitions was used to determine the significance of the branching in a tree. Evolutionary distances were computed by MEGA X using the Tamura-Nei method (Tamura and Nei, 1993).

#### Analysis on the putative enzyme

In order to better understand the variation of the amino acidic sequences of EOMT among *Ocimum* sp., these were evaluated *in silico*. To achieve this, a comparison was made based on the motifs related to the enzymatic activity and substrate specificity in the OMT superfamily according to Ibrahim *et al.* (1998). Following, a search for domains in the enzyme was carried out in UniProt ([www.uniprot.org](http://www.uniprot.org)). Lastly, the protein model was obtained using the homology-modelling server Swiss-Model (<https://swissmodel.expasy.org/>).

#### **2.2.5 Chemical profile analysis**

For obtaining the chemical profile of the reference material, plants grown as described in section 2.1 were analysed. Leaves between positions 3-4 to 7-8 from the apical meristem were sampled and immediately frozen in liquid nitrogen to be either directly processed or stored at -80°C for further analysis. Each biological replicate consisted of three plants. Three biological replicates were taken for volatiles solvent extraction following a modified protocol from Kapteyn *et al.* (2007). Leaves were ground in liquid nitrogen using a mortar and pestle. The powdered leaves were weighed and suspended in ethyl acetate (EtAc) GC grade (Merck GmbH, Darmstadt, Germany) maintaining a 1:5 (w/v) ratio. The samples were incubated overnight at 20°C, shaking (150 rpm) in darkness. Extracts were then filtered using a Chromafil® PET-20/15 MS filter (Macherey-Nagel GmbH & Co. KG, Düren, Germany) and directly analysed by Gas chromatography and flame ionisation detection (7890B GC System, Agilent, Waldbronn, Germany), or stored at -20°C for later analysis. The GC was equipped with a HP-5 nonpolar 30 m x 0.32 mm x 0.25 µm capillary column (Agilent, Waldbronn, Germany). For GC 1 µl of the extract was injected with a syringe using helium as carrier gas at a flow rate of 1.5 ml·min<sup>-1</sup> with a split ratio of 12.5:1. The run was set up with an initial heating at 40°C maintained for 1 min followed by a ramp of rising temperature at a rate of 5°C·min<sup>-1</sup> to 60°C, sustained for one further minute. Following, a second ramp of 3°C·min<sup>-1</sup> until it reached 170°C and was kept for one min. A last third ramp of 30°C·min<sup>-1</sup> raised the temperature up to 270°C, which was then kept for 5 min. Standards (Sigma-Aldrich, Deisenhofen, Germany) of eucalyptol, linalool, estragole, eugenol, methyleugenol and β-caryophyllene were used to estimate the retention time of each compound.

## 2.3 Phenylpropanoid Pathway

### 2.3.1 Phenylpropanoid pathway enzymes and primers design

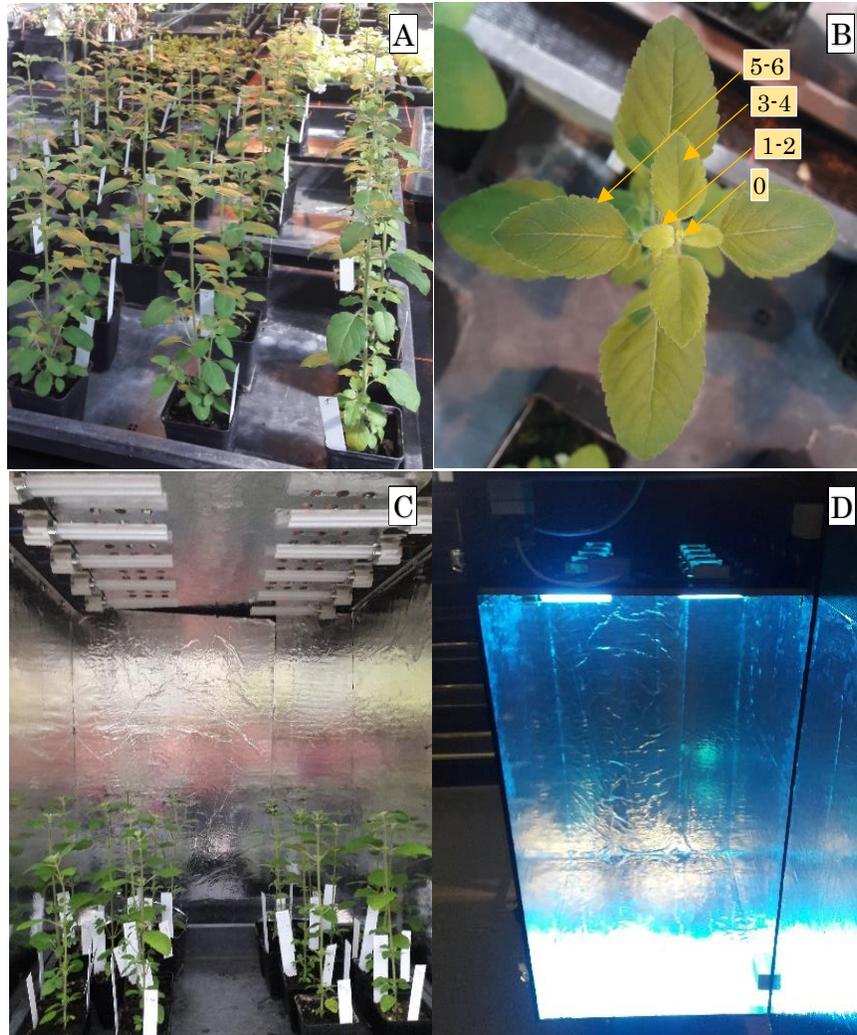
Quantitative PCR (qPCR) was used to investigate the phenylpropanoid metabolic pathway. For the evaluation of gene expression, genes of selected enzymes catalysing reactions in the pathway were chosen: Phenylalanine ammonia lyase (PAL), Cinnamate 4-hydroxylase (C4H), Caffeic acid *O*-methyltransferase (COMT), Caffeoyl-CoA *O*-methyltransferase (CCOMT), Eugenol (and Chavicol) synthase (EGS), and Eugenol *O*-methyltransferase (EOMT) (see **Figure 1.5** for phenylpropanoid pathway details). In addition, Actin was used as a housekeeping gene due to its stable gene expression in UV-B experiments (Davey *et al.*, 2012). qPCR primers for the proteins actin, PAL, C4H and COMT were obtained from bibliography as detailed in **Table 2.3**. To obtain qPCR primers for CCOMT and EGS, the sequence of these enzymes were amplified from cDNA using primers from Rastogi *et al.* (2013). Semi-qPCR reactions were performed as described earlier (section 2.2.3) and Sanger sequencing was done by GATC (change to Eurofins Genomics in the course of the thesis) using the same forward and reverse primers as for the semi-qPCR. The sequence analysis was as described in section 2.2.4. For the design of the EOMT enzyme primers, the sequences obtained (section 2.2) in the plants ID 5751 and 8256 were used as a template. Primers were designed with the help of the programs BioEdit and MEGA X. Melting temperature was checked in the oligo-analysis online tool from Eurofines Genomics (web: <https://eurofinsgenomics.eu/en/ecom/tools/oligo-analysis>). Amplicons length were aimed to be between 75 and 150 bp. All primers were tested in semi-qPCR using cDNA as a template before gene expression experiments and are listed in **Table 2.3**.

**Table 2.3** List of primers used in qPCR experiments in order to obtain sequences for primers design and to quantify gene expression.

Name	Purposed detection	Sequence 5'-3'	Reference
Actinq fw Actinq rev	Expression of Actin as reference gene	TCGTGCTCAGTGGTGGATCA GGGCTGTTATCTCCTTGCTCAT	Rastogi <i>et al.</i> (2013)
PALq fw PALq rev	Expression of PAL	TCCTCCCGGAAAACAGCTG TCCTCCAAATGCCTCAAATCA	Spagnolo <i>et al.</i> (2017)
C4Hq fw C4Hq rev	Expression of C4H	CACCGCAATCTCACCGATTAC AGACGGCGAACACCATGTC	Maurya <i>et al.</i> (2019)
COMTq fw COMTq rev	Expression of COMT	GAGGAACAGGAGCCCACTCA AGTGCGCATCACTCCAGTCA	Maurya <i>et al.</i> (2019)
CCOMT fw CCOMT rev	Obtain sequence of CCOMT	ATGGCAGAAAATGGTGAGCAGCAAA TCAGATGATGCGGCGACACAGG	Rastogi <i>et al.</i> (2013)
CCOMTq fw CCOMTq rev	Expression of CCOMT	GCCAAAAACACCATGGAGATTGGAG CAAAGCAGGGCCTTCTCTGAAGTC	This study This study
EGS fw EGS rev	Obtain sequence EGS	ATGGAGGAAAATGGGATGAAAAGCA TTAAAATGCTGCTGAAGCCGGC	Rastogi <i>et al.</i> (2013)
EGSq fw EGSq rev	Expression of EGS	GGCAATCAAGGTTGCTGGG GCTTCGAATGGCGGCAATG	This study This study
EOMTq fw EOMTq rev	Expression of EOMT	TCCGGTCTATCCCTTCTGCCG ACCGACGGCATCTTTGCATC	This study This study

### 2.3.2 Experiment design and UV-B treatment

In order to better understand the regulation mechanisms of the phenylpropanoid pathway for the synthesis of ME in *O. tenuiflorum*, two chemotypes were used: Krishna (ID 5751) and Rama (ID 8256) (see **Table 2.1** for details). A complete randomized design was established at the Botanical Gardens of KIT (**Figure 2.1a**). Six weeks old plants, grown as described in section 2.1, were used for this test. The experiment consisted of 4 groups of plants per chemotype. Each group had three biological replicates, in turn, each replicate consisted of three individual plants. The control group comprised plants not treated with UV-B radiation (Pre-treatment). The following groups were plants treated with UV-B radiation for a period of 10 min from which leaf samples were taken immediately after radiation (0 h post-treatment), one hour after radiation (1 h post-treatment) and two hours after radiation (2 h post-treatment). The experiment was carried out in September of 2019 and repeated in September of 2020, being the 2 h post-treatment group sampled only in the former date. Data for gene expression comprises only samples taken in September 2019. Data for chemical profile includes samples from 2019 and 2020. For all plant groups, leaves in the 4-5 position counting from the apical meristem (**Figure 2.1b**) were sampled in liquid nitrogen and stored at -80°C for further analysis.



**Figure 2.1.** Details of the UV-B experiment. **a** *Ocimum tenuiflorum* plants grown at the Botanical Gardens in the Karlsruhe Institute of Technology. **b** *O. tenuiflorum* plant 0: apical meristem, 1-2: young leaves, 3-4: expanding leaves, 5-6: fully expanded leaves. **c** *O. tenuiflorum* plants inside UV-B box to be treated. **d** example of UV-B lamps on.

#### UV-B treatment device

The UV-B radiation treatment was performed in a device especially built for this research project (**Figure 2.1 c and d**). A wooden box with a ventilation system was constructed and 10 UV-B lamps (PL-S 9W/01 G23-Philips UV-B/01) were installed on the ceiling of the box. The lamps emit radiation in a narrow wavelength band that ranges between 305-315 nm with a peak at 311 nm, and the radiation intensity was measured at 10 W/m<sup>2</sup>. The box was placed at the Botanical Garden of KIT, in the same glasshouse where the plants were grown to ensure equal temperature and humidity conditions during the UV-B radiation treatment as growing conditions.

### 2.3.3 Gene expression analysis (RNA extraction, cDNA synthesis, qPCR analysis)

RNA extraction and cDNA synthesis was done as described in section 2.2.1. cDNA was diluted 1:10 for further experiments on gene expression. The quantitative real-time PCR (qPCR) reaction protocol used was established by the Molecular Cell Biology from the Botanical Institute at KIT. Briefly, a mastermix for each pair of primers was prepared considering three technical replicates reactions per sample. A 20  $\mu$ l reaction contained 1  $\mu$ l of cDNA (1:10), 4  $\mu$ l GoTaq buffer (Promega GmbH, Mannheim), 11.75  $\mu$ l of nuclease free water, 0.4  $\mu$ l of 10mM dNTPs, 0.4  $\mu$ l of 10  $\mu$ M forward primer, 0.4  $\mu$ l of 10  $\mu$ M reverse primer, 1  $\mu$ l of 50mM MgCl<sub>2</sub> (USB Corporation, Cleveland OH, USA), 0.1  $\mu$ l of GoTaq polymerase (Promega GmbH, Mannheim; Germany) and 0.95  $\mu$ l of SybrGreen (Thermo Fisher Scientific Inc., Waltham MA, USA). The reactions were completed in the CFX96 Touch™ Real-Time PCR Detection System from Bio-Rad Laboratories GmbH (Munich, Germany). The reaction conditions were: an initial denaturation for 3 min at 95°C, followed by 40 cycles of 15 sec at 95°C, 40 sec at 61°C (for all primers except EOMTq at 69.5°C) and a plate-read step; succeeded by a melting curve step of 5 sec 50°C to 95°C at an increment rate of 0.5°C and a plate-read step.

For visualization and evaluation of the qPCR data results, the Bio-Rad CFX Manager version 3.1 (Bio-Rad Laboratories GmbH, Munich; Germany) was used. The relative gene expression (RGE) of the enzymes involved in the phenylpropanoid pathway was calculated by averaging the Ct values of the technical replicates and deducting the housekeeping gene from it, obtaining a delta Ct for each biological replicate in each treatment and controls (**Equation 1**). The relative expression was then calculated with **Equation 2**. Subsequently, the average values of technical and biological replicates for each group were obtained. Standard error for every triple pair of biological replicates was calculated. Significant differences among treatments were determined by two-way ANOVA and t-test post-hoc using the online program SAS® OnDemand for Academics ([https://www.sas.com/de\\_de/software/on-demand-for-academics.html](https://www.sas.com/de_de/software/on-demand-for-academics.html)). Significant differences at  $p$ -value<0.05 were assigned by different letters. For a better visualization on the results, heat maps were constructed using the OriginPro® version 2020b (OriginLab Corporation, Northampton, MA, USA). Finally, to schematize the phenylpropanoid pathway, the program ChemDraw version 20.0 (PerkinElmer Informatics) was used.

$$\Delta Ct = Ct \text{ gene of interest} - Ct \text{ housekeeping gene} \quad (\text{Eq. 1})$$

$$\text{RGE} = 2^{-\Delta Ct} \quad (\text{Eq. 2})$$

### 2.3.4 Chemical profile analysis

#### Volatile compounds

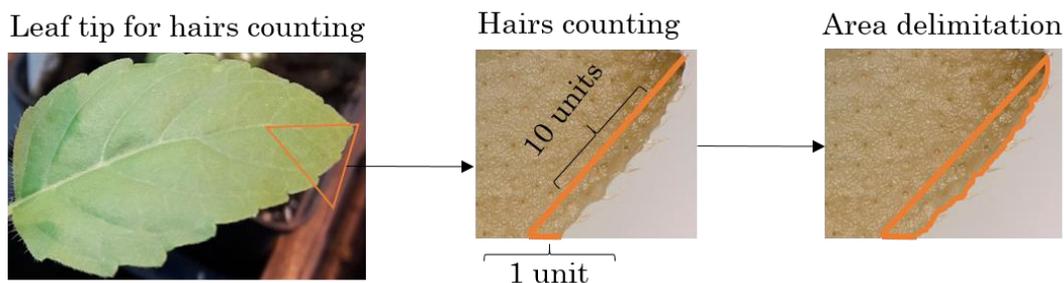
Solvent extraction was done as described in section 2.2.6 with some modifications. Leaves in position 5-6 from the apical meristem were taken as samples. These were ground in liquid nitrogen using a mortar and pestle. The powder was weighed and suspended in EtAc containing 1 mg/ml 1,2,4 trichloro benzene (Sigma-Aldrich, Deisenhofen, Germany) as internal standard, in a ratio 1:5 (w/v). Gas chromatography conditions are described in section 2.2.6. For each biological replicate the analyte/internal standard ratio was calculated together with average and standard error for each treatment. Significant differences among groups were determined by two-way ANOVA and t-test post-hoc using the online program SAS® OnDemand for Academics ([https://www.sas.com/de\\_de/software/on-demand-for-academics.html](https://www.sas.com/de_de/software/on-demand-for-academics.html)). Significance at  $p$ -value<0.05 was specified by different letters.

#### Anthocyanin content

The protocol used was modified from Van Tuinen *et al.* (1999). Samples were ground in liquid nitrogen with a mortar and pestle, weighed and suspended in 4 ml of 1% HCl/methanol solution. The extracts were incubated in darkness overnight at 20°C shaking (150 rpm). After 18 hours, the extracts were centrifuged (Heraeus Pico 17, Thermo scientific, Waltham MA, USA) at 5000g for 5 min and the supernatant was used for analysis. Anthocyanin content was measured in a Jasco V750 spectrophotometer (Jasco Corporation, Tokyo, Japan), at 530 nm using 1% HCl/methanol solution as blank. For each group, biological replicates averages and standard error were calculated. Significant differences among treatments were determined by two-way ANOVA and t-test post-hoc using the online program SAS® OnDemand for Academics ([https://www.sas.com/de\\_de/software/on-demand-for-academics.html](https://www.sas.com/de_de/software/on-demand-for-academics.html)). Significance at  $p$ -value<0.05 was specified by different letters.

### 2.3.5 Trichomes analysis

Hairy trichome analysis was done in order to get an insight into the physiological differences between the two *O. tenuiflorum* chemotypes, Rama and Krishna. Fully expanded leaves from positions 7-8 counting from the apical meristem were harvested in 80% ethanol solution and incubated for 10 weeks at 4°C in order to extract the chlorophyll. Leaves were observed in a VHX-950F digital microscope (Keyence, Osaka, Japan). The images obtained by microscopy were analysed in ImageJ (Schneider *et al.*, 2012). As shown in **Figure 2.2**, a line was drawn from the tip of the leaf to the base of the picture (10-units length) and from the base to the border (1-unit length) in order to standardize the area in each leaf. The trichome number was counted inside the limited zone. After counting, a border line was drawn to obtain a final area. The trichomes per area ratio was finally calculated.



**Figure 2.2.** Schematic representation of section selected for trichrome counting. From left to right it shows first the portion of the leaf where photos were taken followed by the selection delimited for hair counting, and finally the area calculation to obtain a n° hairs/area ratio.

Glandular trichomes (or peltate glands) visualization was aimed in order to determine whether these type of glands are present or not in *O. tenuiflorum* leaves. Peltate glands in *O. basilicum* leaves have been established as the structure where the enzymatic activity of EOMT is higher when compared to the whole leaf (Gang, Simon, *et al.*, 2002). Therefore, to be certain of the presence of these glands in our *O. tenuiflorum* accessions, a nail polish imprinting technique was employed to visualize them. For doing so, translucent nail polished was applied to the adaxial side of leaves. When dried, the nail polished layer was removed and observed under 20x objective in light microscopy using a Zeiss-Axioskop 2 FS with DIC illumination (Zeiss, Jena, Germany).

## 2.4 Eugenol *O*-methyltransferase Subcellular localization

### 2.4.1 Eugenol *O*-methyltransferase full sequence

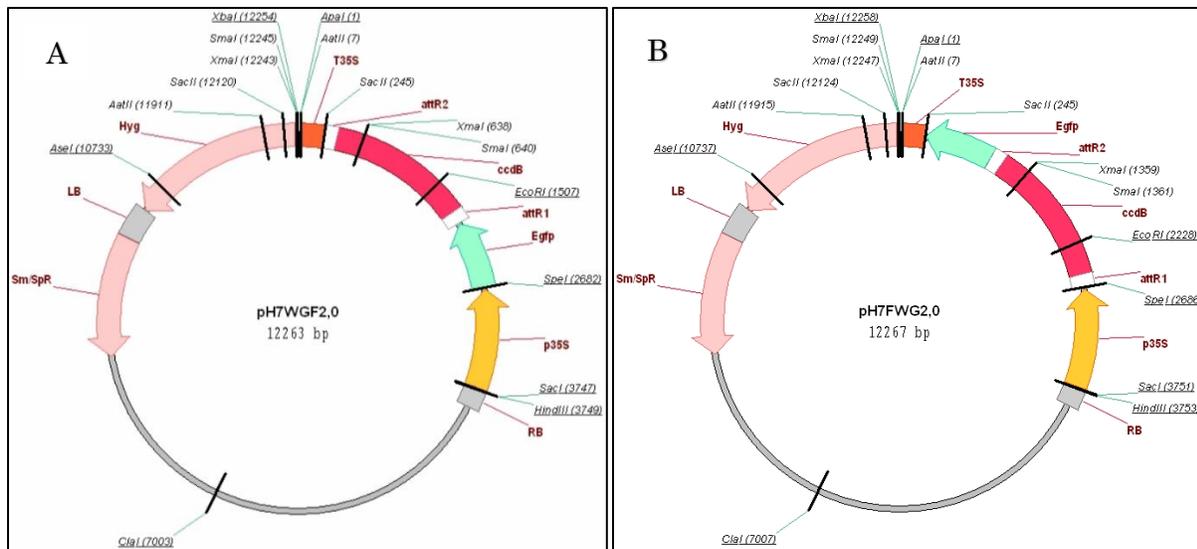
The full sequence of the putative EOMT was amplified with the following primers EOMT*ob* fw 5'-3': ATGGCATTGCAAAAAGTAG and EOMT*ob* rev 5'-3': TTAAGGATAAGCCTCTATGAG. These primers were designed from the full sequence of the recombinant enzyme (Genbank accession number: AF435008.1) obtained by Gang *et al.* (2002) from *Ocimum basilicum*. The amplification of the enzyme was done using cDNA as template from reference plant ID 5751 (see **Table 2.1** for details) and a Q5® *High-Fidelity* DNA polymerase as described in section 2.2.2. The PCR amplicon was checked by gel electrophoresis and the PCR product was purified. The product was cloned into the pGEM-T Easy Vector System (Promega GmbH, Mannheim) as explained in section 2.2.3. After plasmid purification, an aliquot was sequenced using M13 primers. Sanger sequencing was done by Eurofins Genomics (Cologne, Germany). Once the sequence was analysed as described in section 2.2.4, and the putative EOMT was confirmed to be inserted in the vector, this was used in further analysis.

### 2.4.2 Localization peptides search

In addition, localization peptides were searched in the sequence of the enzyme (sequences of gDNA -exons and introns- and cDNA) using the search tools of the websites SignalP 5.0 (<http://www.cbs.dtu.dk/services/SignalP/> -last visited June 2021-). Further, nuclear export sequences (NESs) were searched in the data base provided by the Technical University of Denmark (<http://www.cbs.dtu.dk/services/NetNES/> -last visited July 2021-) and the data base provided by Rostlab from the Technical University of Munich (<https://rostlab.org/services/nlsdb/> -last visited July 2021-).

### 2.4.3 Plasmid construction for *Agrobacterium* transformation

The sequence of the putative EOMT obtained above was cloned by Gateway® Technology using the Clonase™ II system (Invitrogen, Thermo Fisher Scientific, Germany) following the manufacturer's instructions. For doing so, the sequence of the EOMT was amplified with modified EOMT<sub>ob</sub> primers with *attB* sites ends added. The *attB*-PCR product was then used in a BP recombination reaction to be cloned into a donor vector (pDONR™). The reaction was used for *E. coli* (strain DH5- $\alpha$  competent cells) transformation (see section 2.2.3) using Zeocin™ in LB media agar plates as selective antibiotic. The extracted plasmids were sequenced (Sanger sequence by Eurofins Genomics) using M13 primers, in order to verify that the target sequence was inserted in the right direction. The verified entry clone from the BP reaction was then used in different LR recombination reactions having two destination Gateway vectors as systems for stable transformation (**Figure 2.3**). Both vectors contain the sequence for green fluorescence protein (GFP) which can be N-terminally fused (vector pH7WGF2) and C-terminally fused (vector pH7FWG2) to the target protein (Karimi *et al.*, 2002). The entry clone and destination vectors were linearized using the restriction enzyme EcoRI-HF® (New England BioLabs, Frankfurt) to facilitate the reaction. The LR reaction was used for *E. coli* strain DH5- $\alpha$  competent cells transformation (see section 2.2.3) using spectinomycin in LB media agar plates as selective antibiotic. Colonies were selected and individually grown to increase the plasmids copies. After plasmid extraction, the interface target-gene-GFP was verified in both destination vectors by sequencing using the following primers: pEGFP\_N fw 5'-3': CCGTCCAGCTCGACCAG (GATC sequencing primers by Eurofins Genomics) and pEGFP\_N rev 5'-3': CTGGTCGAGCTGGACGG (reverse complement of pEGFP\_N fw). The resulting Gateway vectors having the EOMT sequence were renamed as follow: pH7WGF2 with inserted EOMT: EOMT-*W* and pH7FWG2 with inserted EOMT: EOMT-*F*.



**Figure 2.3.** Schematic view of Gateway vectors used for *Agrobacterium*-mediated plant cell transformation containing a cauliflower mosaic virus (CaMV) 35S promoter (p35S) and green fluorescence protein (GFP) for subcellular localization (Karimi *et al.*, 2002). **a** Vector pH7WGF2.0 used for N-terminally fused GFP (image obtained from: <https://gatewayvectors.vib.be/index.php/collection/ph7wgf2>). **b** Vector pH7FWG2.0 used for C-terminally fused GFP (image obtained from: <https://gatewayvectors.vib.be/collection/ph7fwg2>).

#### 2.4.4 *Agrobacterium* transformation

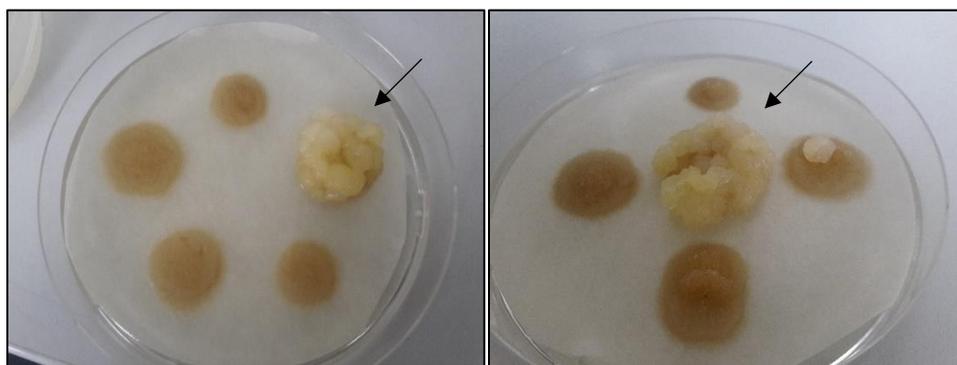
A freeze-thaw transformation protocol for *A. tumefaciens* established in the laboratory of Molecular Cell Biology from the Botanical Institute at KIT was used. The transformation was done separately for both vectors obtained in the previous section: EOMT-*W* and EOMT-*F*. Briefly, about 500 ng of the vectors were gently mixed with bacterial cells and immediately freeze in liquid nitrogen for 5 min. The bacteria were then thawed at 37°C in a water bath for 10 min. Paul media was added to the bacteria and these were incubated at 28°C shaking (280 rpm) for 4 hours. Cells were later centrifuged and most supernatant discarded. The remaining supernatant containing the cells were plated in Paul's agar plates with added antibiotics for colony selection (rifampicin, streptomycin and spectinomycin). After three days, colonies were selected and grown for 4 hours in Paul media. 1 µl of the bacteria suspension was used for a colony PCR as described in section 2.2.3. The resulting cultivated colonies, containing either the vector EOMT-*W* or the vector EOMT-*F*, were used for plant cell transformation.

#### 2.4.5 BY2 cells transformation and EOMT localization

*Nicotiana tabacum* cv. 'Bright Yellow', clone 2 (BY-2) cells (Nagata *et al.*, 1992) were transformed with *A. tumefaciens* following a transformation protocol established at

Molecular Cell Biology from the Botanical Institute at KIT. Wild type BY-2 cells were cultivated in liquid MS media at 26°C. Four days old cultures were used for the transformation. First, 60 ml of BY-2 cells were washed in a Najel device filter with 700 ml Paul's media under sterile conditions avoiding drying of the cells. Then, 6 ml of cells were added to aliquots of 180 µl of transformed *A. tumefaciens* (see section above) and mixed in an orbital shaker for 5 min at 100 rpm. The BY-2 transformed cells were later spotted in MS agar plates containing a filter paper (**Figure 2.4**). The media was supplemented with 60 mg/ml hygromycin to select transformed cells and 300 mg/ml cefotaxime to eradicate *A. tumefaciens*. Antibiotics were added in a ratio of 60 µl antibiotic per 30 ml media. Cells were incubated for a period of three weeks in darkness at 26°C before selecting transformed calluses (**Figure 2.4**). After this time, calluses were delicately sliced and transferred into MS liquid media supplemented with half of the antibiotics previously used.

The localization of EOMT-labelled-GFP enzyme was observed by Spinning Disc Microscopy (Zeiss, Jena, Germany) under a 63x oil immersion objective. The microscope was provided with a CSU-X1 confocal scanner unit (Yokogawa Electric Corp.) and a digital CCD camera (Zeiss AxioCam MRm). BY-2 cells were firstly observed under differential interference contrast (DIC), and later the GFP signal from transformed cells was detected by excitation with blue light using an ArKr laser (Zeiss, Jena, Germany). The images obtained were visualised and analysed using the Zeiss ZEN System Imaging software.



**Figure 2.4.** Examples of three-week plates where BY-2 transformed cells were cultivated, with one of the spots resulting in a viable callus (arrows) that were later transferred into liquid media, and four unsuccessfully transformed cell lines of BY-2.

#### 2.4.6 Protein analysis

Cell cultures from both BY-2 transformed cell lines (EOMT-*W* and EOMT-*F*) were harvested on the third day after sub-cultivation. The protein was extracted from each cell line, following the method established in the laboratory of Molecular Cell Biology from the Botanical Institute at KIT. Homogenised cells by mortar and pestle using

liquid nitrogen, were resuspended in a 1:1 ratio (w/v) with extraction buffer (25 mM MES, 5 mM EGTA, 5 mM MgCl<sub>2</sub>·6H<sub>2</sub>O, pH 6.9, supplemented with 1 mM DTT and 1 mM PMSF) and vortexed vigorously. Lysates were centrifuged at 13500 rpm, for 15 min at 4 °C. The supernatant containing the protein extract was mixed with 3X loading buffer (containing 50 mM Tris/HCl, pH 6.8), 30 % glycerol (v/v), 300 mM DTT, 6 % SDS, and 0.01 % bromophenol blue. The mixtures were incubated at 98 °C for a period of 8 min, then centrifuge at 12000 rcf for 5 min. Samples were separated by SDS-PAGE gel 10 % (w/v).

Subsequently, a western blot analysis was done. The separated proteins were transferred from the SDS-PAGE gel to a PVDF membrane in a Trans-Blot® SD Semi-Dry Transfer Cell (Bio-Rad Laboratories, Inc., Kent, England), that run for 1 h at 20 volts. For the antibody incubation, the membrane was then washed with TBS buffer (20 mM Tris/HCl, 300 mM NaCl, pH 7.4-), and incubated 1h in blocking buffer (TBS buffer supplemented with 5 % (w/v) milk powder (Sigma-Aldrich, Darmstadt)). Following subsequent washing with TBST buffer (TBS buffer with 1 % Tween) and TBS buffer, the primary antibody (anti-GFP antibody, IgG mouse (Sigma-Aldrich, Darmstadt)) was provided in a 1:1000 dilution and incubated overnight at 4°C. After that, the membrane was washed two times in TBST buffer and one time in TBS buffer. A dilution of 1:30000 of the secondary antibody (anti-mouse IgG, goat (Sigma-Aldrich, Darmstadt)) was then added and incubated 1 h at room temperature. The PVDF membrane was then washed four times with TBST buffer, and stained by 15 min incubation with staining buffer (100 mM Tris-HCl, 100 mM NaCl, 500 mM MgCl<sub>2</sub>·H<sub>2</sub>O, pH 9.7). Finally, BCIP/NBT detection reagent (Thermo Fisher Scientific, Germany) was used to visualize the band pattern.

## 2.5 DNA barcoding

The following section was modified from **Ríos-Rodríguez *et al.* 2021**.

### 2.5.1 Plant material

Reference plant material is described in **Table 2.1** and commercial samples are described in **Table 2.4**. Commercial samples were obtained from local stores in the city of Karlsruhe, Germany.

### 2.5.2 DNA extraction

For DNA extraction from reference plants, leaves were collected at the Botanical Garden of KIT in liquid nitrogen and processed as described in section **2.2**. Commercial samples that had declared to contain only Holy Basil as herb, the DNA extraction was completed as previously described. For commercial samples that

declared a mixture of herbs, a selection of the plant material was made. In these, non-leaf like material was discarded (**Figure 2.5**) to avoid interference in the subsequent reactions. Successively, DNA was extracted as explained in section **2.2**.

**Table 2.4.** Commercial samples used in this study to test DNA barcodes; ID assigned; type of commercial product description; declared species.

ID	Type of sample	Declared
TulComm.0001	Tea. Has different types of Tulsi, unmixed with other herbs.	<i>Ocimum</i> sp. unspecified.
TulComm.0002	Tea. Has different types of Tulsi, unmixed with other herbs.	<i>Ocimum</i> sp. unspecified.
TulComm.0003	Tea. Tulsi unmixed with other herbs.	<i>Ocimum</i> sp. unspecified.
TulComm.0004	Tea. Tulsi mixed with different herbs.	<i>Ocimum</i> sp. unspecified.
TulComm.0005	Tea. Tulsi mixed with different herbs.	<i>Ocimum</i> sp. unspecified.
TulComm.0006	Tea. Tulsi mixed with different herbs.	<i>Ocimum</i> sp. unspecified.
TulComm.0007	Tea. Tulsi mixed with different herbs.	<i>Ocimum</i> sp. unspecified.
BasComm.0001	Spice. Dried basil unmixed with other herbs.	<i>Ocimum</i> sp. unspecified.
BasComm.0002	Spice. Dried basil unmixed with other herbs.	<i>Ocimum</i> sp. unspecified.
MenComm.0001	Tea. Mint unmixed with other herbs	<i>Mentha</i> sp. unspecified.



**Figure 2.5.** Example of Holy Basil commercial sample mixed with other herbs, where the plant material was preselected for DNA extraction. In the circled area is shown the non-leaf like material that is discarded.

### 2.5.3 DNA-Barcodes and PCR conditions

#### Trait-independent marker assay

As a trait-independent marker, the plastidic *psbA-trnH intergenic spacer* was selected. This was amplified using the primers described in **Table 2.6**. A standard semi-qPCR reaction was set-up and the thermocycling parameters are described in Jürges *et al.* (2018). The resulting PCR product was subsequently used in a Restriction Fragment Length Polymorphism (RFLP) assay using the restriction enzyme Hinf I (New England BioLabs, Frankfurt) (Jürges *et al.*, 2018). For *psbA-trnH* details of the accessions see appendix **Table A.1**.

### Trait-related marker assay

The gene encoding EOMT was used as trait-related marker. The putative sequence of EOMT was obtained from gDNA as described in section 2.2.2 using the primers EOMTfw<sub>R</sub> and EOMTrev<sub>R</sub> (see Table 2.2). The sequences were obtained and analysed as explained in section 2.2.4. The EOMT sequences from Krishna Tulsi (ID 5751, Table 2.1) and Sweet Basil (ID 9056, Table 2.1) were employed as reference to design diagnostic primers to amplify a fragment of EOMT: EOMTfw and EOMTrev (see Table 2.6). These primers were used together with a third reverse primer, EOMTrev<sub>R</sub> (see Table 2.2), positioned outside of the fragment amplified in a multiplex PCR that, in case of *O. tenuiflorum*, produced an additional diagnostic band to the small amplicon yielded by EOMTfw and EOMTrev. The PCR reaction for these EOMT based primers consisted of an initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 61°C for 30 sec, and elongation at 72°C for 30 sec. The reaction completed with a final elongation at 72°C for 10 min, using a conventional Taq polymerase without a proofreading function (New England Biolabs, Frankfurt). All reactions were supplemented with 10 mg/ml BSA and 5 M betaine. The results were visualised by electrophoresis as explained in section 2.2.

### Combined trait-related and -independent marker assay

In addition, the trait-independent marker (*psbA-trnH*) and the trait-related marker (EOMT) were used in a single multiplex PCR assay. To improve the reaction conditions, the EOMT primers were modified to reach the *psbA-trnH* markers annealing temperature. These modified primers (EOMTfw' and EOMTrev') are described in Table 2.6. The multiplex PCR setting was the same as the one employed for the *psbA-trnH* reaction previously defined.

**Table 2.6.** Primer for DNA barcoding.

Name	Sequence 5'-3'	Reference
<i>psbA</i>	GTTATGCATGAACGTAATGCTC	Sang <i>et al.</i> (1997)
<i>trnH</i>	CGCGCATGGTGGATTACAAATCC	Tate and Simpson (2003)
EOMT fw	TCCGGTCTATCCCTTCTGCCG	This study
EOMT rv	ACCGACGGCATCTTTGCATC	This study
EOMTfw'	TCCGGTCTATCCCTTCTGCC	This study
EOMTrev'	CCGACGGCATCTTTGCATC	This study

### 2.5.4 Chemical profile analysis

In order to evaluate the volatiles in the commercial samples, leaves extracts in ethyl acetate were analysed using gas chromatography as described in section 2.2.6. 1ul of the extract was used for analysis and the retention time was compared to eugenol and methyleugenol standards. The percentage of each compound was calculated in relation to the total area.

## 3 Results

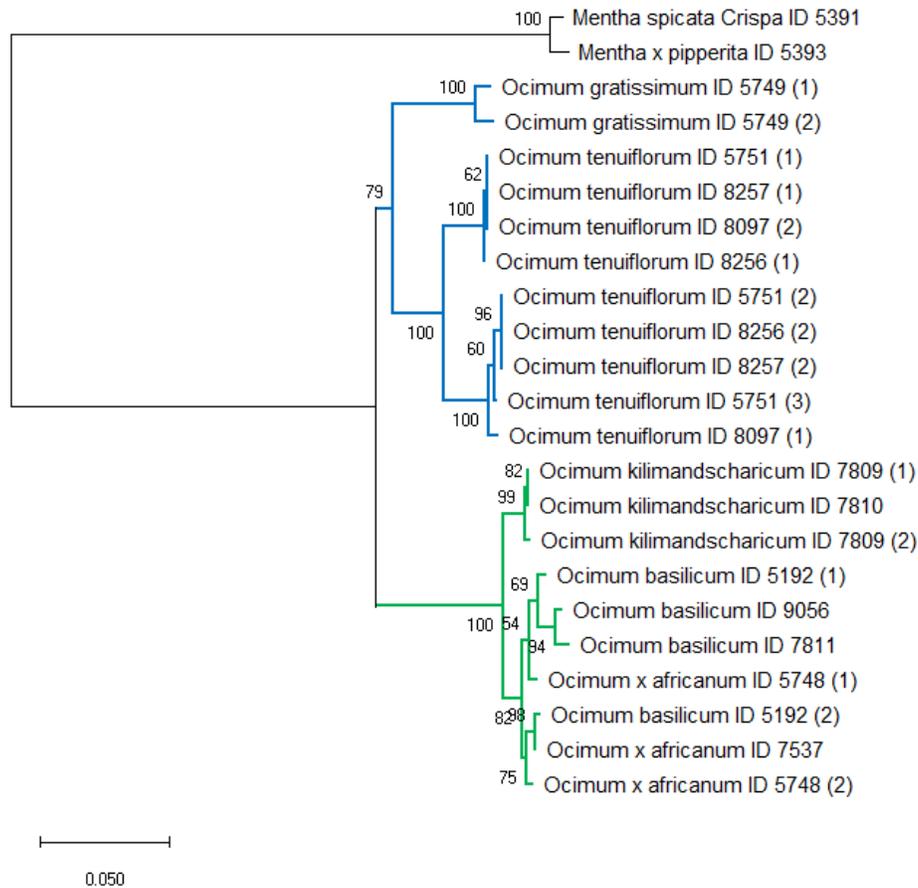
### 3.1 Eugenol *O*-methyltransferase in the genus *Ocimum*

The sequence of the enzyme eugenol *O*-methyltransferase (EOMT) has been analysed and three main results have been found. First, EOMT phylogeny discloses two distinctive groups in *Ocimum* sp., one with species closely related to *O. basilicum* and another with species closely related to *O. tenuiflorum*. Second, the amino acid sequence of EOMT presents residues and motifs potentially relevant for the synthesis of methyleugenol in the genus *Ocimum*. This is supported by the third results, in which only the *O. tenuiflorum* accessions present methyleugenol in their chemical profile. Following, these three results will be revised in detail.

#### 3.1.1 EOMT phylogeny discloses two distinctive groups in *Ocimum* sp.

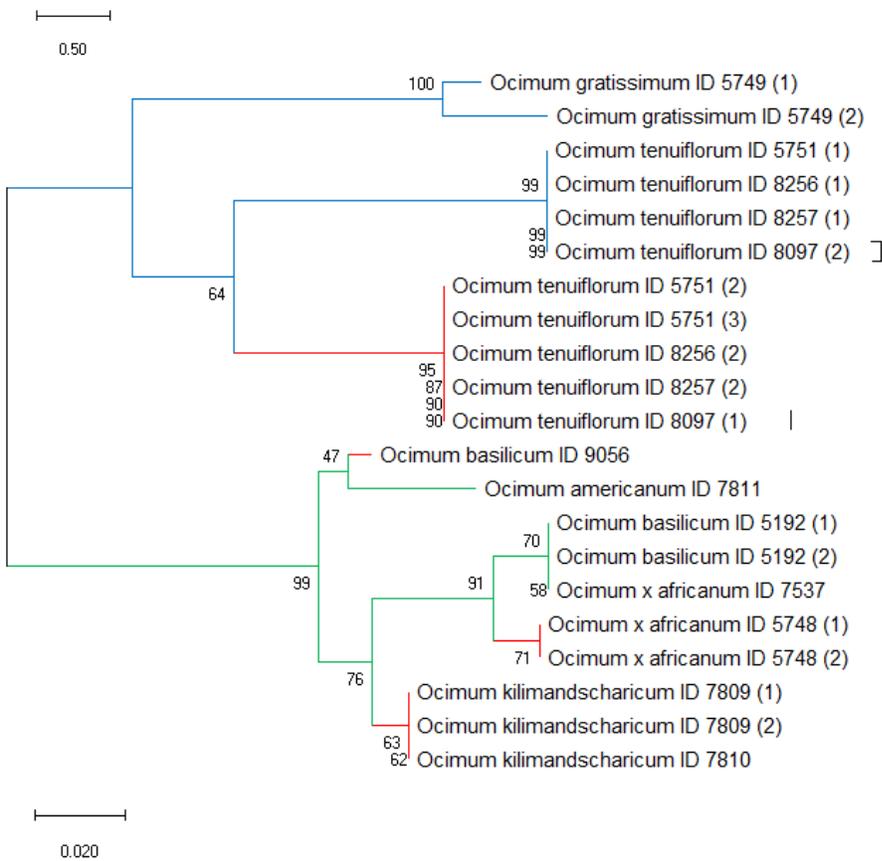
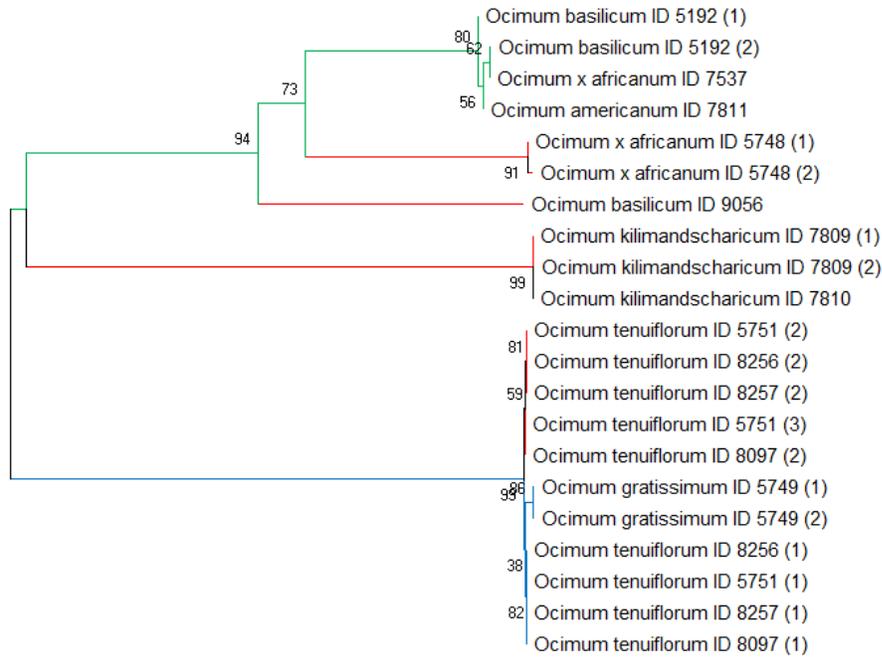
In order to understand the genetic variation of the EOMT enzyme in the genus *Ocimum*, 21 nucleotide sequences were obtained from genomic DNA from six *Ocimum* species: four *O. tenuiflorum*, one *O. gratissimum*, two *O. kilimandscharicum*, two *O. basilicum*, one, *O. americanum* and two *O. x africanum*. The number of obtained EOMT sequences surpasses the number of individuals (12) from which the DNA was obtained because of the presence of isoforms of the gene. Moreover, from the accession number 5751 (*O. tenuiflorum*) three different sequences for EOMT were found. The sequences have been deposited in the database of the Botanical Garden of the KIT (nicklab.de).

The EOMT nucleotide sequences were then analysed using the Neighbour-joining method and two distinctive groups were identified among *Ocimum* species (**Figure 3.1**). In the first group, the EOMT sequences from *O. tenuiflorum* and *O. gratissimum* were clustered together, this was called the Tulsi clade. In the second group, the EOMT sequences from *O. kilimandscharicum*, *O. basilicum*, *O. americanum* and *O. x africanum* were clustered together, this was called the Basil clade.



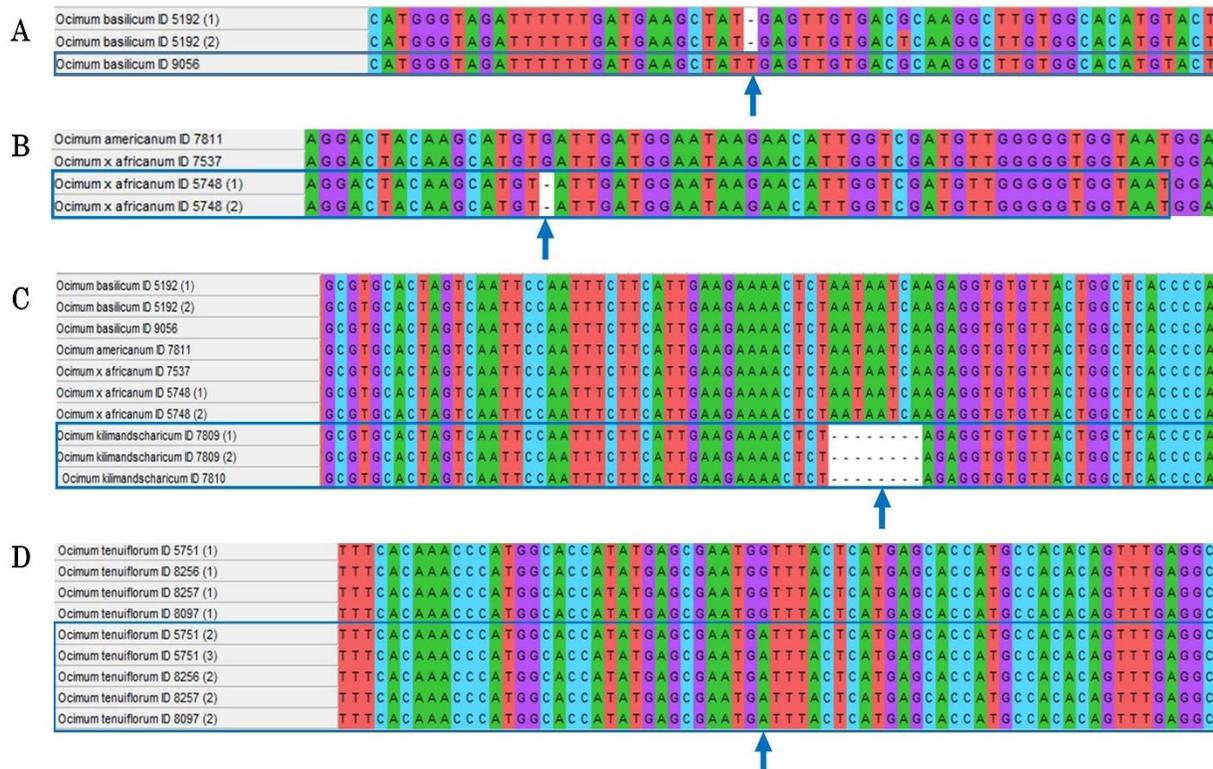
**Figure 3.1.** Phylogeny of *Ocimum* species using the nucleotide sequence (exons and introns) of the enzyme eugenol *O*-methyltransferase. Two *O*-methyltransferase sequences from *Mentha* sp. were used as outgroup. Blue branches indicate the Tulsi clade (*O. tenuiflorum* and *O. gratissimum*) and green branches indicate the Basil clade (*O. basilicum*, *O. americanum*, *O. africanum* and *O. kilimandscharicum*). Neighbour-joining distant tree, bootstrap values (based on 1000 replicates) next to the branches. For details on the accessions refer to **Table 2.1**.

When taking in consideration the individuals within the Basil clade (**Figure 3.1**), *O. basilicum* (ID 9056) is grouped together with *O. americanum* (ID 7811). Contrary to this, in the exons analysis (**Figure 3.2**), the two species are not grouped together inside the clade, but they are again clustered together in the intron analysis also shown in **Figure 3.2**. Similarly, when taking in consideration the individuals within the Tulsi clade (**Figure 3.1**), *O. gratissimum* is grouped separately from the rest of the *O. tenuiflorum* inside the clade. However, in **Figure 3.2 (top)**, *O. gratissimum* is located within the sequences of *O. tenuiflorum*. Similar results were found in the phylogenetic analysis using the sequence of exons and cDNA (see appendix, **Figure A.1**). These examples on the clustering differences suggests the intron sequences influence both clades arrangement.



**Figure 3.2** Phylogeny of *Ocimum* species using the **exons** nucleotide sequence (**upper**) from the eugenol *O*-methyltransferase enzyme. Blue branches indicate the Tulsi clade and green branches indicate the Basil clade. Red branches indicate the sequences of pseudogenes. Phylogeny of *Ocimum* species using the **intron** nucleotide sequence (**lower**) from the eugenol *O*-methyltransferase enzyme. Blue branches indicate the Tulsi clade of (symbol  $\square$ ) indicates pseudogene in the branch). Green branches indicate the Basil clade. Red branches indicate the sequences of pseudogenes (symbol  $\square$  indicates non pseudogene in the branch). Neighbour-joining distant tree, bootstrap values (based on 1000 replicates) next to the branches. For details on the accessions refer to **Table 2.1**.

Subsequently, the EOMT exon sequences were translated into amino acidic sequences, and pseudogenes of the EOMT were found in the two clades mentioned above (see **Figure 3.2**). In *O. basilicum* ID 9056 the EOMT sequence obtained was a pseudogene in which a base insertion shifted the reading frame creating later a stop codon (**Figure 3.3 a**). Both sequences found in *O. x africanum* ID 5748 were pseudogenes (**Figure 3.3 b**), where a deletion of one base shifted the reading frame creating a later stop codon. All off the sequences found in *O. kilimandscharicum* ID 7809 and ID 7810 were pseudogenes; here an eight base deletion shifted the reading frame resulting in a stop codon (**Figure 3.3 c**). In the *O. tenuiflorum* clade, one sequence from each *O. tenuiflorum* ID 5751, ID 8256, ID 8257 and ID 8097 were pseudogenes. In all of them, the pseudogene was caused by a point mutation where a base transition was observed: guanine was substituted adenine, resulting in a nonsense mutation (stop codon) (**Figure 3.3 d**). The remained sequences of the functional protein obtained from gDNA were analysed together with cDNA.



**Figure 3.3** Mechanisms for pseudogene formation in the gene of the putative eugenol *O*-methyltransferase (EOMT) in *Ocimum* species. **a** Pseudogene of EOMT in *O. basilicum* ID 9056 (framed in rectangle shape) by one base insertion (arrow). **b** Pseudogenes of EOMT in *O. x africanum* ID 5748 (framed in rectangle shape) by deletion of one base (arrow). **c** Pseudogenes of EOMT in *O. kilimandscharicum* IDs 7809 and 7810 (framed in rectangle shape) by deletion of eight bases (arrow). **d** Pseudogenes of EOMT in *O. tenuiflorum* IDs 5751, 8256, 8257, and 8097 (framed in rectangle shape) by nonsense mutation (G→A point mutation) creating a TGA stop codon (arrow).

The nucleotide and amino acidic sequences of the putative EOMT obtained from gDNA were compared and analysed together with the enzyme sequences obtained from cDNA. As a result, several EOMT gene variants were found (**Table 3.1**). When comparing the two *Ocimum* clades mentioned above, in the Basil clade few gene sequences were found in the cDNA, and in some accessions such as *O. basilicum* ID 9056, no functional protein sequences were found at all. Contrary to this, in the accessions belonging to the Tulsi clade, only in *O. gratissimum*, the enzyme sequence was not found in cDNA. And all of the accessions from this clade had EOMT gene variants that were functional. In addition, the maximum number of the gene variants for EOMT (five) was found in *O. tenuiflorum* ID 5751. Lastly, gene variants found in the cDNA of *O. basilicum* ID 9052 and *O. tenuiflorum* ID 5751 were not found in the gDNA.

**Table 3.1.** Summary of the number of gene variants of the putative eugenol *O*-methyltransferase (EOMT) enzyme found in *Ocimum* sp. The gene sequences were obtained from genomic DNA (gDNA) and complementary DNA (cDNA). The functional gene sequences and pseudogenes of the EOMT were obtained by analysing the amino acid sequence using the translation tool of MEGA X. This analysis summary excludes intron differences.

<i>Ocimum</i> sp. (ID)	gDNA EOMT gene variants (pseudogenes)	cDNA EOMT gene variants	Total EOMT gene variants (gDNA +cDNA)	Total functional EOMT gene variants
<i>O. basilicum</i> (5192)	2 (-)	2	3	3
<i>O. basilicum</i> (9056)	1 (1)	-	1	-
<i>O. americanum</i> (7811)	1 (-)	-	1	1
<i>O. x africanum</i> (5748)	2 (2)	1	3	1
<i>O. x africanum</i> (7537)	1 (-)	-	1	1
<i>O. kilimandsch.</i> (7809)	2 (2)	-	2	-
<i>O. kilimandsch.</i> (7810)	1 (1)	-	1	-
<i>O. gratissimum</i> (5749)	2 (-)	-	2	2
<i>O. tenuiflorum</i> (5751)	3 (2)	3	5	3
<i>O. tenuiflorum</i> (8256)	2 (1)	1	2	1
<i>O. tenuiflorum</i> (8257)	2 (1)	1	3	2
<i>O. tenuiflorum</i> (8097)	2 (1)	1	2	1

As mentioned in **Table 3.1**, the intron sequences were not included in the gene variations to determine differences in the functional enzyme. However, the introns were analysed as already shown in the phylogeny analysis (**Figure 3.2**). Moreover, other differences and patterns were found among the intron sequences (see appendix, **Figure A.2**). Besides variation in the nucleotide sequences of the intron, there were

differences in the length. The largest intron was from *O. gratissimum* (98 bp), whereas the shortest intron was found in *O. tenuiflorum* pseudo-sequences mostly (86 bp). Among the similarities, besides the common base pairs within the sequences in *Ocimum* sp., the position of the intron was quite conserve in all the sequences obtained, at around the base pair 760 in the EOMT (~1,100 bp).

### 3.1.2 Protein analysis reveals *aa* residues and motifs in EOMT which are potentially relevant for the synthesis of methyleugenol in the genus *Ocimum*

In order to better understand the genetic variation of the EOMT enzyme, the sections of the amino acid sequence thought to play an important role in the synthesis of methyleugenol were chosen for this analysis. In this frame, five motifs found in the *O*-methyltransferase superfamily, shown in **Table 3.2**, are suggested to be part of the active site of the enzyme (Gang, Lavid, *et al.*, 2002; Ibrahim *et al.*, 1998). Similar but not identical motifs were also found in the EOMT from the *Ocimum* sp. analysed in this study (**Table 3.2**). This includes the boxes I, II, III, suggested to be important for SAM binding sites, and box V suggested to be important for catalysis (Gang, Lavid, *et al.*, 2002).

Moreover, and to support the previous idea, between motifs III and IV is located the amino acid residue Serine –position 261 in EOMT *aa* sequence – (see appendix, **Table A.2**), which is a mutagenic site that when altered to a Phenylalanine instead, the enzyme changes from a eugenol *O*-methyltransferase to a chavicol *O*-methyltransferase (Gang, Lavid, *et al.*, 2002).

Taking in consideration all of the *aa* residues between and including box I and V, all of the sequences analysed presented a total of 126 *aa* residues. From those 113 *aa* residues are identical for all the sequences from all *Ocimum* species. This highlights that variation of only 13 *aa* residues could make a difference on the chemical profile (see appendix, **Table A.2**).

**Table 3.2.** Motifs found in different sequences of eugenol *O*-methyltransferase from *Ocimum* species. **1** Box number for the motifs. **2** Motifs in *O*-methyltransferase superfamily suggested by Ibrahim *et al.* (1998). **3** Motifs in putative Eugenol *O*-Methyltransferase in *Ocimum* sp. **4** Amino acid residue position in *O. tenuiflorum*. Blue coloured amino acid residues indicate a match between suggested motifs for OMT superfamily and EOMT from *Ocimum* sp.

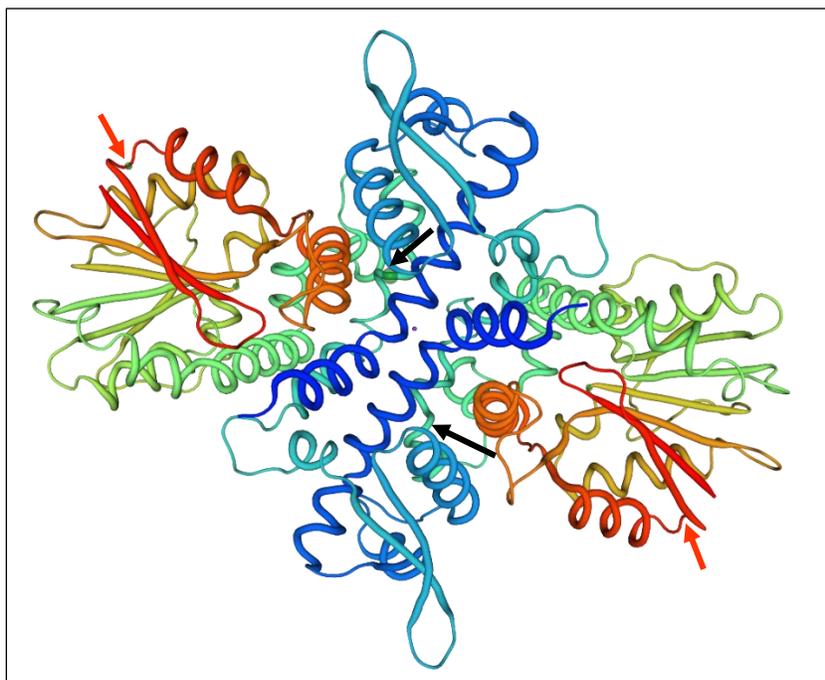
1	I	II	III	IV	V
2	L V D V G G G x G	G I N F D L P H V	E H V G G D M F	N G K V I	G G K E R T
3	L V* D V G G G x G	C T V x D L P H V	S Y I G G D M F	G G K V I	N A K E R T
4	199 200 201 202 203 204 205 206 207	222 223 224 225 226 227 228 229 230	241 242 243 244 245 246 247 248	285 286 287 288 289	320 321 322 323 324 325

\* One EOMT sequence out of 20 had an Alanine instead of Valine in motif I, position 200. Details are shown on **Table A.2** in appendix.

Interestingly, *O. gratissimum* EOMT sequences have more than a few amino acids residues, in the frame of these five motifs, that are identical to the amino acids residues found the EOMT sequences from *O. basilicum* related species (details in appendix, **Table A.2**). For better understanding of these similarities, the chemical profile has been obtained.

Further analysis in UniProt (last visited: 21.06.2021), using the whole sequence of the EOMT from *O. tenuiflorum* (Krishna Tulsi 5751), suggested that the 357 aa residues enzyme possess two important domains, a dimerization domain (comprising residues 33-81) and an *O*-methyltransferase domain (comprising residues 132-338) (see appendix, **Figure A.3**). Following, and considering the five motifs described in **Table 3.2**, these motifs are located within the *O*-methyltransferase domain.

Subsequently, a model of the enzyme was obtained using Swiss-model (last visited: 21.06.2021), which is shown on **Figure 3.4**. EOMT is modelled as a homodimer based on a (S)-norcoclaurine 6-*O*-methyltransferase (template ID: 5icc1.A) (Robin *et al.*, 2016), with a sequence identity of 43.7%.

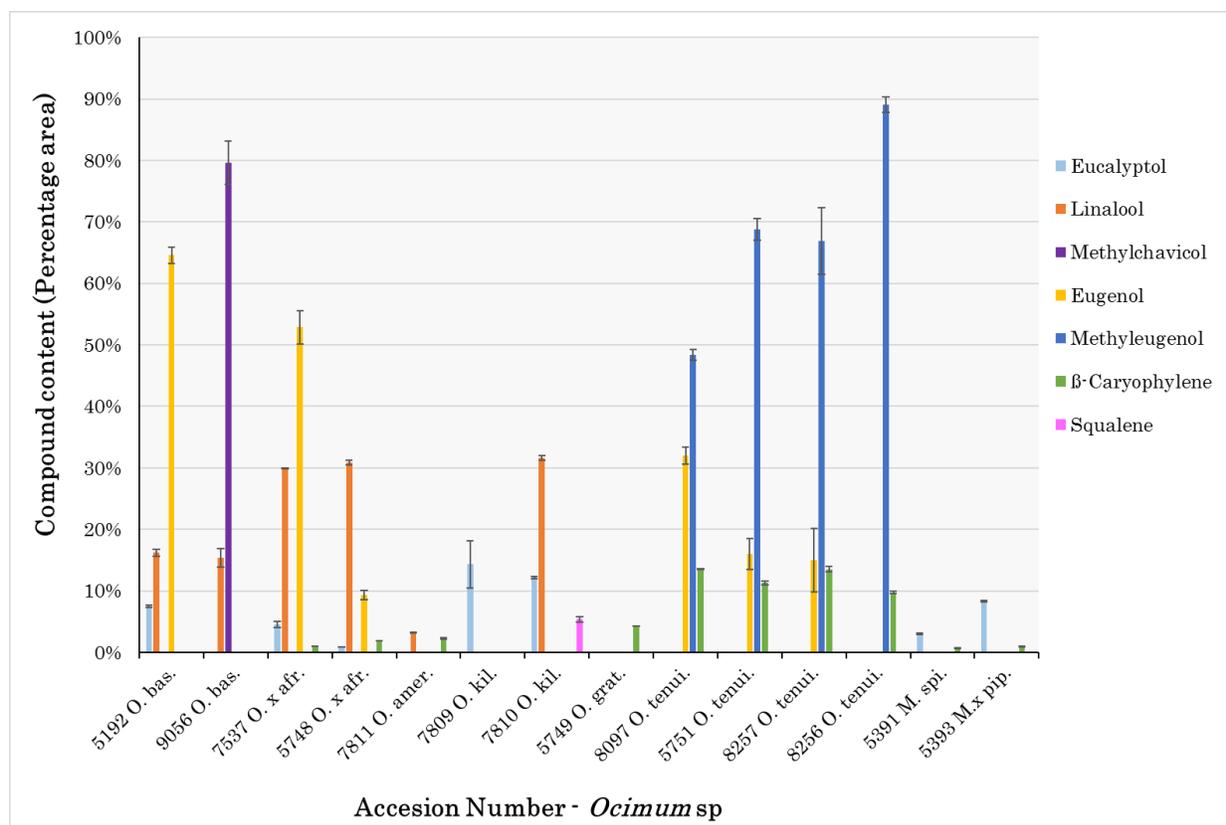


**Figure 3.4** Suggested model for a eugenol *O*-methyltransferase homodimer from *O. tenuiflorum*, obtained from the server Swiss-model using a (S)-norcoclaurine 6-*O*-methyltransferase as template (template ID: 5icc1.A). Cold (blue) to warm (red) amino acidic residues colours define N- to C-terminus. Arrows show delimitation of the *O*-methyltransferase domain in each monomer (comprising amino acidic residues 132-338). Black arrow indicates amino acidic residue position 132, red arrow indicates amino acidic residue position 338.

### 3.1.3 Chemical profile suggest *O. tenuiflorum* is a methyleugenol producer species compared to other *Ocimum* sp.

The volatiles' profile done in *Ocimum* species shows important differences among the individuals (**Figure 3.5**). The compounds here identified can be summarized in two groups, terpenes and phenylpropenes. The terpenes found were the monoterpenes eucalyptol and linalool, the sesquiterpene  $\beta$ -caryophyllene and the triterpene squalene. The phenylpropenes found were methylchavicol, eugenol and methyleugenol.

Particularly between the Tulsi clade (*O. tenuiflorum* and *O. gratissimum*) and the Basil clade (*O. basilicum*, *O. americanum*, *O. x africanum* and *O. kilimandscharicum*) identified in **section 3.1.1**, there are clear differences. For instance, eucalyptol, linalool, methylchavicol and squalene can be found only in the Basil clade. Eugenol and  $\beta$ -caryophyllene can be found in both clades, and methyleugenol can be found only in the Tulsi clade. In this last clade however, methyleugenol can be found only in *O. tenuiflorum*, not in *O. gratissimum*. This is particularly interesting since as mentioned above, the amino acid profile of the EOMT sequence from *O. gratissimum* has similarities with the sequences from the Basil clade. Now, one can hypothesise that the different *aa* found between *O. tenuiflorum* and *O. gratissimum*, that are at the same time similar to those found in the Basil clade, might be linked to the methyleugenol presence/absence in those species.

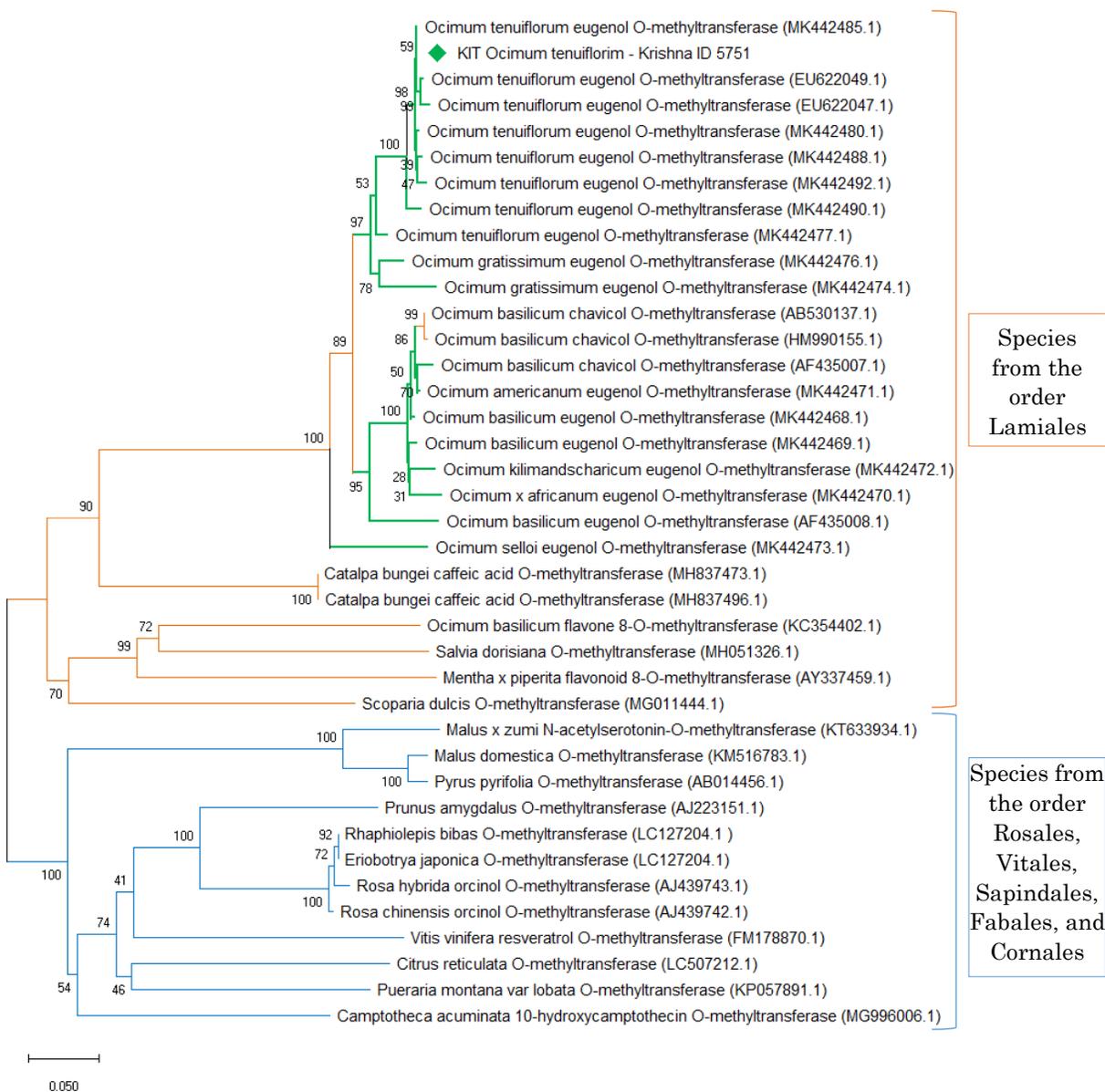


**Figure 3.5** Chemical profile of different *Ocimum* species extracts, and *Mentha* species as an outgroup. The analysis was done by gas chromatography. The volatiles compounds were identified according to their retention time (RT) compared to standards: eucalyptol (RT: 13.253), linalool (RT: 16.934 min), methylchavicol (RT: 21.636 min), eugenol (RT: 28.885 min), methyleugenol (RT: 30.985 min), β-caryophyllene (RT: 31.364 min) and squalene (RT: 53.168 min). The compounds content was calculated as percentage area from the total analytes in each chromatogram. Error bars indicate standard error of three biological replicates. For details on the accessions refer to **Table 2.1**.

### 3.1.4 OMT-based phylogeny with other species than *Ocimum* shows cluster of order Lamiales

A nucleotide blast search in NCBI web site was carried out using the EOMT sequence obtained from *O. tenuiflorum* ID 5751. The OMT-based matches were from the following species (family): *Ocimum* (Lamiaceae), *Salvia* (Lamiaceae), *Mentha* (Lamiaceae), *Catalpa* (Bignoniaceae), *Scoparia* (Plantaginaceae), *Malus* (Rosaceae), *Pyrus* (Rosaceae), *Prunes* (Rosaceae), *Rhaphiolepis* (Rosaceae), *Eriobotrya* (Rosaceae), *Rosa* (Rosaceae), *Vitis* (Vitaceae), *Citrus* (Rutaceae), *Pueraria* (Fabaceae) and *Camptotheca* (Nyssaceae). The phylogeny of the found OMTs is shown in **Figure 3.6**. EOMT from *O. tenuiflorum* is clustered with the sequences of other EOMT also belonging to *Ocimum* sp. The same pattern found in this analysis (**Figure 3.6**), where species related to *O. basilicum* are clustered apart from *O. tenuiflorum*, was previously observed when analysing only EOMT sequences from the genus *Ocimum* (**Figure 3.1**) Moreover, the Tulsi clade only has EOMT sequence, whereas the Basil clade is formed by sequences from EOMT and chavicol *O*-methyltransferase (CVOMT).

The OMTs phylogeny results also show that the species belonging to the families Lamiaceae, Bignoniaceae and Plantaginaceae are clustered together. The three families belong to the order Lamiales (**Figure 3.6**). The other species that are together in one clade, have no EOMT but other OMT enzyme types, and these were obtained from species belonging to several families belonging to the orders Rosales, Vitales, Sapindales, Fabales, and Cornales (**Figure 3.6**).



**Figure 3.6** *O*-methyltransferase (OMT) genes based phylogenetic tree. Eugenol *O*-methyltransferase from *Ocimum* sp. labelled in green branches. OMT from species belonging to the Lamiales order labelled in orange branches. OMT from species belonging to the Rosales, Viales, Sapindales, Fabales, and Cornales orders are labelled in blue branches. Neighbour-joining distant tree, bootstrap values (based on 1000 replicates) next to the branches.

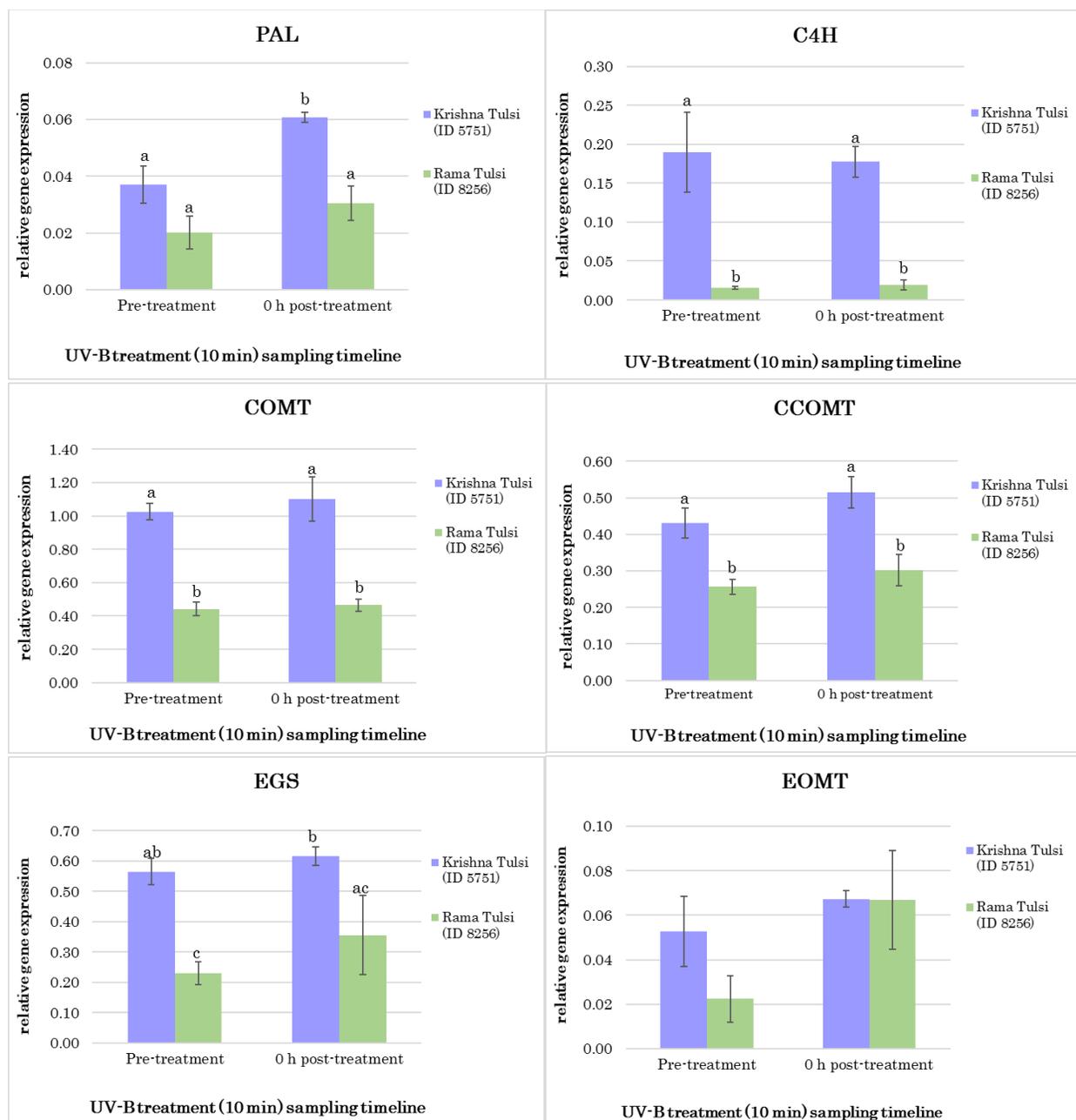
## 3.2 Phenylpropanoid Pathway in *O. tenuiflorum*

The phenylpropanoid pathway in *O. tenuiflorum* was studied in two chemotypes, Krishna Tulsi and Rama Tulsi. The key findings suggest that the two chemotypes have different regulation mechanisms for the synthesis of methyleugenol. Following, the results leading to the finding will be revised in detail.

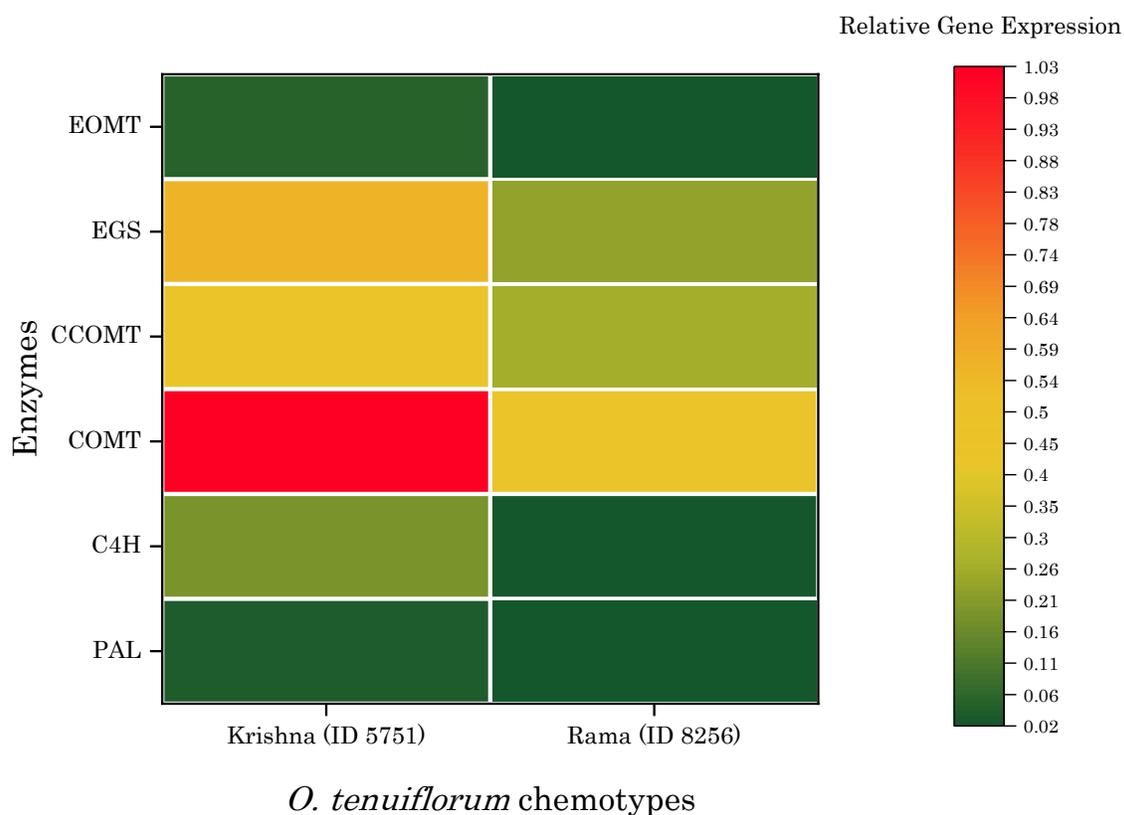
### 3.2.1 Steady-state relative gene expression of enzymes from the phenylpropanoid pathway reveals similarities and differences in two *O. tenuiflorum* chemotypes

In order to better understand the regulation mechanisms for the synthesis of methyleugenol in *Ocimum* sp. and to establish the differences related to the phenylpropanoid pathway between two *O. tenuiflorum* chemotypes -Krishna Tulsi (ID 5751) and Rama Tulsi (ID 8256)-, UV-B stress was used to study this pathway. The relative gene expression (RGE) of selected enzymes from the phenylpropanoid pathway were analysed from leaf samples obtained from plants under control conditions, and plants that were exposed to 10 min of UV-B radiation (10 W/m<sup>2</sup>). As a result, Krishna and Rama show similarities and differences in the relative gene expression of the selected enzymes (see **Figure 3.7**).

The main similarities among both chemotypes are: **(1)** the lower expression of PAL, C4H and EOMT compared to the rest of the enzymes; **(2)** the over expression of COMT compared to all of the rest of enzymes; and **(3)** the increase on the RGE of the enzymes after being exposed to 10 min of UV-B radiation suggesting post-transcriptional/translational regulations (see **Figure 3.7** and **Figure 3.8**). From the last point, PAL was the enzyme that showed the highest increment on its RGE immediately after the UV-B treatment. In contrast, the main difference between the two chemotypes is the lower RGE of each enzyme in Rama compared to its equal in Krishna (see **Figure 3.8**).

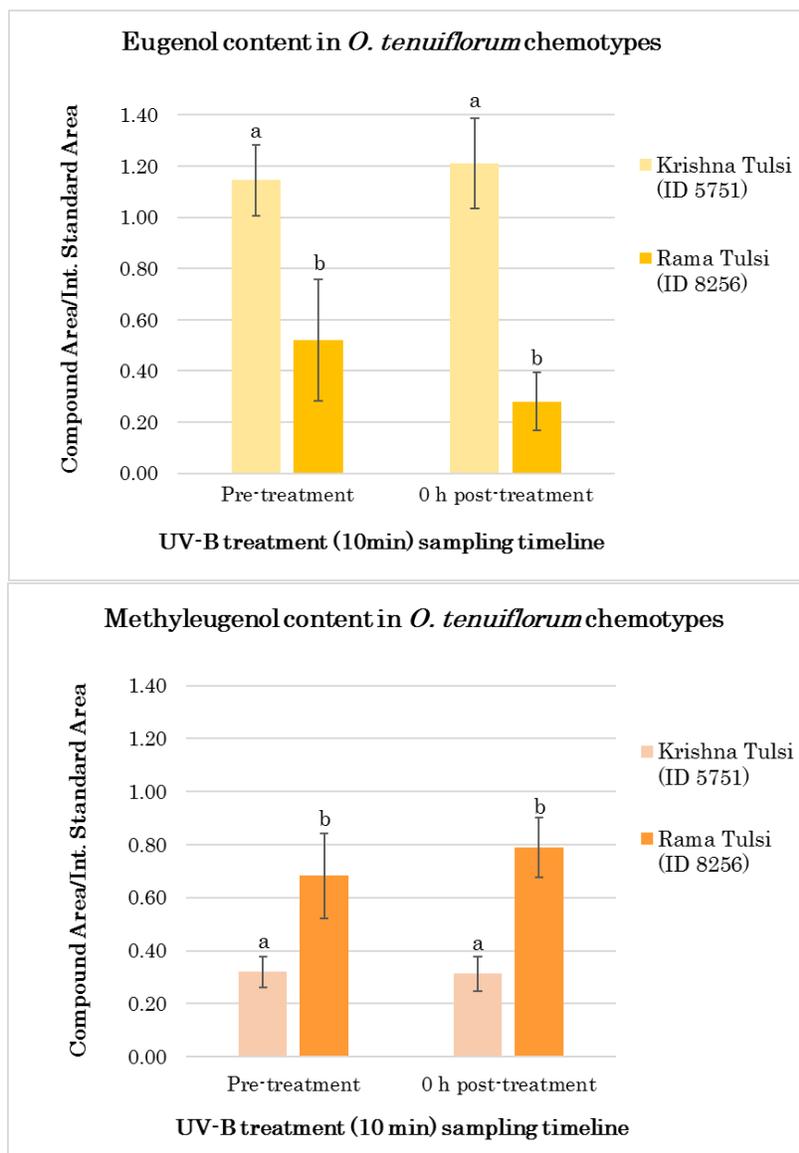


**Figure 3.7.** Relative gene expression of selected enzymes from the phenylpropanoid pathway in two *Ocimum tenuiflorum* chemotypes, Krishna Tulsi (ID 5751) and Rama Tulsi (ID 8256), in the frame of a UV-B radiation experiment. Relative gene expression was measured in expanded leaves from plants not exposed to UV-B radiation (Pre-treatment) and immediately after 10 min of UV-B radiation (0 h post-treatment). The enzymes analysed in this experiment were: Phenylalanine ammonia lyase (PAL), Cinnamate 4-hydroxylase (C4H), Caffeic acid *O*-methyltransferase (COMT), Caffeoyl-CoA *O*-methyltransferase (CCOMT), Eugenol (and Chavicol) synthase (EGS), and Eugenol *O*-methyltransferase (EOMT). Error bars indicate standard error of three biological replicates. Different letters indicate statistically significant differences ( $p < 0.05$ ).



**Figure 3.8.** Heat map summarizing the results of the steady state relative gene expression of selected enzymes from the phenylpropanoid pathway in two *Ocimum tenuiflorum* chemotypes, Krishna (ID 5751) and Rama (ID 8256). Relative gene expression was measured in expanded leaves from plants not exposed to UV-B radiation. The enzymes analysed in this experiment were: Phenylalanine ammonia lyase (PAL), Cinnamate 4-hydroxylase (C4H), Caffeic acid *O*-methyltransferase (COMT), Caffeoyl-CoA *O*-methyltransferase (CCOMT), Eugenol (and Chavicol) synthase (EGS), and Eugenol *O*-methyltransferase (EOMT).

Further, the content of the volatiles eugenol and methyleugenol were determined for both Tulsi chemotypes in the frame of the UV-B experiment (**Figure 3.9**). Both phenylpropenes show no significant variability during the experiment, in both chemotypes. However, differences between the two chemotypes were observed. Krishna has significant higher eugenol content compared to Rama, this in line with the relative gene expression results, shown above. Contrary to this, Krishna had significantly lower methyleugenol content than Rama, though the RGE of EOMT was higher in Krishna, which could suggest a more efficient EOMT in Rama.



**Figure 3.9.** Eugenol (upper) and Methyleugenol (lower) content in two *Ocimum tenuiflorum* chemotypes, Krishna Tulsi (ID 5751) and Rama Tulsi (ID 8256), in the frame of a UV-B radiation experiment. Volatiles were measured in expanded leaves from plants not exposed to UV-B radiation (Pre-treatment) and immediately after 10 min of UV-B radiation (0 h post-treatment). Error bars indicate standard error of three biological replicates. Different letters indicate statistically significant differences ( $p < 0.05$ ).

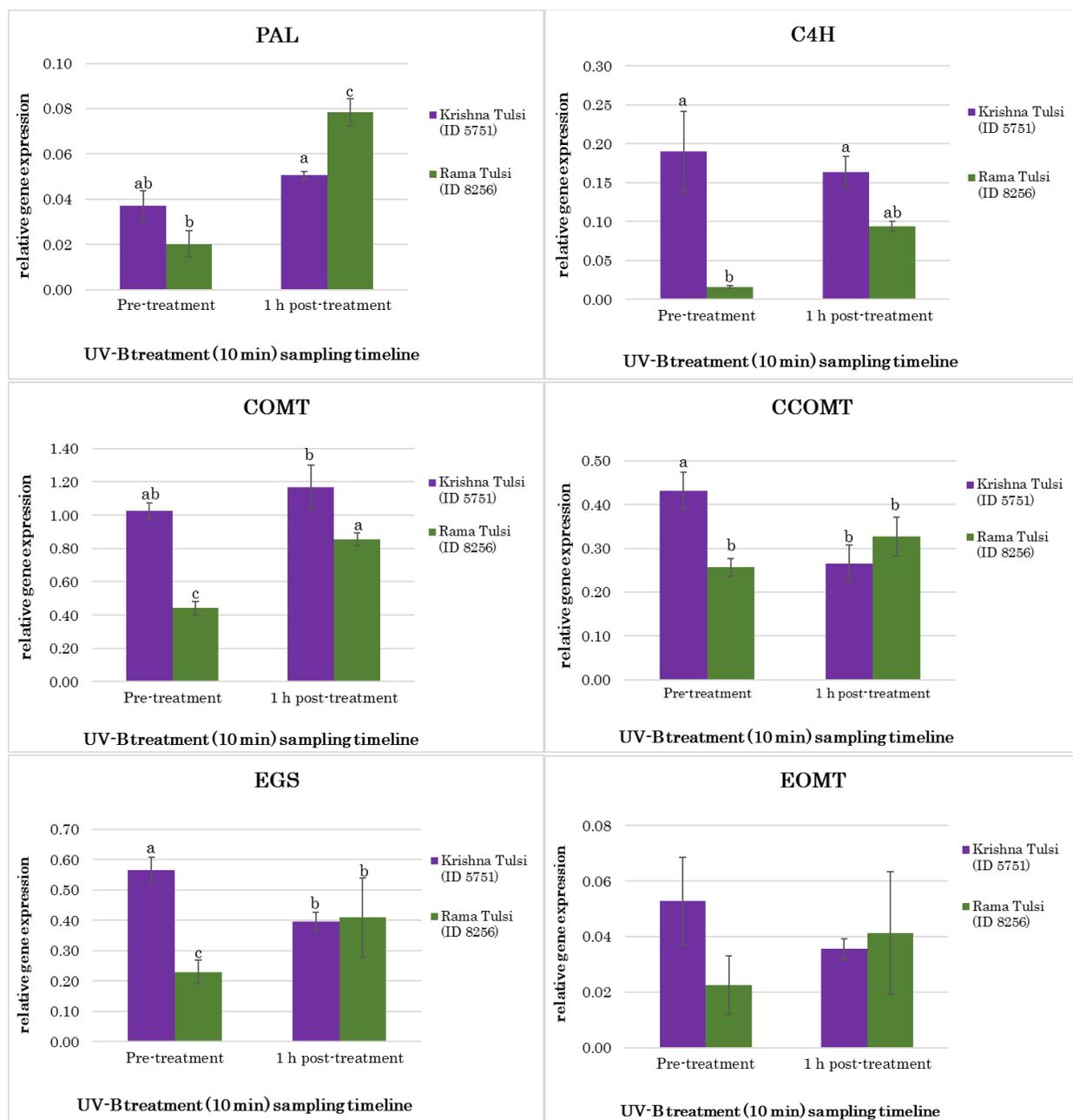
### 3.2.2 Enzymes relative gene expression shows differential regulation mechanisms for methyleugenol synthesis as response to UV-B stress in two *O. tenuiflorum* chemotypes

The phenylpropanoid pathway for the synthesis of methyleugenol in *Ocimum* sp. (see **Figure 1.5**) has been suggested by different authors. However, the regulation mechanisms are yet to be understood. Therefore, the relative gene expression (RGE) of selected enzymes from the phenylpropanoid pathway were analysed from leaf samples obtained from plants under control conditions, and plants that were exposed to 10 min of UV-B radiation (10 W/m<sup>2</sup>) and sampled one hour after the treatment. The results show different regulation mechanisms for the methyleugenol synthesis as response to UV-B stress between the two *O. tenuiflorum* chemotypes, Krishna and Rama (see **Figure 3.10** and **Figure 3.11**).

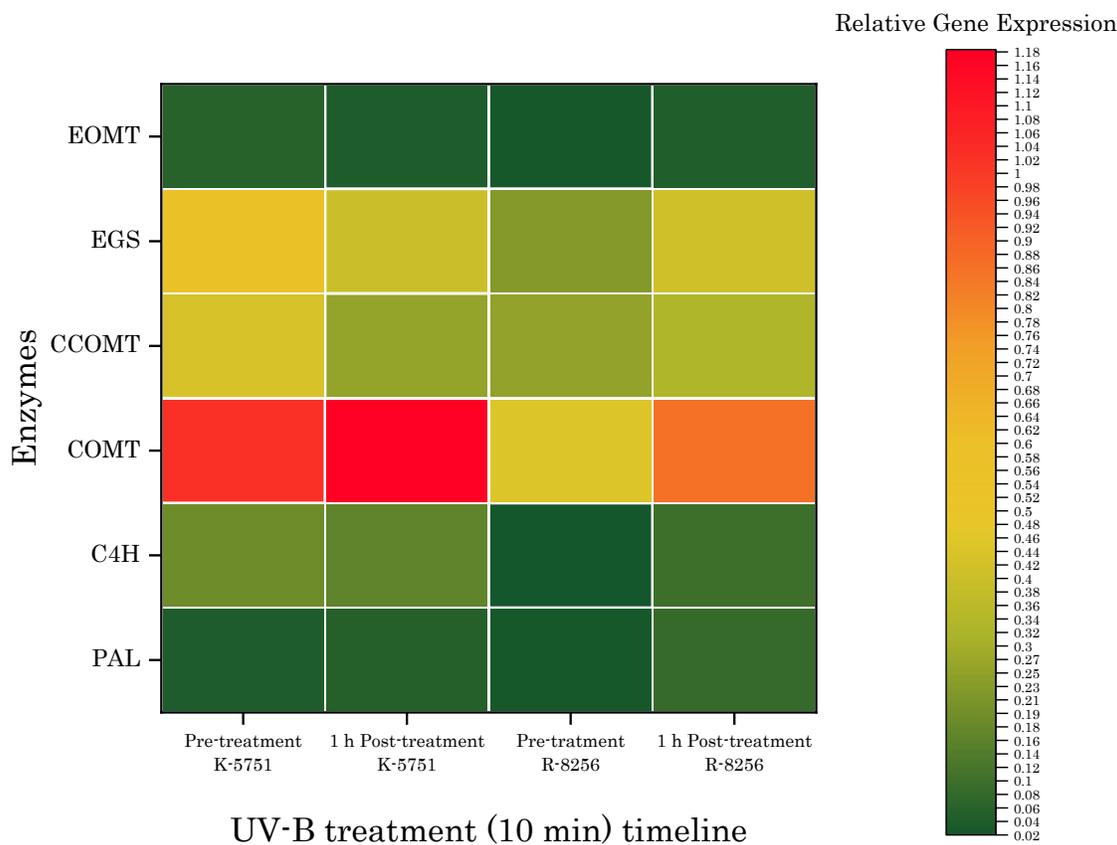
In Krishna Tulsi the phenylpropanoid pathway for the synthesis of methyleugenol was down-regulated under UV-B stress. The RGE of the enzymes PAL and COMT increased but not significantly (**Figure 3.10**). Contrary to this, the RGE of CCOMT and EGS, and significantly decrease during the UV-B experiment (**Figure 3.10**). Whereas EOMT RGE decrease but not significantly. These results seem to suggest that in Krishna Tulsi, CCOMT could be playing an important role in the synthesis of ME.

In Rama Tulsi the phenylpropanoid pathway for the synthesis of methyleugenol was up-regulated under UV-B stress. The RGE of PAL, COMT and EGS significantly increased during the UV-B experiment (**Figure 3.10**). While the RGE of C4H, CCOMT and EOMT an upregulation was observed, but no significant (**Figure 3.10**). These results suggest that in Rama Tulsi, COMT could be more important for the synthesis of ME than CCOMT.

Overall, the results on the Tulsi chemotypes one hour after the UV-B treatment (**Figure 3.11**), shows in Krishna: higher RGE of C4H and COMT, similar levels of RGE for CCOMT, EGS and EOMT (being these three enzymes down regulated in Krishna and upregulated in Rama), and lower levels of RGE of PAL when compared to Rama. These results suggest different regulation mechanisms for the regulation of the synthesis of ME in *O. tenuiflorum* chemotypes, and different responses to UV-B stress.



**Figure 3.10.** Relative gene expression of selected enzymes from the phenylpropanoid pathway in two *Ocimum tenuiflorum* chemotypes, Krishna Tulsi (ID 5751) and Rama Tulsi (ID 8256), in the frame of a UV-B radiation experiment. Relative gene expression was measured in expanded leaves from plants not exposed to UV-B radiation (Pre-treatment) and one hour after 10 min of UV-B radiation (1 h post-treatment). The enzymes analysed in this experiment were: Phenylalanine ammonia lyase (PAL), Cinnamate 4-hydroxylase (C4H), Caffeic acid *O*-methyltransferase (COMT), Caffeoyl-CoA *O*-methyltransferase (CCOMT), Eugenol (and Chavicol) synthase (EGS), and Eugenol *O*-methyltransferase (EOMT). Error bars indicate standard error of three biological replicates. Different letters indicate statistically significant differences ( $p < 0.05$ ).

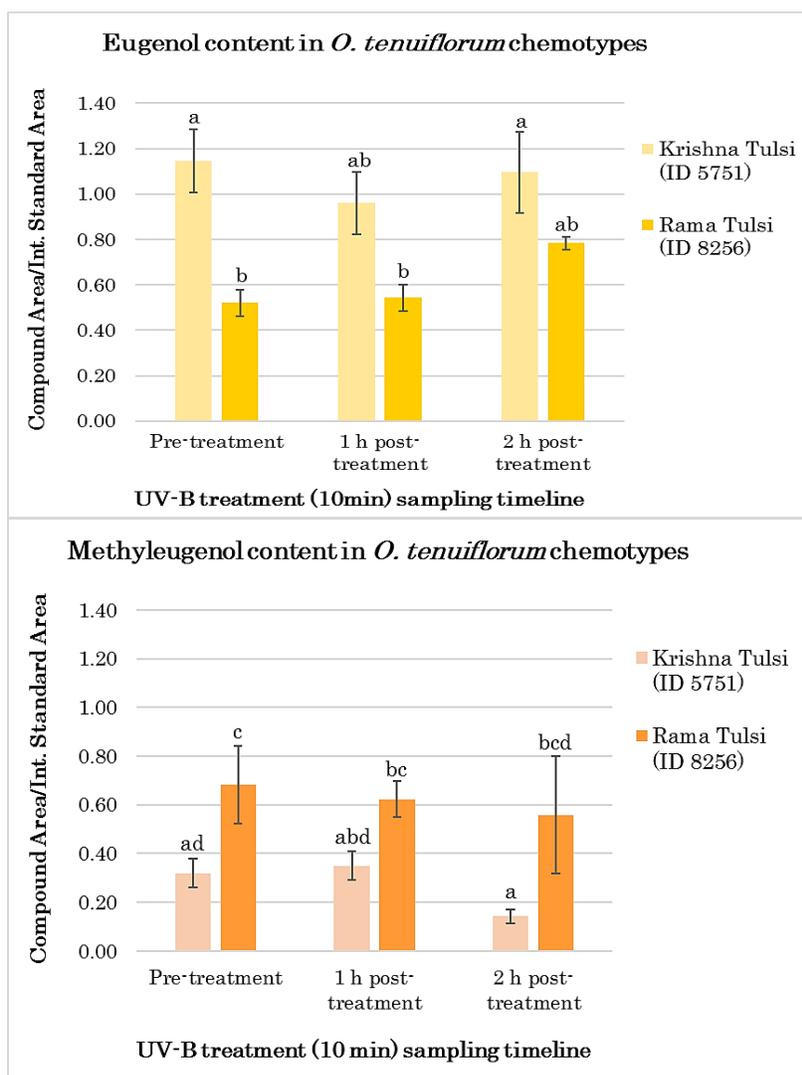


**Figure 3.11.** Heat map summarizing the results of the relative gene expression of selected enzymes from the phenylpropanoid pathway in two *Ocimum tenuiflorum* chemotypes, Krishna (K-5751) and Rama (R-8256), in the frame of a UV-B radiation experiment. Relative gene expression was measured in expanded leaves from plants not exposed to UV-B radiation (Pre-treatment), and one hour after 10 min of UV-B radiation (1 h post-treatment). The enzymes analysed in this experiment were: Phenylalanine ammonia lyase (PAL), Cinnamate 4-hydroxylase (C4H), Caffeic acid *O*-methyltransferase (COMT), Caffeoyl-CoA *O*-methyltransferase (CCOMT), Eugenol (and Chavicol) synthase (EGS), and Eugenol *O*-methyltransferase (EOMT).

### 3.2.3 Chemical profile analysis supports findings on differential gene expression and response to UV-B stress for two *O. tenuiflorum* chemotypes

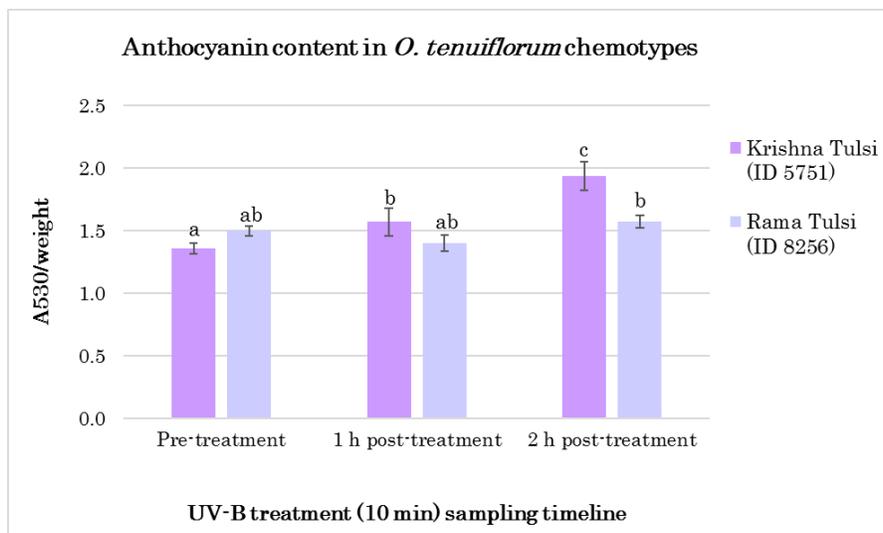
The content of the volatiles eugenol and methyleugenol were determined for both Tulsi chemotypes in the frame of the UV-B experiment (Figure 3.12). Comparing the two chemotypes, the content of eugenol in Krishna is generally higher than in Rama. However, two hours after the UV-B stress the levels of eugenol in Rama increase to the point that it is not anymore significantly different from Krishna. This is in line with the relative gene expression results shown above, where the phenylpropanoid pathway for the synthesis of methyleugenol in Krishna is suppressed.

The content of methyleugenol in Krishna is generally lower than in Rama. However, the significant differences are not aligned with the gene expression suggesting the EOMT from Krishna could be less efficient on synthesising ME.



**Figure 3.12.** Eugenol (upper) and Methyleugenol (lower) content in two *Ocimum tenuiflorum* chemotypes, Krishna Tulsi (ID 5751) and Rama Tulsi (ID 8256), in the frame of a UV-B radiation experiment. Volatiles were measured in expanded leaves from plants not exposed to UV-B radiation (Pre-treatment), one hour after 10 min of UV-B radiation (1 h post-treatment) and two hours after 10 min of UV-B radiation (2 h post-treatment). Error bars indicate standard error of three biological replicates. Different letters indicate statistically significant differences ( $p < 0.05$ ).

The anthocyanin content was determined for both Tulsi chemotypes in the frame of the UV-B experiment (**Figure 3.13**). Except for the anthocyanin content in Krishna which significantly increased 2 hours after the UV-B treatment, the anthocyanin content showed mostly no significant variability during the experiment, in both chemotypes. The results for Krishna might explain the down-regulation of genes related to the synthesis of methyleugenol and the lower methyleugenol content found in this chemotype.

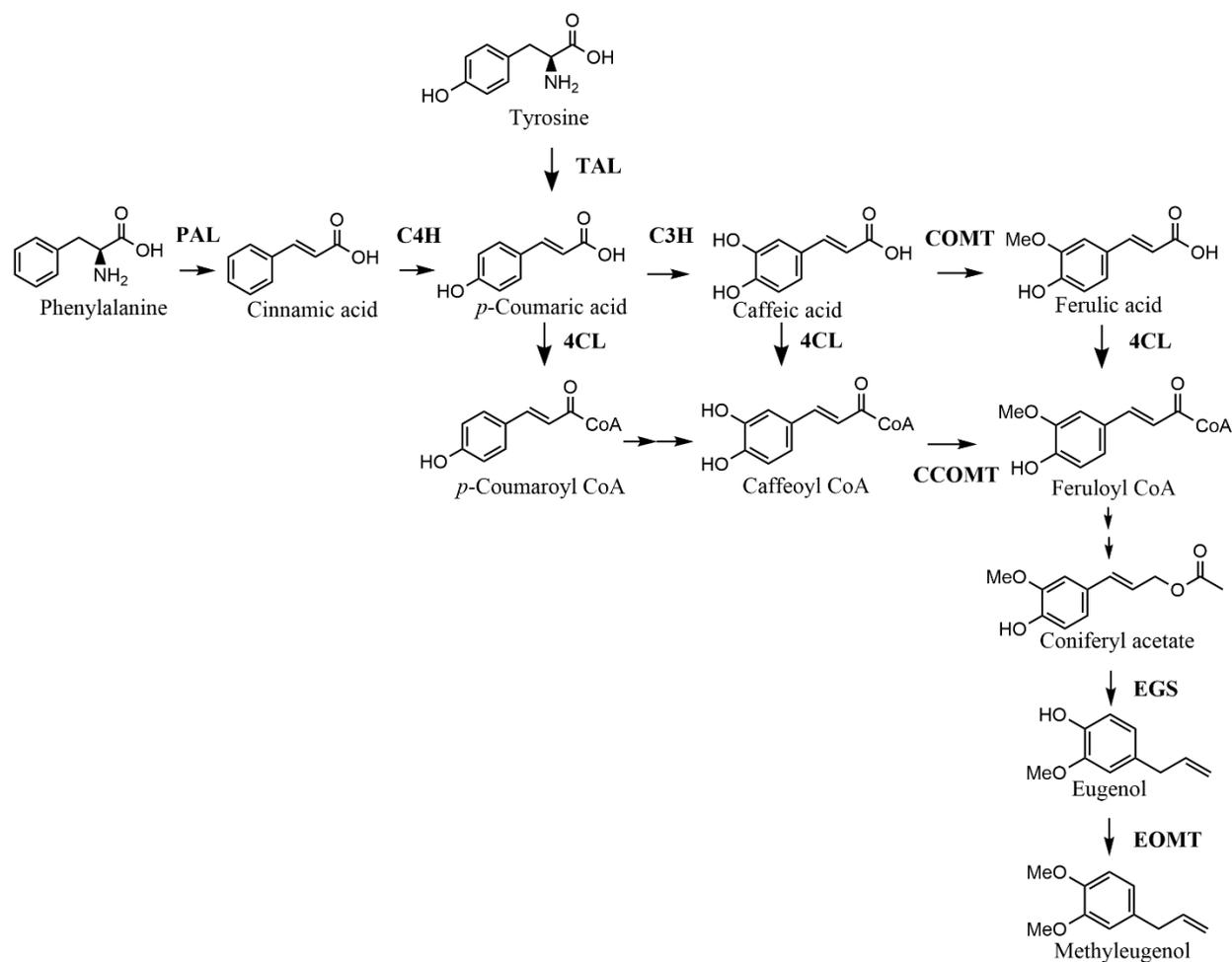


**Figure 3.13.** Anthocyanin content in two *Ocimum tenuiflorum* chemotypes, Krishna Tulsi (ID 5751) and Rama Tulsi (ID 8256), in the frame of a UV-B radiation experiment. Anthocyanin content was measured in expanded leaves from plants not exposed to UV-B radiation (Pre-treatment), one hour after 10 min of UV-B radiation (1 h post-treatment) and two hours after 10 min of UV-B radiation (2 h post-treatment). Error bars indicate standard error of three biological replicates. Different letters indicate statistically significant differences ( $p < 0.05$ ).

### 3.2.4 Model of the phenylpropanoid pathway for the synthesis of methyleugenol in two *O. tenuiflorum* chemotypes

From the results shown in this section, a model for the phenylpropanoid pathway in the two *O. tenuiflorum* chemotypes is suggested (**Figure 3.14**). The main differences are that first; it seems that the upstream reactions to the synthesis of methyleugenol in Rama Tulsi is highly regulated by the enzyme caffeic acid *O*-methyltransferase (COMT). Whereas in Krishna Tulsi, the enzyme caffeoyl-CoA *O*-methyltransferase (CCOMT) seems to also play an important role in regulation.

Second, the RGE level of enzymes such as COMT and EGS in control samples suggest this pathway is active, but rather independently from PAL at this stage. Therefore, another source of substrates to justify the higher expressions of COMT and EGS could be tyrosine, catalysed by the enzyme tyrosine ammonia lyase (TAL).



**Figure 3.14.** Phenylpropanoid pathway for the synthesis of methyleugenol in *Ocimum tenuiflorum*. Double arrows indicate intermediated suggested reactions not shown. The enzymes shown in the pathway are: Phenylalanine ammonia lyase (PAL), Cinnamate 4-hydroxylase (C4H), Tyrosine ammonia lyase (TAL), 4-coumarate 3-hydroxylase (C3H), 4-Coumarate-CoA ligase (4CL), Caffeic acid *O*-methyltransferase (COMT), Caffeoyl-CoA *O*-methyltransferase (CCOMT), Eugenol (and Chavicol) synthase (EGS), and Eugenol *O*-methyltransferase (EOMT).

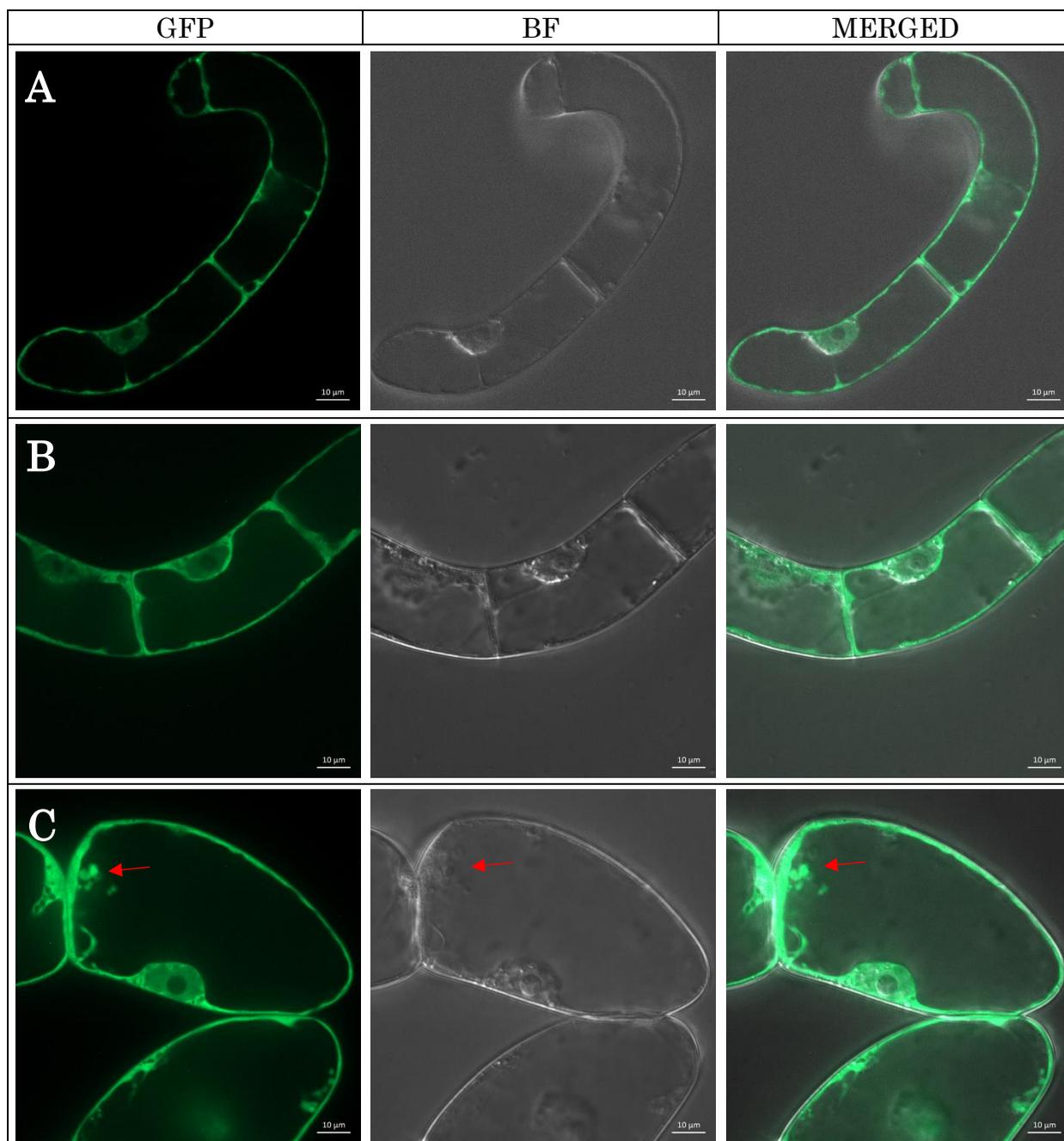
### 3.3 The subcellular localization of *O. tenuiflorum* EOMT is in the cytosol and the nucleus

For a better comprehension on the regulation mechanisms for the synthesis of ME in *Ocimum* sp., the EOMT from *O. tenuiflorum* was fused with green fluorescence protein and expressed in BY-2 cells in order to determine its subcellular localization. Results show that EOMT is localized in the nucleus and in the cytosol of BY-2 transformed cells (**Figure 3.15** and **Figure 3.16**). These results correspond to transformed cells with a construct containing a N-terminally (**Figure 3.15**) and C-terminally fused GFP to EOMT (**Figure 3.16**).

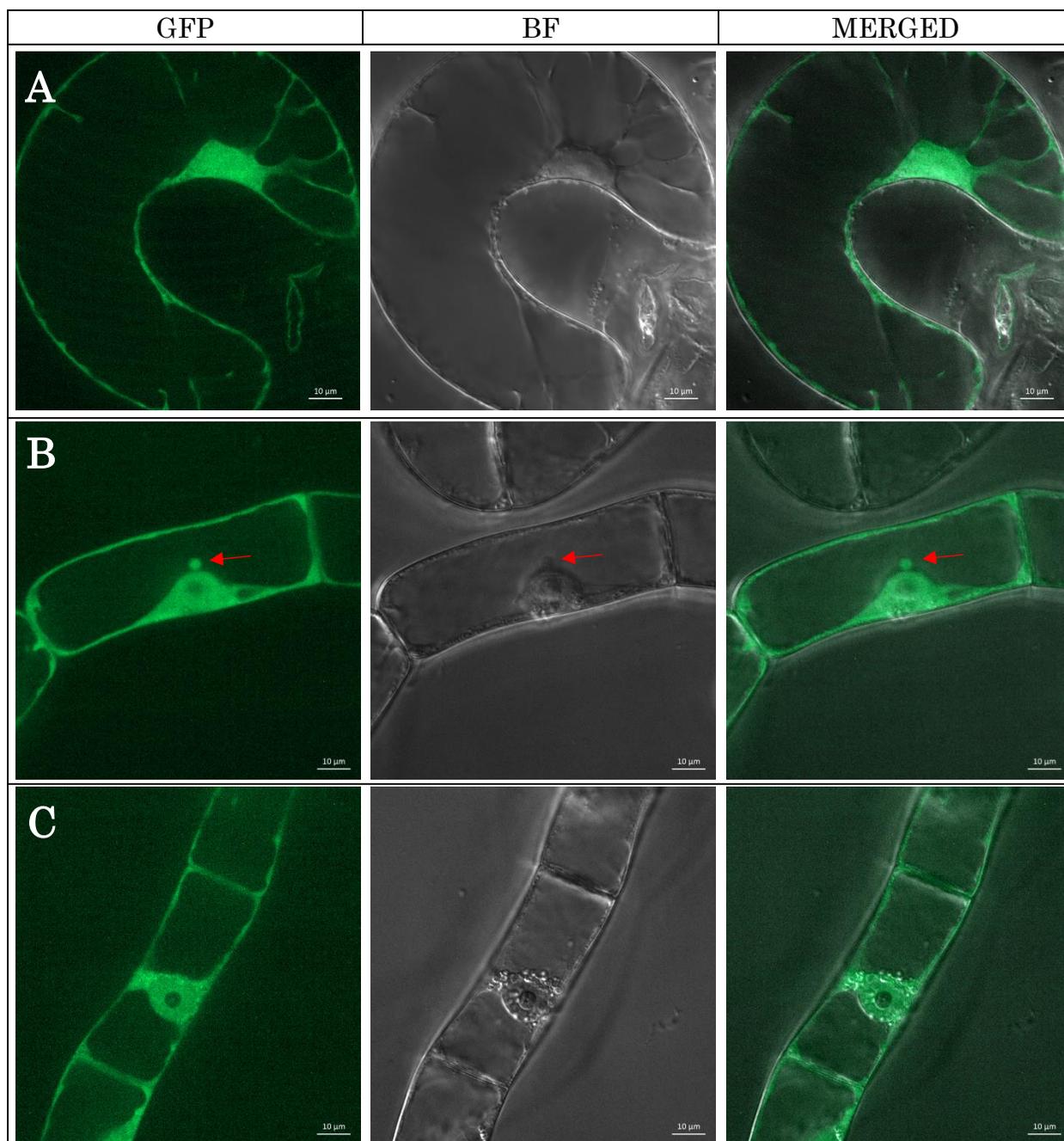
In addition, the examination for localization peptides using the search tools of the websites SignalP 5.0 (<http://www.cbs.dtu.dk/services/SignalP/> -last visited June 2021-) gave no positive results. Similarly, the search for nuclear export sequences (NESs) at the database from TUD (<http://www.cbs.dtu.dk/services/NetNES/>) and Rostlab-TUM (<https://rostlab.org/services/nlsdb/>), gave no positive results.

Together with these results, the images on transformed BY-2 cell also show vesicles in the cytoplasm that emit the GFP signal near the plasma membrane (**Figure 3.15c**) and near the nucleus (**Figure 3.16b**). Since GFP can be found free in the cytosol, this suggests the EOMT enzyme can move inside the cell via vesicle transport.

Further, protein was extracted from transformed BY-2 cell cultures, in order to do western blot assays. The western blot results (see appendix **Figure A.6**) showed bands at around 70kDa, corresponding to the expected size of EOMT-GFP fusion. These bands were not found in the positive control done with BY-2 wild type cells. However, the results also show several unspecific antibody binding to non GFP proteins. Therefore, more assays are needed in order to confirm that EOMT is effectively localized in the nucleus and is not a misleading result caused by non-fused GFP.



**Figure 3.15.** Live cell images of BY2 transformed cells (A, B and C), showing cytoplasmic and nuclear localization of the enzyme eugenol *O*-methyltransferase (EOMT) from *Ocimum tenuiflorum*. BY-2 wild type cells were transformed using the Gateway® vector system pH7WGF2.0 with N-terminally fused GFP to EOMT (called EOMT-*W* in this study). From left to right is shown in each case (A, B and C) the green fluorescence signal (GFP), the bright field signal (BF), and the combined GFP and BF signals (Merged). Red arrow indicate vesicles with the GFP signal.



**Figure 3.16.** Live cell images of BY2 transformed cells (A, B and C), showing cytoplasmic and nuclear localization of the enzyme eugenol *O*-methyltransferase (EOMT) from *Ocimum tenuiflorum*. BY-2 wild type cells were transformed using the Gateway® vector system pH7FWG2.0 with C-terminally fused GFP to EOMT (called EOMT-*F* in this study). From left to right is shown in each case (A, B and C) the green fluorescence signal (GFP), the bright field signal (BF), and the combined GFP and BF signals (Merged). Red arrow indicate a vesicle with the GFP signal.

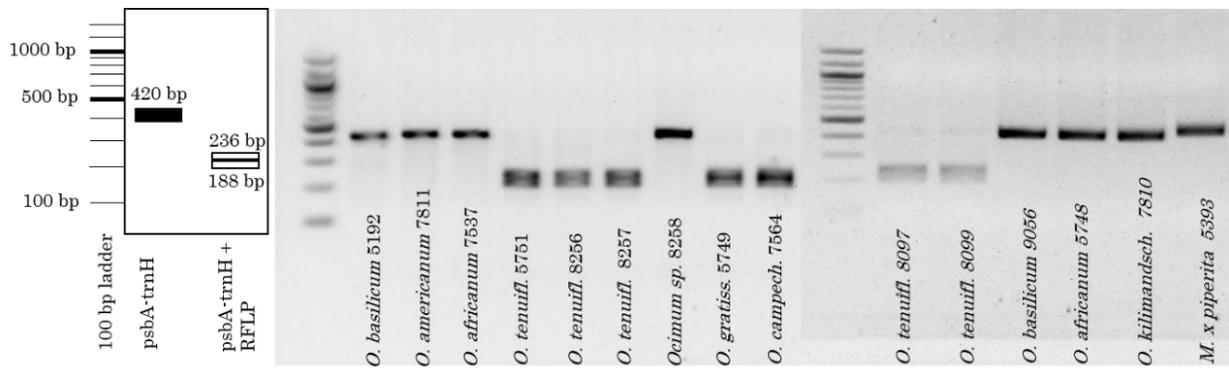
### 3.4 DNA barcoding for *O. tenuiflorum* discrimination

Discrimination of *O. tenuiflorum* from other species in order to identify surrogating species in commercial samples was done using DNA barcoding techniques. A one single reaction assay using a trait-independent marker (*psbA-trnH* *igs*) coupled together with a trait-related marker based on the enzyme eugenol *O*-methyltransferase (EOMT<sup>®</sup>) have shown positive results for the discrimination of *O. tenuiflorum* from other *Ocimum* sp. reference plants and in commercial samples. This results' section was published in Ríos-Rodríguez *et al.* (2021).

#### 3.4.1 Trait-independent marker *psbA-trnH* coupled with RFLP allows to discriminate *O. tenuiflorum* from other species in a two-step protocol

In order to distinguish *O. tenuiflorum* from other *Ocimum* species a DNA barcoding assay based on the amplification of the *psbA-trnH* *igs* marker followed by RFLP was developed (Jürges *et al.*, 2018). The validation of the method was done by using reference plants (**Figure 3.17**). Five *O. tenuiflorum* accessions gave a double band pattern at around 200 bp, predicted from the presence of a specific single-nucleotide polymorphism creating a restriction site for the enzyme Hinf I. Contrary to this, Hinf I did not digest the *psbA-trnH* PCR products from *O. basilicum*, *O. americanum*, *O. x africanum*, *O. kilimandscharicum*, and a non-identify *Ocimum* sp. obtained as “Vana Tulsi” (ID 8258). Therefore, these species presented a band at around 400 bp. Similarly, the *psbA-trnH* amplicon of *Mentha x piperita* was not digested by the enzyme.

As a result, the DNA barcoding assay permits to discriminate *O. tenuiflorum* from three other *Ocimum* sp. but shares same patterns with *O. gratissimum* and *O. campechianum*. In addition to this, the method disallows to discriminate Tulsi in a single PCR reaction.

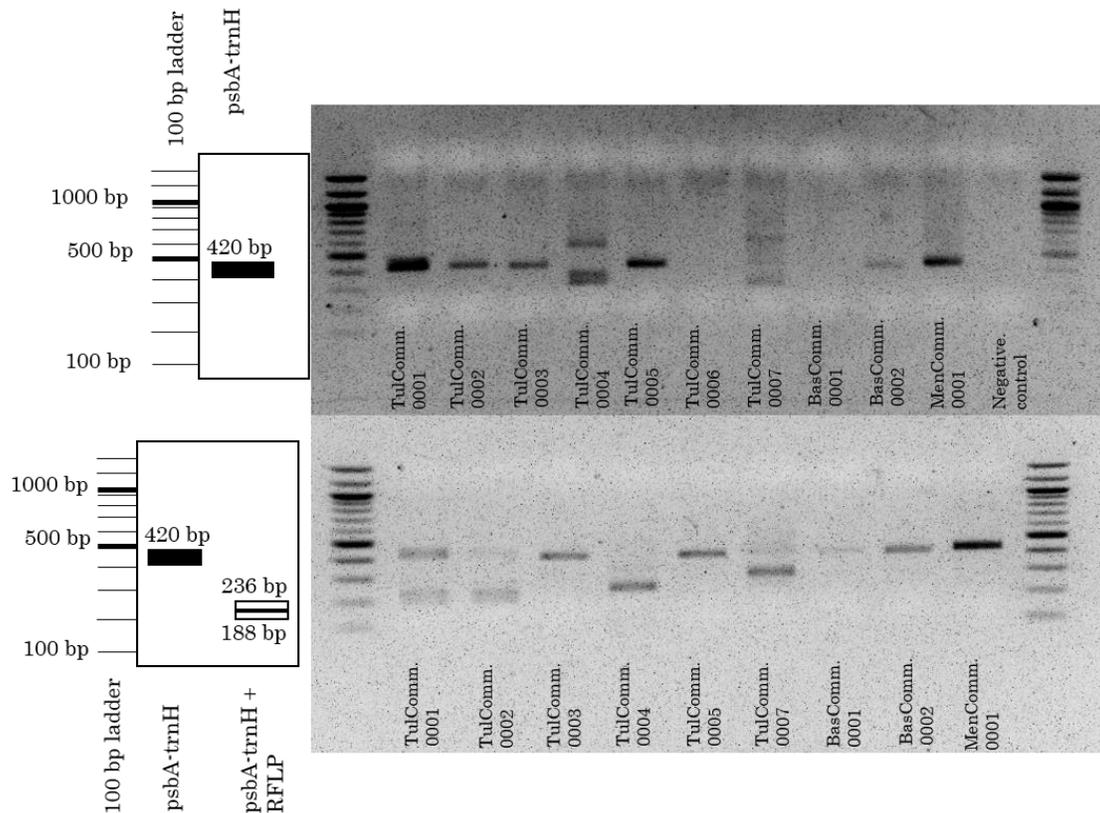


**Figure 3.17.** Discrimination of *Ocimum tenuiflorum* based on RFLP of the *psbA-trnH* DNA barcode. Left: Banding RFLP pattern predicted from the sequence of the *psbA-trnH* marker in *O. tenuiflorum*. Right: Representative gel showing amplicates for *psbA-trnH* DNA barcode followed by digestion with *Hinf*I on different validated reference plants of *Ocimum* and *Mentha* (Jürges *et al.*, 2018; Ríos-Rodríguez *et al.*, 2021).

The trait-independent marker assay was then tested in commercial samples (**Figure 3.18**). From seven commercial samples allegedly having Tulsi, only TulComm.0001 and TulComm.0002, resulted with the RFLP pattern of the *psbA-trnH* barcode. No digestion by *Hinf* I was observed for all of the other commercial samples. TulComm.0004 and TulComm.0007 had both a small band, but was not the RFLP double-band pattern of *O. tenuiflorum*. Also, these two samples had a band that was higher as the estimated 420 bp for *psbA-trnH*, probably resulting from the mixed herbs in the samples. Sample TulComm.0006 had no band at all, suggesting difficulties given by DNA degradation.

As predicted, samples BasComm.0001 and BasComm.0002, declared as Sweet Basil, were not digested. And the negative control MinComm.0001, containing Mint tea, resulted in the expected undigested amplicon. Therefore, these results suggest that commercial samples TulComm.0003, TulComm.0004, TulComm.0005 and TulComm.0007 do not contain *O. tenuiflorum*.

Though this method achieves the detection of *O. tenuiflorum* surrogates, when managing large amount of samples, a two-step protocol is laborious and requires more consumables, impacting the assays profitability. Since a one-step protocol is more attractive, alternatives using a trait-related marker were developed.



**Figure 3.18.** Discrimination of *Ocimum tenuiflorum* based on RFLP of the *psbA-trnH* marker in commercial samples. **Left top:** *psbA-trnH* marker banding for *Ocimum* sp. **Right top:** Representative gel showing amplicons for *psbA-trnH* on commercial samples. **Left down:** Banding RFLP pattern predicted from the *psbA-trnH* marker in *O. tenuiflorum*. **Right down:** Representative gel showing RFLP of the *psbA-trnH* DNA barcode with *Hinf*I on commercial samples, sample TulComm.0006 was omitted. TulComm.0001: Tea with different Tulsi types; TulComm.0002: Tea with different Tulsi types; TulComm.0003: Tulsi tea; TulComm.0004, TulComm.0005, TulComm.0006 and TulComm.0007: different brands of Tulsi tea mixed with other herbs; BasComm.0001: Dried basil; BasComm.0002: Dried basil; MenComm.0001: Mint tea. For details on samples declarations see **Table 2.4** (Ríos-Rodríguez *et al.*, 2021).

### 3.4.2 The trait-related marker EOMT has potential for discriminating *O. tenuiflorum* from other species in a one-step protocol

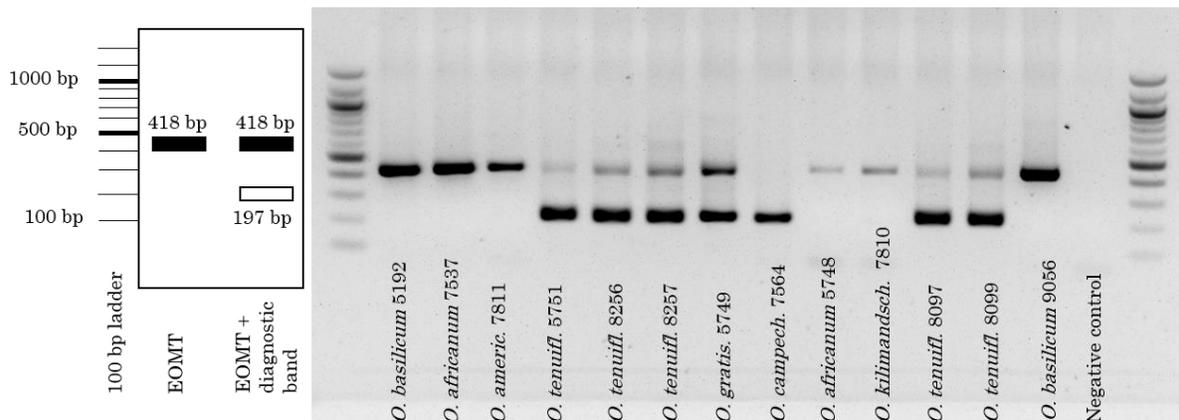
A bibliographic review on the occurrence of methyleugenol and its precursor, eugenol, in *Ocimum* species (Table 3.3) reveals a large variation on the compounds content, among and within the species. Based on this review and the fact that this research involved the analysis of the sequence of EOMT (enzyme responsible for the synthesis of ME), this enzyme was therefore chosen as trait-related marker for discriminating *O. tenuiflorum* from other species.

**Table 3.3.** Literature reports for methyleugenol (ME) contents and the related eugenol (EU) content in different species of *Ocimum* (Ríos-Rodríguez *et al.*, 2021).

<i>Ocimum</i> species	ME content* (%)	EU content* (%)	References
<i>O. tenuiflorum</i> L.	n.d. - 3.1	n.d. - 73	Jirovetz <i>et al.</i> (2003), EMEA (2005), Rajeswara Rao <i>et al.</i> (2011), Tan and Nishida (2012), Joshi (2013), Chowdhury <i>et al.</i> (2017), Raina and Misra (2018)
	20.1	51.2	
	52	27	
	56.18	1.66	
	72.5	0.9	
	73 - 92.4	2.4 - 5.8	
<i>O. basilicum</i> L.	n.d.	n.d. - 8.3	Jirovetz <i>et al.</i> (2003), Runyoro <i>et al.</i> (2010), Rajeswara Rao <i>et al.</i> (2011), Tan and Nishida (2012), Chowdhury <i>et al.</i> (2017), Raina and Misra (2018)
	0.1	7.4 - 19.3	
	1.1	n.d.	
	0.29 - 0.3	n.d. - 0.13	
	5.6 - 12.3	n.d.	
	15.53	n.d.	
<i>O. x africanum</i> Lour.	n.d.	n.d. - 0.3	Chowdhury <i>et al.</i> (2017), Raina and Misra (2018)
	0.1	0.2	
<i>O. americanum</i> L.	n.d.	n.d.	Jirovetz <i>et al.</i> (2003), Chowdhury <i>et al.</i> (2017), Raina and Misra (2018)
	0.02	1.12	
<i>O. gratissimum</i>	n.d.	5.2 - 84.1	Jirovetz <i>et al.</i> (2003), Zoghbi <i>et al.</i> (2007), Benitez <i>et al.</i> (2009), Rajeswara Rao <i>et al.</i> (2011), Joshi (2013), Chowdhury <i>et al.</i> (2017), Rawat <i>et al.</i> (2017), Raina and Misra (2018)
	0.1	62.2	
	0.28	63.36	
	14.54	61.72	
<i>O. campechianum</i> Mill.	n.d.	3.8 - 46.55	Zoghbi <i>et al.</i> (2007), Benitez <i>et al.</i> (2009), Figueiredo <i>et al.</i> (2018)
	0.2	48.3 - 60.6	
	0.3	32.2	
	9.5	40.8	
	12.0	9.0	
	60.6	7.4	
	69.5	0.2	
	80 - 87	(-)	
<i>O. kilimandscharicum</i> G.	n.d. - 0.1	n.d.	Runyoro <i>et al.</i> (2010), Rajeswara Rao <i>et al.</i> (2011), Lawal <i>et al.</i> (2014), Chowdhury <i>et al.</i> (2017)
	53.9	(-)	

\*Content from leaves oil/extracts. n.d.: not detected. (-) not mentioned.

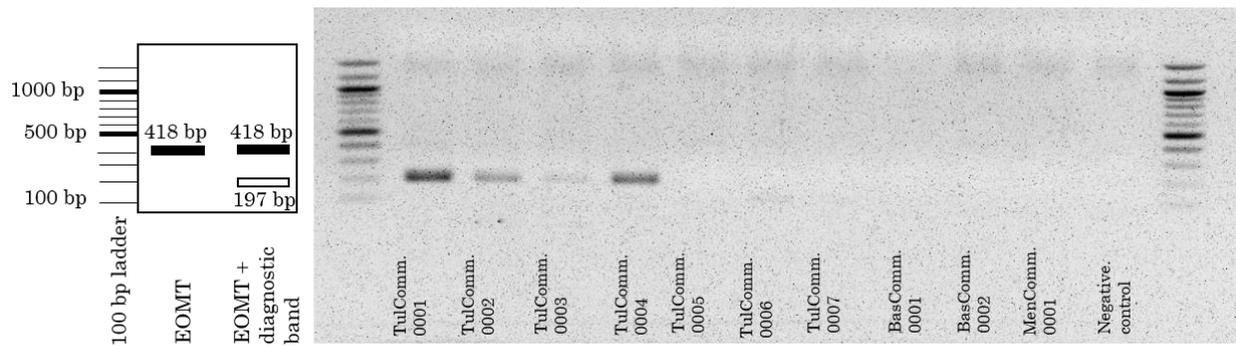
The type of the DNA barcode aimed was an Amplification Refractory Mutation System (ARMS) assay. For designing the trait-related ARMS, based on the sequence of EOMT, more than one SNP was considered. This would allow a further robust differentiation of the diagnostic ARMS primer. As a result, from the primers design, the expected pattern was a full-length amplicon of 418 bp for *Ocimum* sp. and a diagnostic second band at 197 bp in addition for *O. tenuiflorum*. The marker tested in 13 validated reference *Ocimum* accessions (**Figure 3.19**). The diagnostic band at around 197 bp was amplified in all five *O. tenuiflorum* accessions, in *O. gratissimum* and in *O. campechianum*.



**Figure 3.19.** Discrimination of *Ocimum tenuiflorum* based on the trait-related eugenol *O*-methyltransferase (EOMT) using a duplex ARMS strategy. Left: Predicted pattern for *O. tenuiflorum*. Right: Representative gel showing amplicates obtained for a core-collection of validated reference plants for *Ocimum*, where several informative SNPs in *O. tenuiflorum* allow for binding of a diagnostic primer, such that a smaller band at around 200 bp appears (Ríos-Rodríguez *et al.*, 2021).

Commercial samples declared to contain Tulsi were analysed with the EOMT DNA barcode with unsatisfactory results (**Figure 3.20**). Samples containing Tulsi should present a two band pattern as shown in **Figure 3.19**. Yet, in only four out of ten samples a single band at around 200 bp was observed, probably representing the diagnostic band. In addition, and contrary to the reference accessions results, none of the samples had the full-length band at around 400 bp, reducing the validity of the assay. Under this frame, the most plausible explanation is DNA degradation, meaning that the amount of template is insufficient to be amplified by PCR.

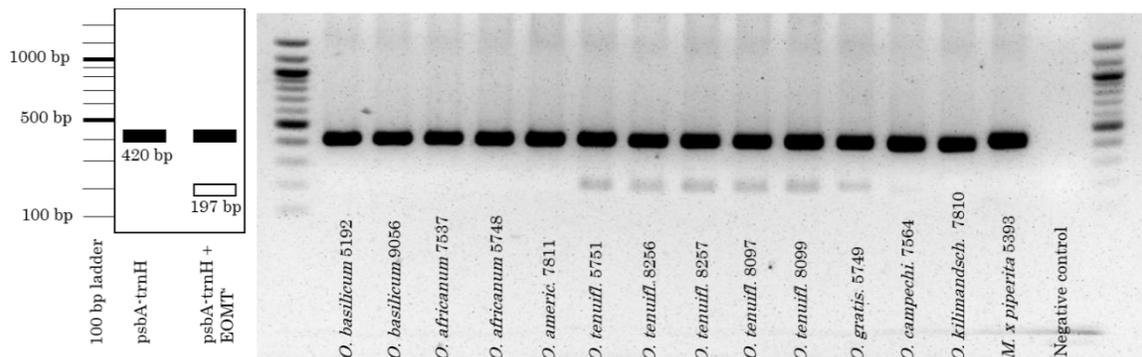
Overall, the EOMT marker in an ARMS assay is not effective in commercial samples. This poses a disadvantage compared to the RFLP of the *psbA-trnH* marker, where a result was observed in most tested samples (compare **Figure 3.18** and **Figure 3.20**). Consequently, the EOMT marker, even though shows clear results for reference plants, is misleading and does not allow to identify *O. tenuiflorum* commercial samples. Therefore, in order to improve the DNA barcoding assay, the trait-independent *psbA-trnH* marker was tested together with the trait-related EOMT marker in a multiplex PCR.



**Figure 3.20.** Discrimination of *Ocimum tenuiflorum* based on the trait-related eugenol O-methyltransferase (EOMT) using a duplex ARMS strategy. Left: Predicted pattern for *O. tenuiflorum*. Right: Representative gel showing amplification results for the trait-related marker EOMT for commercial Basil products and Mint as outgroup. TulComm.0001: Tea mixture with different Tulsi types; TulComm.0002: Tea mixture with different Tulsi types; TulComm.0003: Tulsi tea; TulComm.0004, TulComm.0005, TulComm.0006 and TulComm.0007: different brands of Tulsi tea mixed with other herbs; BasComm.0001: Dried basil; BasComm.0002: Dried basil; MenComm.0001: Mint tea. For details on samples declarations see **Table 2.4** (Ríos-Rodríguez *et al.*, 2021).

### 3.4.3 Multiplexing trait-independent marker (*psbA-trnH*) and trait-related marker (EOMT') allows for *O. tenuiflorum* discrimination in a one-step protocol

The multiplex PCR assay consisted in a single PCR reaction having the trait-independent *psbA-trnH* marker and a modified version of the previous trait-related marker primers (EOMTfw' and EOMTrev'), ensuring that both pair of primers had a similar annealing temperature. Reference plants were then analysed in order to validate the assay (**Figure 3.21**). All reference material had the expected *psbA-trnH* marker amplicon at around 400 bp. In addition, a lower band at around 200 bp, corresponding to the predicted size of the diagnostic EOMT' band, was observed in the five *O. tenuiflorum* accessions and in *O. gratissimum*. A faint diagnostic band was also observed in *O. campechianum*. The outgroup *M. x piperita*, supposedly not a ME producer species, only yielded a band at around 400 bp, which represents the *psbA-trnH* marker reporting a successful reaction. Hence, the multiplex PCR results were consistent with the RFLP results (**Figure 3.18**) and with the EOMT - ARMS results (**Figure 3.20**), showing once again that *O. tenuiflorum*, *O. gratissimum*, and *O. campechianum* present a similar pattern which is different from the other Basils, especially *O. basilicum*.

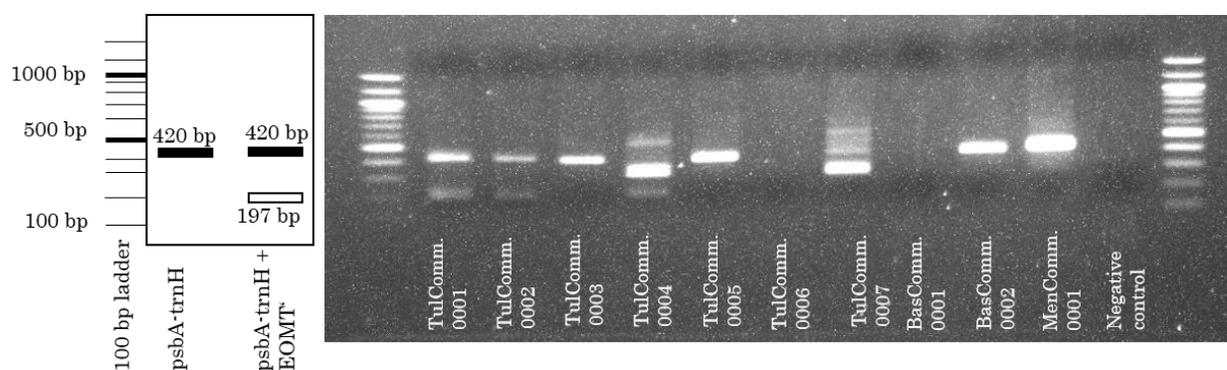


**Figure 3.21.** Discrimination of *Ocimum tenuiflorum* based on the diagnostic ARMS primer for the trait-related eugenol *O*-methyltransferase (EOMT) and the trait-independent markers *psbA-trnH*. a Predicted pattern for *O. tenuiflorum*. b Representative gel showing amplicates obtained for a core-collection of validated reference plants for *Ocimum* along with *Mentha x piperita* as outgroup (Ríos-Rodríguez *et al.*, 2021).

The DNA barcoding method was later tested in commercial samples (**Figure 3.22**). As a result, the *psbA-trnH* amplicon (at around 400 bp) was observed in most samples, except TulComm.0006 and BasComm.0001. Sample TulComm.0004, allegedly containing Tulsi mixed with other herbs, had one strong band considerably smaller than 400 bp, and a weaker band larger than 400 bp. These might be a result of overlay bands from the other undeclared herbs. However, sample Tulcomm.0007 presented a

similar situation and the expected *psbA-trnH* band at around 400 bp is well observed. Samples TulComm.0001 and TulComm.0002 had the diagnostic band at around 200 bp, indicating the presence of *O. tenuiflorum*, and matching the RFLP results (Figure 3.18). Hence, these two samples apparently contain *O. tenuiflorum*. Sample TulComm.0004 also presented the diagnostic band, which in this case was inconsistent with the RFLP result (Figure 3.18). Perhaps due to the presence of other herbs that are ME producers in the samples. In the other samples that were supposed to contain Tulsi, TulComm.0003, TulComm.0005 and TulComm.0007, the diagnostic band was lacking. Similarly, the samples containing Basil and Peppermint were negative for the diagnostic band as expected.

Overall, the results yielded by the trait-independent marker *psbA-trnH* coupled with the trait-related marker EOMT<sup>r</sup> in a multiplex PCR, were consistent with those for the RFLP (Figure 3.18). Moreover, the DNA barcoding assay was completed in a one-step protocol. Therefore, this method is an acceptable alternative for *O. tenuiflorum* identification.



**Figure 3.22.** Discrimination of *Ocimum tenuiflorum* based on the diagnostic ARMS primer for the trait-related eugenol *O*-methyltransferase (EOMT) and the trait-independent markers *psbA-trnH*. Left: Predicted pattern for *O. tenuiflorum*. Right: Representative gel showing amplicons of the trait-related EOMT and the trait-independent marker *psbA-trnH* for commercial Tulsi and Basil products, using Mint as outgroup. TulComm.0001: Tea mixture with different Tulsi types; TulComm.0002: Tea mixture with different Tulsi types; TulComm.0003: Tulsi tea; TulComm.0004, TulComm.0005, TulComm.0006 and TulComm.0007: different brands of Tulsi tea mixed with other herbs; BasComm.0001: Dried basil; BasComm.0002: Dried basil; MenComm.0001: Mint tea. For details on samples declarations see **Table 2.4** (Ríos-Rodríguez *et al.*, 2021).

#### 3.4.4 Commercial samples chemical profile highlights the specificity of the trait-related marker

*Ocimum* species present variable amounts of methyleugenol (**Table 3.3**), therefore the eugenol *O*-methyltransferase from *Ocimum tenuiflorum* was used to design the trait-related marker. However, a potential relationship between the trait-related marker and the methyleugenol content was not confirmed. Further, in commercial samples, an added variable influencing the ME content is the plant material processing, which can promote metabolites degradation. Therefore, in order to find out a possible association between the DNA-barcode and the methyleugenol content, the chemical profile of commercial samples was analysed (**Table 3.4**). All commercial samples that declared to have Tulsi contained ME, though only TulComm.0001 and TulComm.0002 had more than trace amounts (**Table 3.4**). The same samples were positive for the trait-related EOMT marker used in the multiplex PCR assay and in the RFLP assay. The previous results suggest that the trait-related EOMT marker can identify *Ocimum tenuiflorum* even in dried processed mixtures of herbs. The results showed as well that samples containing undeclared Basil species (BasComm.0001 and BasComm.0002) had also substantial amounts of ME, although these were negative for the trait-related EOMT marker and the RFLP. Hence, the presence of the EOMT diagnostic band in a sample does not automatically mean that the sample contains ME. On the other hand, samples lacking the EOMT diagnostic band are not guaranteed to be void of ME.

In addition, the eugenol content was also measured in order to better understand the occurrence of ME in the samples. From all commercial samples only TulComm.0001, TulComm.0002 and TulComm.0004 had more than trace amounts (**Table 3.4**). As mentioned above, these three samples declared to have Tulsi and were positive for the trait-related EOMT marker.

**Table 3.4.** Methyleugenol (ME) and Eugenol (EU) content in commercial samples extracts detected by Gas Chromatography. For the details of the declarations on the samples refer to **Table 2.4** (Ríos-Rodríguez *et al.*, 2021).

<b>Commercial Samples</b>	<b>ME Content *</b>	<b>EU Content *</b>
TulComm.0001	11.58 %	21.10 %
TulComm.0002	10.20 %	30.10 %
TulComm.0003	trace amounts	trace amounts
TulComm.0004	trace amounts	14.15 %
TulComm.0005	trace amounts	trace amounts
TulComm.0006	trace amounts	trace amounts
TulComm.0007	trace amounts	trace amounts
BasComm.0001	5.36 %	trace amounts
BasComm.0002	6.66 %	trace amounts
MenComm.0001	(-)	(-)

\* Content: percentage from total chromatogram area. (-) compound not present in the sample.

### 3.5 Results summary

#### Eugenol *O*-methyltransferase (EOMT) in the genus *Ocimum*

- EOMT phylogeny discloses two distinctive groups in *Ocimum* sp. that were categorized in two groups, the Tulsi clade (*O. tenuiflorum* and *O. gratissimum*) and the Basil clade (*O. basilicum*, *O. americanum*, *O. africanum* and *O. kilimandscharicum*).
- *O. gratissimum* belonging to the Tulsi clade, is the only accession in the clade that does not produce methyleugenol under the given growing conditions. And none of the Basil clade accessions produces methyleugenol under the given growing conditions.
- Similarities in the amino acid sequence of EOMT between *O. gratissimum* and *O. basilicum* are hypothesised to be relevant for the synthesis of methyleugenol in the genus *Ocimum*.

#### Phenylpropanoid Pathway in *O. tenuiflorum*

- Two *O. tenuiflorum* chemotypes (Krishna and Rama Tulsi) show different regulation mechanisms of the phenylpropanoid pathways for the synthesis of methyleugenol, when using UV-B as stress factor. These results are supported by gene expression and the chemical profile. The upstream reactions to the synthesis of methyleugenol in Rama Tulsi seems to be highly regulated by the enzyme caffeic acid *O*-methyltransferase (COMT). Whereas in Krishna Tulsi, the enzyme caffeoyl-CoA *O*-methyltransferase (CCOMT) seems to also play an important role.
- In addition, given the low transcript levels of PAL and C4H, it is hypothesised that besides phenylalanine, tyrosine might also be a relevant first substrate of the pathway.

#### Subcellular localization of *O. tenuiflorum* EOMT

- The subcellular localization of EOMT obtained from Krishna Tulsi was found to be in the nucleus and the cytosol of transformed BY-2 cells.
- Western-blot analysis shows that EOMT fused with GFP was positively expressed in transformed BY-2 cells, however, more analyses are needed in order to confirm that free GFP is absent.

#### DNA barcoding for *O. tenuiflorum* discrimination

- A one-reaction assay using a trait-independent marker (*psbA-trnH* *igs*) together with a trait-related marker based on the enzyme eugenol *O*-methyltransferase (EOMT') is able to discriminate *O. tenuiflorum* from other *Ocimum* sp. in reference plants and in commercial samples. The EOMT' marker is not related to methyleugenol occurrence in the samples.

## 4 Discussion

### 4.1 EOMT variability in the genus *Ocimum*

Nucleotide and amino acid sequences of the enzyme eugenol *O*-methyltransferase (EOMT) obtained from different *Ocimum* sp. have been analysed and four main results have been found. First, there are variants of the EOMT gene in several *Ocimum* sp. encoding the functional enzyme and pseudogenes. Second, EOMT-based phylogeny discloses two distinctive groups in *Ocimum* sp. that were categorized in two groups, the Tulsi clade (*O. tenuiflorum* and *O. gratissimum*) and the Basil clade (*O. basilicum*, *O. americanum*, *O. africanum* and *O. kilimandscharicum*). Third, *O. gratissimum* belonging to the Tulsi clade, is the only accession on the clade that does not produce methyleugenol, similar to the Basil clade accessions. And fourth, similarities on the amino acid residues between *O. gratissimum* and the Basil clade accessions, both non-ME producers, are potentially relevant for the synthesis of methyleugenol in the genus *Ocimum*.

#### 4.1.1 EOMT gene variability among and within *Ocimum* species

Previous phylogeny work has already identified clades among *Ocimum* species (Jürges *et al.*, 2018). By using four trait-independent barcodes, *matK*, *psbA-trnH*, *trnL-F* and *rbcL*, four haplotypes have been determined in a selected group of *Ocimum* species: haplotype I (*O. basilicum*, *O. americanum*, *O. x africanum* and *O. kilimandscharicum*), haplotype II (*O. tenuiflorum* and *O. campechianum*), haplotype III (*O. gratissimum*), and haplotype IV (*O. filamentosum*). The mentioned DNA barcodes are recommended for being universal, have power of discrimination and present a sequence quality that would allow identification (Hollingsworth, 2011; Hollingsworth *et al.*, 2009; Hollingsworth *et al.*, 2011; CBOL Plant Working Group, 2009). However, not all the mentioned DNA barcodes had the desired characteristics when analysing *Ocimum* species. For example, when comparing the discrimination power of two plastidic markers, *psbA-trnH* *igs* had a higher discrimination power than *rbcL*. (Jürges *et al.*, 2018).

When using the sequence of EOMT -a nuclear trait-related marker- for phylogeny, two distinctive clades were found, the Tulsi clade, comprising haplotype II and III, and the Basil clade, comprising haplotype I. For the Tulsi clade, the two haplotypes are not completely branched together, which is expected for these two distinctive groups. However, in the Basil clade, the haplotype I designated by means of trait-independent markers -nuclear and plastidic-, *O. basilicum* and *O. kilimandscharicum*, have been found to be branched together (Jürges *et al.*, 2018). Contrary to this, when using the sequence of EOMT, they are belonging to the same

clade but branched apart. This highly suggests that EOMT has a greater power of discrimination than the rest of the nuclear and plastidic markers (trait-independent) previously used. In addition, the sequences of the EOMT for *O. kilimandscharicum* resulted to be pseudo-sequences, which might further explain the greater power of discrimination of EOMT as a barcode. These pseudo-sequences were not exclusively found in *O. kilimandscharicum*, but in most *Ocimum* sp. accessions.

Pseudogenes of the putative EOMT obtained from genomic DNA were found in accessions of *O. tenuiflorum*, *O. basilicum*, *O. x africanum* and *O. kilimandscharicum*. Pseudogenes are non-functional duplicates of a gene, they are originated by gene duplication which serves an evolutionary purpose of generating genomic diversity (Panchy *et al.*, 2016). However, duplication of genes can be redundant and it is counteracted with pseudogenization achieved by different mutation mechanisms (Panchy *et al.*, 2016). In this research, insertions, deletions and point mutations have been found to be the mechanisms for pseudogenes generation of the putative EOMT in *Ocimum* species. Moreover, interesting patterns for these mechanisms were found. For instance, deletions/insertions were present in the EOMT sequences from the Basil clade (haplotype I), whereas point mutations to produce a stop codon was the only pseudogenization mechanism found in all the EOMT sequences from *O. tenuiflorum* accessions (haplotype II).

In this frame, in most of the *Ocimum* accessions one or two copies of the EOMT gene were found in the gDNA, being one of them the pseudogene (**Table 3.1**). However, *O. tenuiflorum* ID 5751 was an exception, where 3 copies of the gene (two pseudogenes) were found from gDNA. This event could be explained by what other authors have already suggested, the hypothesis of polyploidization in the genus *Ocimum* (Carović-Stanko *et al.*, 2010; Pyne *et al.*, 2018). Allopolyploidy has been suggested by studies on *O. basilicum*, *O. americanum* and their hybrid *O. x africanum* in the basis of chromosome duplication (Carović-Stanko *et al.*, 2010; Pyne *et al.*, 2018). However, polyploidy in *O. tenuiflorum* has been more difficult to explain. One of the reasons is the difficulty to establish the chromosome number (2n) and the basic chromosome number (x) in *O. tenuiflorum*. Reports of the chromosome number for *O. tenuiflorum* have indicate specimens with 2n equal to 16, 32, 36, and 76 (Carović-Stanko *et al.*, 2010; Pyne *et al.*, 2018; Rastogi *et al.*, 2014). These difficulties on obtaining the chromosome number have been granted not only to the small size of the chromosomes that makes it hard to count them, but it has also been suggested the existence of different genotypes/cytotypes within the species (Carović-Stanko *et al.*, 2010). Further experiments crossing different genotypes/cytotypes individuals of *O. tenuiflorum* resulting in chromosome duplication would be needed in order to better understand potential polyploidy mechanisms in Tulsi. However, this approach might not completely explain the even greater gene variation for EOMT found in the cDNA.

Particularly in *O. basilicum* and *O. tenuiflorum* accessions, some EOMT gene variants were found in the cDNA but not in the gDNA (**Table 3.1**). One reason for this could lie on the limitation of the cloning procedure. Besides amplifying the EOMT, the primers used were likely to amplify other OMT. Therefore, cloning was used in order to obtain single sequences. This technical condition might have contributed to have a limited number of EOMT sequences cloned into a vector and positively transform *E. coli*. However, it could also be that finding EOMT sequences in the cDNA, that were not present in the gDNA, was a result of the presence of isoforms.

Gene isoforms are a variant form of a gene which are either encoded in the genomic DNA, or are generated in the transcription processes, derived from a single gene that has undergone alternative splicing, resulting in protein isoforms (Buchanan *et al.*, 2015; GEP, 2015). In the present study, EOMT gene variants found in the cDNA of *Ocimum* sp. and not in the gDNA, could have been generated from a single gene. Under the assumption that this isoforms theory is a possibility, why would the plant produce them? A potential reason for these mutations to be useful for the plant is the promiscuous nature described for OMT. For example, as shown by Gang *et al.* (2002a), the EOMT from *O. basilicum* is an enzyme capable of using chavicol, eugenol, *t*-isoeugenol and caffeic acid as substrates. A similar case happens with a *t*-isoeugenol *O*-methyltransferase found in Anise, that was also capable of methylating *t*-anol to form *t*-anethole (Koeduka *et al.*, 2009). Therefore, the different EOMT isoforms might be responsible of methylating different substrates or having different catalytic reaction parameters. Particularly, because even though the differences between the isoforms found in this research are based on point mutations only, these change the *aa* profile as revised in the next section. Moreover, how these isoforms are formed and whether alternative splicing occurs or not, remains unknown, although the introns participation in this process is a plausible option.

Introns variability was found in EOMT sequences among *Ocimum* sp. These variabilities were enough to rearrange the accessions in the EOMT-based phylogeny studies within *Ocimum* clades (compare **Figure 3.1** and **3.2**). However, the reasons for these differences remain unclear. In addition to these nucleotide variabilities, there was an intron positional conservation observed at around the base pair 760 (section **3.1.1**). This intron position lies in the *O*-methyltransferase domain between motifs III and IV (see appendix **Table A.2** and **Figure A.3**). Intron positional conservation is believed to safeguard a particular function of an intron, since chances to be eliminated will be reduced if the location is conserved (Chorev and Carmel, 2012). The role of introns, which initially were labelled as junk DNA, has gradually been investigated and elucidated. Introns have been suggested to serve different functions in transcription regulation such as initiation, termination and splicing, also they might participate in localization and nuclear export, either for having signalling peptides or via isoforms formation (Chorev and Carmel, 2012). Moreover, it has been

reported that PAL and 4CL isoforms expressed accordingly to the final product of the pathway and to the tissue where the phenylpropanoid pathway will be active (Biała and Jasiński, 2018; Rastogi *et al.*, 2013). At the moment it is unknown, whether or not the variability and positional conservation of introns in EOMT enzyme from *Ocimum* species is responsible for the protein isoforms found or whether the intron has specific functions.

#### 4.1.2 EOMT across the genus *Ocimum*: amino acid residues variability and their potential influence on the chemical profile

The profile of the amino acid residues shows differences among the genus *Ocimum*. First, the EOMT *aa* sequences from *Ocimum* species presented two domains, a dimerization and an *O*-methyltransferase domain. Second, within the *O*-methyltransferase domain, five motifs described for the OMT superfamily were found (Ibrahim *et al.*, 1998). And third, similarities between the *aa* sequences from *O. gratissimum* (from the Tulsi clade) and the accessions of the Basil clade, suggest that these similarities in the *aa* profile are linked to similarities in the chemical profile. However, more research on the enzyme is needed to confirmed the findings.

The protein homology analysis based on several other structures from *O*-methyltransferase, suggested that EOMT had a dimerization domain and an *O*-methyltransferase (see appendix **Figure A.3**). Therefore, it was also suggested that EOMT was a homodimer, which is consistent with studies on the recombinant EOMT enzyme from *O. basilicum* (Gang, Lavid, *et al.*, 2002). Further, in the terms of this study, the differences found in the *aa* sequences were located in the *O*-methyltransferase domain and, thus, might affect the specificity of the enzyme but not the dimerization process.

Within the EOMT *O*-methyltransferase domain five motifs described for the OMT superfamily were found, but these were not quite identical to the original suggested motifs (Ibrahim *et al.*, 1998). A possible explanation for this outcome, is that the research done by Ibrahim *et al.* (1998) took place more than two decades ago, when not so many OMT sequences were available in the data base used. Interestingly, the motifs are quite similar among *Ocimum* species. Furthermore, *O. gratissimum* accessions, which does not produce ME in the given growing conditions, have key similarities on the amino acidic profile with sequences from the Basil clade accessions, also non-ME producers in the given conditions (see **Figure 3.5** and appendix **Table A.2**). Based on this results, the hypothesis is that the amino acids that are similar between the EOMT sequences from *O. gratissimum* and the EOMT sequences from Basil clade accessions, might have an influence on determining whether a plant will be a ME producer or not. It is important to stress that the labelled species as ‘non-ME’ producers, do not produce ME under the given growing

conditions in the glasshouse, therefore not influenced by UV radiation (which influences phenylpropenes synthesis). Therefore, hypothetically, the *aa* sequences differences are expected to influence the substrate specificity/product formation, but not completely halt the EOMT catalytic reaction.

Slight differences in the chemical profile among *O. tenuiflorum* accessions were also found. The results show that Rama Tulsi produces more ME (around 90% of the extract) than the other *O. tenuiflorum* (between 50 and 70% of the extract) (**Figure 3.5**). These differences could be due to subtle differences in the amino acid profile from the enzyme found in Rama compared to Krishna. However, Krishna is a different chemotype that produces other compounds in higher quantities such as anthocyanins, and in this case, higher amounts of eugenol and  $\beta$ -caryophyllene (Jürges *et al.*, 2018). Therefore, either the enzyme present in Rama is more efficient in converting eugenol to methyleugenol, or the regulation mechanisms in Krishna are such that eugenol availability is desired and, therefore, it is not fully converted to methyleugenol. Here, the recombinant studies on all of the isoforms of the EOMT found in *O. tenuiflorum* would be helpful on giving an insight on these open questions.

#### 4.1.3 EOMT beyond the genus *Ocimum*: OMT-based phylogeny

A phylogeny analysis including OMT from other species than *Ocimum* was done (**Figure 3.6**). In this analysis the Tulsi and Basil clades were also distinctively found. Further, the Basil clade had sequences from EOMT and chavicol *O*-methyltransferase (CVOMT). Studies on the recombinant EOMT from *O. basilicum*, showed high similarities between EOMT and CVOMT (Gang *et al.*, 2002). Hence, these two enzymes from *O. basilicum* are closer related than to the EOMTs from the Tulsi clade. Moreover, the OMT-based phylogeny shows that the EOMT from *Ocimum* species seemed to derive from flavonoid *O*-methyltransferases. This matches with theories that suggest that land plants first developed the flavonoid pathway for adaptation purposes such as defence against UV radiation and later on, the phenylpropanoid pathway was developed (Emiliani *et al.*, 2009).

In addition, the OMT-based phylogenetic tree showed that OMT sequences from the order Lamiales were clustered together. These results contradict what others have found, where plants from other orders are clustered together with Lamiales (Gang, 2005). The findings in this research suggest that the OMT from the Lamiales order might have a common ancestor, and at some point it differentiated from the OMT from other orders. This could partially explain differences in the promiscuity profile of OMTs from different orders. For example, an isoeugenol *O*-methyltransferase from Anise (*Pimpinella anisum*, family Apiaceae, order Apiales) can use *t*-isoeugenol and *t*-anol as substrate, whereas an EOMT from Basil (*Ocimum basilicum*, family Lamiaceae, order Lamiales) can use eugenol, *t*-isoeugenol, chavicol and caffeic acid

as substrates (Gang, *et al.*, 2002; Koeduka *et al.*, 2009). Yet, more interrogates arise from these findings such as how the phylogeny could explain the occurrence of methyleugenol in the different species. ME can be found in several plants, for instance in the order Rosales, where none of the OMT was clustered together with the OMTs from Lamiales (Tan and Nishida, 2012). Perhaps, chemical profile analyses along with phylogeny, both including species of different orders could help to better comprehend the clades arrangement here presented. Moreover, the OMT clusters might reflect how the phenylpropanoid pathway evolved in different directions depending on the plant order. Therefore, knowledge on the phenylpropanoid pathway as a whole and species-specific, it would be crucial for understanding the results here presented.

#### 4.2 Understandings on the Phenylpropanoid Pathway for the synthesis of Methyleugenol in *O. tenuiflorum*

Results of the study on the phenylpropanoid pathway in two *O. tenuiflorum* chemotypes, Krishna Tulsi and Rama Tulsi, suggest that there are differential regulation processes for the synthesis of methyleugenol in both chemotypes.

For defining the phenylpropanoid pathway for the synthesis of methyleugenol in *O. tenuiflorum*, suggested pathways from other researches in the field were revised. This bibliography research had to undergo filters. The first filter was settled by the end product. The final metabolites of the phenylpropanoid pathway are numerous, such as the production of lignin, phenols, flavonoids, carotenoids, to name a few (Vogt, 2010; Rastogi *et al.*, 2013; Buchanan *et al.*, 2015). Thus, in order to study the phenylpropanoid pathway in this study, an assumption on the compounds participating only for the formation of methyleugenol was done. This is linked to the assumption that the enzymes here analysed produced the desired metabolite. However, they might have been participating in other reactions.

The second filter was narrowing the plant genus. Plants use the phenylpropanoid pathway for the synthesis of lignin and flavonoids forming part of the plant physiology and as a defence mechanism (Vogt, 2010; Buchanan *et al.*, 2015). However, not all plants have the enzymatic machinery to produce all of the possible secondary metabolites that are originated from phenylalanine/tyrosine. Therefore, because the phenylpropanoid pathway varies among species, only information related to the phenylpropanoid pathway described for the genus *Ocimum* was considered.

The last filter, settled by the literature findings on the genus *Ocimum*, was the start substrate. The phenylpropanoid pathway has been described to start normally with the amino acid phenylalanine, therefore the name, or in some cases even in parallel with the amino acid tyrosine (Emiliani *et al.*, 2009; Vogt, 2010). Tyrosine is catalysed

by the enzyme tyrosine ammonia lyase (TAL), to synthesise *p*-coumaric acid (see **Figure 3.14**). However, in all of the revised bibliography, there were no suggestions for the participation of tyrosine in the phenylpropanoid pathway in *Ocimum* sp. as starting substrate. This led to not include TAL in the selected enzymes evaluated in this study.

Having settled the filters, the phenylpropanoid pathway for *Ocimum* sp. was found to contain diverse paths -within it- for the synthesis of ME: via caffeoyl-CoA and via ferulic acid (see section 1.3.5 and **Figure 1.5**). Out of these, the path via caffeoyl-CoA would be considered the 'primary pathway' because its precursor, *p*-coumaroyl-CoA, has been described as the principal branching point for the synthesis of a variety of phenylpropenes (Vogt, 2010). Unlike in this and other research, these paths together were not considered in all studies. Thus, different studies on the genus *Ocimum* had focused in slightly different phenylpropanoid pathways for the synthesis of methyleugenol. The results of this research suggest that, these both paths -within the phenylpropanoid pathway- are indeed most likely occurring in *Ocimum* sp. for the synthesis of methyleugenol (see section 3.2.5). Moreover, when comparing two *O. tenuiflorum* chemotypes (Krishna and Rama Tulsi), the enzymes transcripts' levels suggest that different regulation mechanisms on the phenylpropanoid pathway occur for each chemotype.

#### Phenylpropanoid pathway in *O. tenuiflorum*: the basis

The steady state of the enzymes analysed revealed ground differences between the two chemotypes, Krishna and Rama, at the transcript level and the metabolite level. Absolutely all of the six enzymes transcripts' levels in Krishna were higher than in Rama, and four out of six were significantly higher (see **Figure 3.7** and **Figure 3.8**). Interestingly, these differences at the transcript level were not precisely found at the metabolite level (see **Figure 3.9**). The eugenol concentration did match the over expression of EGS in Krishna compared to Rama. However, the concentration of methyleugenol in Rama was significantly higher than in Krishna. This could suggest a predisposition of Rama to produce more ME, via metabolic control, or perhaps the EOMT in Rama could be more efficient than in Krishna. While different isoforms of the EOMT were found from gDNA and cDNA in Krishna, only one isoform was found in Rama (see appendix **Table A.2**). This could be an advantage for Rama on producing methyleugenol. At aside these differences, there were also similarities.

Particularly interesting, was the higher (nearly double) transcript level of COMT over CCOMT in both chemotypes (see **Figure 3.8**). This suggests that COMT might have rather important role on providing substrates for the phenylpropanoid pathway, though it is important to point out that these enzymes do not only participate in the synthesis of intermediates for phenylpropenes, but also for lignin. Nonetheless, the COMT/CCOMT ratio is interesting because, as stated above, the pathway via

caffeoyl-CoA, catalyzed by CCOMT, would be considered the 'primary pathway' due to the importance of its precursor, *p*-coumaroyl-CoA. Thus, studies on the phenylpropanoid pathway for the synthesis of methyleugenol should consider both enzymes, which so far has been neglected by several -not all- studies.

Another consideration, that might also have been neglected in *Ocimum* research, is the potential role of tyrosine. The transcription levels of both, PAL and C4H, are lower than the transcription levels of COMT, CCOMT, and EGS, in Krishna and Rama Tulsi. If phenylalanine were the principal amino acid as starting point on the phenylpropanoid pathway, the levels of PAL and C4H transcript should be higher in order to provide enough substrate for the other enzymes. But, that was not the case. Therefore, based on these findings, it is suggested that tyrosine -and the enzyme TAL- could be playing an important role for the synthesis of phenylpropenes in *O. tenuiflorum*.

#### Phenylpropanoid pathway in *O. tenuiflorum*: insights on its regulation

Analysis of control samples versus samples taken immediately after being exposed to 10 min of high UV-B radiation levels suggest that there are regulations of the transcripts' levels in the two Tulsi chemotypes. Most enzymes showed an increment in the transcript level after 10 min of UV-B exposure. Particularly, PAL in Krishna shows significant increase in the transcript level. Research in maize has shown that several transcripts are indeed upregulated after 10 min of UV-B exposure, however, these were mostly suggested to be from transcription factors and signal related proteins (Casati *et al.*, 2011). In this research, the transcripts are from phenylpropanoid pathway enzymes, that in general are found to be later increased, depending on the plant species and the stress treatment (Casati *et al.*, 2011; Oravec *et al.*, 2006). Before studies linked the UVR8 receptor to UV-B response in plants (see section 4.2.1), research had been done in early UV-B response. Particularly, parsley cell lines containing a vector with the promoter of CHS ligated to a luciferase gene, as reporting protein, demonstrated that as early as 5 min after UV-B radiation the promoter was activated (Frohnmeier *et al.*, 1999). It is suggested that this fast response is triggered by changes on the concentration of free calcium in the cytosol ( $[Ca^{2+}]_{cyt}$ ), which increases under UV-B stress and might play a transient role in the early regulation of genes (Frohnmeier *et al.*, 1999). Moreover,  $[Ca^{2+}]_{cyt}$  has been suggested to promote PAL transcripts, however, not in the frame of UV-B experiments (Lecourieux *et al.*, 2002; Lecourieux *et al.*, 2006). Therefore, one possibility for the fast transcript level increase, could be mRNA synthesis of phenylpropanoid enzymes *via* promoter activation with  $[Ca^{2+}]_{cyt}$  playing a role on the process. This suggests a mechanism of fast transient response at the transcriptional level. However, post-transcriptional/translational regulations affecting the transcript levels are likely to play a role as well. For instance, a potential regulation could be

the hinder of mRNA degradation allowing transcripts to accumulate. mRNA degradation can be achieved by micro RNA (miRNA), which are small non-coding RNA molecules involved in several metabolic activities (Buchanan *et al.*, 2015). A variety of miRNA have been found to be influenced by UV-B and linked to the regulation of phenylpropanoid pathway enzymes' transcripts (Sunitha *et al.*, 2019; Yanjun Yang *et al.*, 2020). However, no study has looked/found early miRNA responses linked to PAL. Further, the fast response found in this study, suggesting particular kinetics of transcript accumulation in Tulsi, are attributed to UV-B as stressor as one factor. However, the high intensity of the radiation used could also be considered a fast response promoting factor. Studies applying different UV-B intensity as stress factor have shown that higher levels of UV-B result in higher levels of relative gene expression of several gene transcripts, including PAL (Contreras *et al.*, 2019). Overall, the accumulation of transcripts from enzymes of the phenylpropanoid pathway as a fast response, after a 10 min high dose UV-B radiation treatment, is not yet understood. However, a transcript accumulation response one hour after 10 min of UV-B radiation treatment, can be attributed to signal transduction and transcriptional regulation mechanisms.

In this study, the response of genes to UV-B stress was also measured in samples taken one hour after 10 min of high doses of UV-B treatment. All of the six enzymes evaluated showed differences at the transcript level one hour after the radiation (see **Figure 3.10** and **Figure 3.11**). Similarly, several early responsive genes to UV-B stress have been identified in *Arabidopsis*, among them PAL, C4H, 4CL and chalcone synthase (CHS) from the phenylpropanoid pathway (Oravecz *et al.*, 2006; Ulm *et al.*, 2004). These genes were responsive 45 min after a 15 min UV-B treatment (Oravecz *et al.*, 2006). Furthermore, in *O. tenuiflorum*, the transcripts' regulation was found to be different in both genotypes. In Krishna Tulsi PAL was upregulated, while CCOMT, EGS and EOMT were down-regulated. In Rama PAL, C4H, COMT, CCOMT, EGS and EOMT were up-regulated. These differences might have its basis on the fact that these two are indeed different chemotypes.

When grown under natural conditions, Krishna Tulsi produces such amount of anthocyanins that the colour of the leaves turn purple at early stage, whereas Rama Tulsi leaves do not, the leaves are green (Jürges *et al.*, 2018). The pathway leading to the formation of anthocyanins, has its start in the phenylpropanoid pathway via *p*-coumaroyl-CoA (Emiliani *et al.*, 2009; Vogt, 2010). Hence, it makes sense that the enzymatic machinery from Krishna Tulsi is prepared and adapted to favour the synthesis of secondary metabolites via caffeoyl-CoA whose precursor is *p*-coumaroyl-CoA, as suggested by the results of this research. On the contrary, as mentioned above, Rama Tulsi is not reported to be an anthocyanin compounds producer, as Krishna Tulsi is. Therefore, it is plausible that Rama Tulsi produces these phenylpropenes also via ferulic acid, perhaps optimizing the same synthesis route of

lignin synthesis, using part of the same enzymatic machinery (Buchanan *et al.*, 2015). However, no causality can be assigned for the increase of COMT/CCOMT on the transcript levels of EGS or EOMT, because in order to have post-translation feedback, a greater time frame should have been considered to measure gene expression. It could be, though, a regulation feedback took place at the transcript level.

The anthocyanin and the volatiles phenylpropenes content support the gene expression data. Tendencies for the increase and decrease of metabolites content were clearly observed, though not all of them are significant. The absence of significant differences in the metabolic profile could be for different reasons. The most likely reason is the time period of the metabolism response between the signal reception and the metabolite synthesis, although the actual response time is so far unknown from UV-B reception to methyleugenol synthesis (Buchanan *et al.*, 2015). As example of the time frame for having significant changes in the chemical profile, research in the UV-B radiation in *O. basilicum* has produced significant changes in the chemical profile after 2 weeks of treatment (Johnson *et al.*, 1999). In this research, it was clear that for Krishna Tulsi the synthesis of anthocyanins required at least 2 hours after 10 min exposure to high doses of UV-B radiation, and equal frame period for diminishing the methyleugenol synthesis. The enzyme CHS, a key enzyme in the synthesis of anthocyanins, was found to accumulate between 1 to 2 hours after continuous lower doses of UV-B radiation (Oravec *et al.*, 2006). Therefore, finding an increment on the content of anthocyanin in Krishna can be attributed as a direct response of UV-B stress and gene expression. Contrary to this, in Rama Tulsi, a greater increase in the content of eugenol was observed 2 hours after the UV-B treatment. The preference of producing greater amounts of eugenol compared to methyleugenol could be based on the volatile properties. Eugenol is less volatile than methyleugenol, therefore better as protection inside the cells against UV-B.

Within this section it was discussed the existence of a differential regulation on the phenylpropanoid pathway enzymes for two *O. tenuiflorum* chemotypes. But, how does this regulation take place from the UV-B signal reception to the expression of the gene? This will be revised in the next section.

#### 4.2.1 Insights on UV-B stress and the Regulation of the Phenylpropanoid Pathway

Plants conquering land had to evolve and, as part of that evolution, the phenylpropanoid pathway developed as a defence mechanisms against microorganisms, predators and UV irradiation (Emiliani *et al.*, 2009; Stratmann, 2003). The UV irradiation can cause mutations, for instance, dimers formation by pyrimidine fusion, which can be repaired through photolyases triggered via blue light (Buchanan *et al.*, 2015). This is a mechanism to maintain the DNA integrity, but more than one mechanism is needed for protection. Therefore, several metabolites play an important role against UV radiation. Flavonoids, anthocyanins, and lignin are

believed to be the first developed as defence mechanism. Particularly, UV absorbing compounds are important for protection against UV radiation, these can be compounds such as phenols and terpenes (Vanhaelewyn *et al.*, 2020). Among these, aromatic phenolic compounds have a benzene ring which absorbs UV radiation and thus, they provide protection against UV absorption by DNA, for example (Vanhaelewyn *et al.*, 2020). However, which compounds are to be produced by plants and through which pathway, is unknown since each plant species has its own particular metabolism.

In general, the pathway from the UV perception by the plant to the synthesis of phenylpropanoids as defence mechanism, is poorly understood so far. Several studies have efforted on elucidating first, the perception of UV by plants, followed by the signal transduction to gene expression; and second, from the gene expression to the physiological/chemical effects (Biever and Gardner, 2016; Müller-Xing *et al.*, 2014; Yin and Ulm, 2017). On one side, the perception of UV-B has been proposed to be by DNA, which is damaged and can go through repair, or not be repaired causing cell cycle arrest and eventually activate several signalling pathways (Biever and Gardner, 2016; Müller-Xing *et al.*, 2014). Another option for the UV-B perception by the plant is through photoreceptors. One photoreceptor for UV-B, the UV Resistance Locus 8 (UVR8) receptor, was discovered in *Arabidopsis* and related to signalling and regulation of the UV-B response more than a decade ago, and later in other plants (Brown *et al.*, 2005; Rizzini *et al.*, 2011; Tossi *et al.*, 2019). Since then, the understanding on the mechanisms behind the signal perception has advanced.

The UVR8 photoreceptor is a dimer protein that when absorbing UV-B radiation, becomes a monomer, the active monomer interacts and forms complexes with different proteins involved in metabolism regulation such as: Constitutively Photomorphogenic 1 (COP1), Myeloblastosis (MYB73/77), and WRKY (WRKY36) (Liang *et al.*, 2019; Rizzini *et al.*, 2011; Yu Yang *et al.*, 2020). These UVR8-monomer+protein complexes are suggested to directly regulate gene transcription (promotion/inhibition), or -particularly the complex with COP1- activates transcription factors (TF), such as Elongated Hypocotyl 5 (HY5), a bZIP family TF which is suggested to regulate about 4000 different genes (Bhatnagar *et al.*, 2020; Rizzini *et al.*, 2011; Ulm *et al.*, 2004). However, whether HY5 or other TF might participate in the synthesis of phenylpropenes in *O. tenuiflorum*, will be discussed.

Transcriptions factors such as MYB family TF and bZIP family TF have been categorized as UV-B-responsive early genes (Ulm *et al.*, 2004). Moreover, *Arabidopsis cop1* mutants and *hy5* mutants show decrease in a few MYB TF (MYB12, MYB13, and MYB111) known to promote the phenylpropanoid/flavonoid pathway (Dubos *et al.*, 2010; Oravec *et al.*, 2006; Pandey *et al.*, 2014). These findings suggest that not only the UVR8-COP1 complex could directly regulate MYB TF, but because the *hy5* mutants also show a decrease in the expression of MYB TF, this suggests an

interaction between HY5 TF and MYB TF. Moreover, the promoter region of a protein could have elements that are recognized by more than one type of TF. For example, the enzyme chalcone synthase (CHS), key for the synthesis of flavonoids, has a promoter region that can be regulated by MYB TF and HY5 (Ang *et al.*, 1998; Dubos *et al.*, 2010).

In this research, the enzymes PAL, C4H, and COMT were up-regulated due to UV-B stress in both chemotypes. It could be that HY5 was promoting the transcripts increase in these enzymes. *Arabidopsis* mutants *cop1* and *hy5* show a decrease in the transcript levels of several enzymes related to phenylpropanoid pathway, compared to control plants exposed to UV-B radiation, such as PAL, C4H, 4CL, and cinnamyl-alcohol dehydrogenase (CAD) (Oravecz *et al.*, 2006). It is unclear if the regulation of the mentioned enzymes is exclusively due to promoter regions that recognize directly HY5 or some other interaction via MYB TF. Regarding the enzymes CCOMT, EGS, and EOMT, given the notorious down-regulation of the genes in Krishna and up-regulation in Rama, perhaps these are not directly regulated by the HY5, but by MYB TF as the other enzymes in the phenylpropanoid pathway.

An alternative signalling pathway besides the UV perception by the UVR8 receptor mentioned above, could be triggered by accumulation of reactive oxygen species (ROS). ROS accumulation has been found to happen in the plant, independent from the UVR8 receptor, leading to the activation of different hormone signalling pathways such as salicylic acid, jasmonic acid and ethylene (Müller-Xing *et al.*, 2014; Yokawa *et al.*, 2016). Among these hormones, methyl jasmonate activates MYB TF, regulating enzymes from the phenylpropanoid pathway (Gális *et al.*, 2006; Zhou *et al.*, 2021). In addition, jasmonic acid and methyl jasmonate have been studied for increasing the amount of eugenol and methyleugenol in *Ocimum basilicum* (Kim *et al.*, 2006; Li *et al.*, 2007; Milan *et al.*, 2017; Złotek *et al.*, 2016). The jasmonic acid signaling pathway is activated as response to wounding, and methyleugenol has insecticidal properties (Stratmann, 2003; Tan and Nishida, 2012). Thus it might be possible that ME synthesis is activated via jasmonic acid signalling pathway, which could be also triggered by UV-B radiation. Similar hypotheses have been suggested as an overlapping response mechanism by plants for both, UV-B and wounding (Stratmann, 2003).

The phenylpropanoid pathway regulation mechanism has been discussed from the substrate to the key enzymes that participate on the synthesis of ME, and from the stressor signal to the phenylpropenes synthesis. However, where the metabolic pathway is located in the cell is also important knowledge for understanding its regulation. Therefore, the subcellular localization of EOMT was studied and it is discussed in the next section.

### 4.3 Comprehension of the subcellular localization of the phenylpropanoid pathway and EOMT

Studies in the subcellular localization of EOMT from *O. tenuiflorum* have not been reported so far. Here it has been reported that the subcellular localization of EOMT was found to be in the cytoplasm and the nucleus of BY-2 transformed cells.

A better understanding on the regulation mechanisms of metabolic pathways is achieved when the subcellular localization of the enzymes participating on the pathway is known. The reason is that enzymes' catalytic reactions can be affected by several conditions. Factors such as the solvent, substrate and cofactors availability and interaction, and enzyme-enzyme interaction are directly related with the environment provided within the cell and/or its compartments (Hrazdina and Jensen, 1992). Therefore, to better understand the mechanisms under which ME is produced in Tulsi, the subcellular localization of EOMT was studied.

In general, the phenylpropanoid pathway subcellular localization has been suggested to be in the cytoplasm (Biała and Jasiński, 2018). Though, the enzyme C4H was found to be localized in the endoplasmic reticulum, whereas PAL was in the cytosol but colocalized with C4H, presumably as a manner to optimize the metabolism of secondary compounds (Achnine *et al.*, 2004; Biała and Jasiński, 2018; Dastmalchi *et al.*, 2016). After the synthesis of *p*-coumaric acid by C4H, there are about seven enzymes that participate in the synthesis of intermediates before the synthesis of methyleugenol. Out of these, two enzymes have been localized in the cytosol. First, the cinnamoyl-CoA reductase (CCR; EC 1.2.1.44), an enzyme that takes as substrate *p*-coumaroyl-CoA and feruloyl-CoA, was localized in the cytosol (Kapteyn *et al.*, 2007; Muhlemann *et al.*, 2014). Second, the enzyme eugenol synthase (EGS) from *O. basilicum* has been found to be located in the cytosol (Reddy *et al.*, 2021). The cytosolic subcellular localization of EOMT found in this research is consistent with to the previous enzymes localization. Moreover, the EOMT was found to be active in glandular trichomes of *O. basilicum*, that have been described to be 'highly cytoplasmic' (Gang *et al.*, 2002). However, why EOMT was also localized in the nucleus is not yet understood.

The nuclear localization of EOMT was an unexpected result. However, is not an isolated case that a protein, alleged exclusively cytoplasmic, is also found to be located in the nucleus. For instance,  $\alpha$ - and  $\beta$ -tubulin  $\gamma$ -microtubules subunits- are mainly cytoplasmic that up until recently were found to be located also in the nucleus (Schwarzerová *et al.*, 2019). The research on tubulin found no acknowledged nuclear localization signals (NLS), however, nuclear export sequences (NESs) were found and explored. Similar to tubulin, in EOMT no NLS or any subcellular localization peptides were found using the search tools of several databases. The search for localization peptides included the intron sequences, because nuclear export and cytoplasmic

localization have been reported to be functions of the introns (Chorev and Carmel, 2012), however, in EOMT this was not the case. Contrary to tubulin, NESs were not found in the EOMT sequence from Tulsi. Further, the role of nuclear localization is unknown. Therefore, it would be important to first, corroborate the nuclear localization of EOMT, by for example, improve the western blot protocol. And second, design experiments in order to better understand the role of the EOMT nuclear localization.

As stated in the scope of this study (see section 1.4) having Tulsi as a superfood poses a challenge for understanding the accumulation of the genotoxic compound methyleugenol in *O. tenuiflorum*, and for facing the challenge of food fraud. Therefore, in sections 4.1, 4.2 and 4.3 the focus was on discussing the results related to the synthesis of methyleugenol in *O. tenuiflorum*. In the last section of the discussion, the results related to DNA barcoding for detecting food fraud in Tulsi, will be discussed.

#### 4.4 DNA barcoding for discrimination of *O. tenuiflorum* in commercial samples

Discrimination of *O. tenuiflorum* from other species in order to identify surrogates in commercial samples was done using DNA barcoding techniques. A one single reaction assay using a trait-independent marker (*psbA-trnH igs*) coupled together with a trait-related marker based on the enzyme eugenol *O*-methyltransferase (EOMT) has shown positive results for the discrimination of *O. tenuiflorum* from other *Ocimum* sp., in reference plants and in commercial samples.

##### 4.4.1 DNA-based authentication using trait-independent marker needs optimization

The plastidic *psbA-trnH igs* marker has been widely used for plant species identification (Hollingsworth *et al.*, 2011; CBOL Plant Working Group, 2009). This plastidic trait-independent marker possesses, in theory, the three essential requirements described for DNA barcodes which are (1) universality, (2) power of discrimination or selectiveness and (3) sequence quality (Hollingsworth *et al.*, 2011). However, for species identification sequencing is needed and when these techniques are used for commercial purposes, optimization in order to lower the amount of sequencing is desired. In this research, it was possible to discriminate different species of *Ocimum* by the amplification of the *psbA-trnH* marker followed by a RLFP. The results of this assay were as predicted, a specific single-nucleotide polymorphism is present in the the *psbA-trnH* sequence of *O. tenuiflorum* (haplotype II), *O. campechianum* (haplotype II) and *O. gratissimum* (haplotype III), which generates a restriction site for *HinfI* (Figure 3.17). As a result, a double band pattern is visualized

for the mentioned species. Contrary to this, the enzyme does not digest the amplicons from *O. basilicum*, *O. americanum*, *O. x africanum*, and *O. kilimandscharicum*, all belonging to haplotype I (Jürges *et al.*, 2018). When this RFLP approach was applied to commercial samples, the *psbA-trnH* amplification coupled with RFLP assay also allowed to discriminate Tulsi samples (**Figure 3.18**). Therefore, RFLP is an optimization for the *psbA-trnH* barcode, since it does allow to distinguish *O. tenuiflorum* from other species in commercial samples. This type of optimizations is not necessary when species have a different size of the *psbA-trnH* region. However, in this case optimization is desirable because the pattern among species were similar, since all *Ocimum* species tested have a similar size amplicon of the *psbA-trnH* (around 400 bp), and therefore limiting a rapid discrimination (Ríos-Rodríguez *et al.*, 2021).

Even though RFLP allows the discrimination of samples before sequencing for identification, it is a two-step protocol, meaning that it is time-consuming and hence costly. Therefore, one-step protocols are desired, for instance the Amplification-Refractory Mutation System (ARMS) (Horn *et al.*, 2014). Nonetheless, in the case of *Ocimum* species, the *psbA-trnH* region, potentially diagnostic SNP are localised in AT-rich areas, disallowing ARMS primer design. In addition, other commonly used independent marker such as *matK* and *rbcL* have been previously used in *Ocimum* sp. showing that they are not a plausible alternative since their sequences is too conservative within the species (Jürges *et al.*, 2018). The former markers have not enough discrimination power, even though they are widely used for plant species identification, in the case of differentiation within *Ocimum* species. This particular case, highlights the need to look for a proper marker for a given situation. Subsequently, a trait-related marker was therefore an option for authentication (Ríos-Rodríguez *et al.*, 2021).

#### 4.4.2 DNA-based authentication using a trait-related marker and its potential

The occurrence of methyleugenol in *Ocimum* sp. largely varies in content, even within a given species (**Table 3.2**). These differences could be a consequence of genetic factors (chemotypes), and/or environmental conditions. For instance, previous studies have shown changes on the chemical profile given the light quality and draught stress in *Ocimum* sp. (Jürges *et al.*, 2018; Johnson *et al.*, 1999; Khakdan *et al.*, 2017). Yet, environmental influences have its limits under genetic control. Therefore, since methyleugenol is present in several *Ocimum* species, the enzyme that synthesises ME, the EOMT, it was hypothesized to be used as a trait-related marker (Ríos-Rodríguez *et al.*, 2021).

The EOMT sequence was therefore used to develop an ARMS test, in which more than one SNP in the diagnostic primer would result in a more robust differentiation.

The idea was tested in 13 validated *Ocimum* reference plants (**Figure 3.19**). All *Ocimum* plant accessions corresponding to haplotype I (Jürges *et al.*, 2018) produced only the expected full-length amplicon, while a second band was found in *O. tenuiflorum* accessions. The double band pattern of Tulsi was shared with *O. gratissimum* and *O. campechianum* (representing haplotypes III and II). It is interesting that similarities between these haplotypes are not exclusively delimited to trait-independent markers like in the previous RFLP case. Especially, when the chemical profile (see **Figure 3.6**) of *O. gratissimum*, for example, suggests a greater differentiation with *O. tenuiflorum*. However, perhaps the extent of DNA-based differentiation assays has its limitations, even though trait-related markers are used, due to the presence of pseudogenes. Potential hybridization can cause gene duplication, resulting in several pseudo copies of genes encoded in the gDNA, producing finally higher similarities among species.

In any case, the results of this assay tested in commercial products were not as well-defined as for the reference plants, and two potential reasons for this outcome will be discussed. First, probably DNA degradation occurs through food products processing (Stoeckle *et al.*, 2011; Ivanova *et al.*, 2016), being the smaller diagnostic ARMS fragment less susceptible to it. A second reason adding to this inferior performance compared to the trait-independent *psbA-trnH* *igs* marker, could be the lower gene copies (only two) of EOMT given that it is a nuclear gene, whereas there are several copies of the plastidic *psbA-trnH* marker, meaning that not enough DNA from EOMT exists as template after partial degradation attributable to processing (Hollingsworth *et al.*, 2011). Consequently, from these limitations existing in commercial samples, the assay is prone to false negative results. Moreover, using EOMT as trait-related marker, generates a problem of “over-specificity”, hence is ambiguous and discrimination of commercial samples is not possible (Ríos-Rodríguez *et al.*, 2021). However, the reference plants discrimination using the trait-related marker gave a good result. Therefore, to improve the assays performance, the trait-independent marker together in a multiplex PCR with the trait-related marker, was thought to be stronger against the negative effect of low gene copies and degraded DNA (Ríos-Rodríguez *et al.*, 2021).

#### **4.4.3 DNA-based authentication by combined trait-related and -independent makers is a feasible optimized option in commercial products**

One reaction having the *psbA-trnH* trait-independent marker primers together with adapted EOMT trait-related marker primers was tested in the reference plants (**Figure 3.21**) (Ríos-Rodríguez *et al.*, 2021). As expected, the result from this multiplex PCR was consistent with the results from the RFLP and the EOMT-ARMS assays. The same reaction set up was tested in commercial samples and again results were equivalent with those obtained by RFLP (**Figure 3.22**), but in this case a one-step

assay was enough. Therefore, the coupled trait-independent and -related markers used together in a PCR were indeed more robust when analysing commercial samples and are a plausible alternative for *O. tenuiflorum* discrimination and identification of commercial samples. Summarizing again the desirable traits for primers used in DNA barcoding, the *psbA-trnH* has the characteristics of being universal, has power of discrimination (after sequencing), and has a quality sequence. However, the power of discrimination through an immediate evaluation by gel electrophoresis before sequencing is not possible in the case of *Ocimum* sp. In this sense, the EOMT primer added into the reaction solved the problem, allowing discrimination by gel electrophoresis.

#### 4.4.4 The EOMT DNA marker is not related to methyleugenol content

The ME content in commercial samples (**Table 3.3**) suggest that the EOMT trait-related marker is able to identify *O. tenuiflorum* chemotypes that are ME producers. In this regards, and as shown in **Table 3.2**, there are also *O. tenuiflorum* chemotypes that are not ME producers. However, whether the EOMT DNA marker here developed can identify those chemotypes or not, is a hypothesis yet to be proven. Further, because commercial samples of common Basil also had ME, lacking of the EOMT diagnostic band is no assurance of the presence of ME in the sample. This proves that the intended specificity of the EOMT marker to identify *O. tenuiflorum* from other *Ocimum* sp. works, and it cannot be used to determine the presence/absence of ME in *Ocimum* sp. (Ríos-Rodríguez *et al.*, 2021). Equally, a sample should not be identified as *O. tenuiflorum* just because of the presence of methyleugenol, which can easily be a false positive in such case. Therefore, the EOMT marker positively discriminating samples cannot be used as a prediction tool for the methyleugenol content, but only as authentication tool of *O. tenuiflorum* (Ríos-Rodríguez *et al.*, 2021).

## 5 Conclusion

The botanical study of *O. tenuiflorum* was initiated in order to better comprehend the toxicological nature of Tulsi due to its currently increased use as food, even positively labelled as 'superfood'. This study has particularly brought new insights into the current comprehension of the synthesis of methyleugenol via the phenylpropanoid pathway. The phenylpropanoid pathway is one of the oldest pathways developed in plants when conquering land, and one of the first and more studied. This pathway gives birth to a variety of secondary metabolites that served different functions in the plant. Among these is methyleugenol, which is synthesised by the enzyme eugenol *O*-methyltransferase (EOMT). A relevant question here was the hypothesised existence of genetic variation of the EOMT across the genus *Ocimum*, and if there were, could such variation influence the methyleugenol synthesis. In this research, the findings on genetic variation in EOMT among *Ocimum* sp. seemed to be linked to differences on plants' chemical profile. The hypothesis, therefore stands, and a set of specific amino acids in the sequences are now suggested to play a role on the enzymatic function. Further, this variability extends to single individuals, that present EOMT isoforms not described before. Following these findings, this research was directed into the phenylpropanoid pathway regulation.

The knowledge on the phenylpropanoid pathway regulation is vast and variable, depending on the plant species, the stress/signal that triggers the pathway, and the final product. Therefore, to better understand the occurrence of methyleugenol in Tulsi, it was necessary to particularly assess *O. tenuiflorum*. The results of this study suggest distinctive regulation mechanisms for the phenylpropanoid pathways in two different *O. tenuiflorum* chemotypes. In Rama Tulsi, the upstream reactions to methyleugenol seemed to be highly driven by the enzyme caffeic acid *O*-methyltransferase (COMT). While in Krishna Tulsi, the enzyme caffeoyl-CoA *O*-methyltransferase (CCOMT) seemed to also play an important role. Moreover, because of the low transcripts' levels of phenylalanine ammonia lyase (PAL) and cinnamate 4-hydroxylase (C4H) in both chemotypes, tyrosine it is hypothesised to be a more important actor in the phenylpropanoid pathway in *O. tenuiflorum*. In addition, using UV-B radiation to trigger this pathway, allowed for studying fast response mechanisms used by plants as defence: transcripts' levels changed as early as 10 min after UV-B radiation, suggesting post-transcriptional/translational regulation. However, the regulation mechanisms from the signal reception to the metabolites synthesis are not fully understood, and it highlights the complexity of the actors playing a role on the pathway regulation. For instance, the type of stress, the species and genotypes in the study, the time frame of response, and the localization of the pathway among others. Particularly, the enzyme localization inside the cell

was a question in this research. The subcellular localization of the enzyme eugenol *O*-methyltransferase was unknown, and found to be in the cytosol and the nucleus of transformed BY-2 cells. The cytosolic localization was consistent with suggestions and findings by other studies, that have assessed the subcellular localization of a number of phenylpropanoid pathway enzymes. However, the nuclear localization of EOMT was puzzling and leaves the door wide open for further research.

Last but not least, to address the challenge of food fraud in commercial Tulsi products, the question was whether it was possible or not to optimize currently used DNA barcodes. This optimization was achieved by an assay using the commonly employed trait-independent marker (*pasbA-trnH igs*), together with a trait-related marker (EOMT) that was developed in this study. This assay presented a higher power of discrimination than *pasbA-trnH igs* alone and allowed for differentiation of *O. tenuiflorum* from surrogates in commercial samples that declared to contain Tulsi. However, it was not related to the occurrence of methyleugenol in commercial samples, which can only be verified by metabolites detection methods such as gas chromatography.

Overall, the case study of Tulsi, highlights the complex issue of considering a traditional medicinal plant -containing toxicologically relevant compounds- as foods or 'superfoods'. This study exposed that even though there is great knowledge on the phenylpropanoid pathway, there is still more to discover. It also showed that even when some species are normally used in food, when chemically evaluating the genus -in this case *Ocimum*- more than one toxic compound might be present in the genus/species genotypes. This is particularly important when using medicinal plants as foods or 'superfoods'. Furthermore, surrogation of medicinal plants is a problem that accompanies this transition of TMP to food. And tools, either for risk avoidance or detection of food fraud, are necessary. Finally, this botanical study on Tulsi -the 'superfood'- supports the idea that precautionary research should be done prior turning traditional medicinal plants into new traditions.

## 6 Outlook

This study has achieved several research goals associated to the phenylpropanoid pathway and the comprehension of the synthesis of methyleugenol in *O. tenuiflorum*. In addition, specific DNA barcodes have been developed for discriminating Tulsi in commercial samples. However, several questions remain open, and therefore, additional research is here suggested.

Concerning the enzyme eugenol *O*-methyltransferase, variation on the amino acidic profile was found across the genus *Ocimum*, and among *O. tenuiflorum* genotypes. Research to understand how this *aa* variation might influence the chemical profile of plants is proposed. For instance, substrate specificity and product formation analysis of several isoforms would help to identify the promiscuity degree of the protein isoforms. In addition, crystal structure analysis of the enzyme would help to elucidate if the variability does or does not have influence on the conformation of the enzyme and on the formation of the homodimer. Also, studies on the role of the introns could help to understand, for instance, their role on isoforms' formation.

Regarding the phenylpropanoid pathway, it is advisable to include more enzymes and current enzymes isoforms in the evaluation of gene expression. For example, tyrosine ammonia lyase (TAL) and chalcone synthase (CHS). Furthermore, it would be interesting to explore the signalling path between the receptor (UVR8) and the final product (ME), including potential intermediates such as jasmonic acid. It would be also interesting to extend the studies on the fast gene response, such as, the potential role of free cytosolic  $Ca^{2+}$  on phenylpropanoid enzymes transcription levels increment. Moreover, extending the time frame of the study, including sampling several hours after radiation, can help to better comprehend the metabolism regulation and defence mechanisms.

About the subcellular localization of the enzyme eugenol *O*-methyltransferase, it would be of interest to explore the localization in *O. tenuiflorum* leaves to confirm the results obtained in transformed BY-2 cells. Additionally, co-localization studies including other enzymes of the phenylpropanoid pathway, can help to elucidate the enzymatic dynamics in the cell. Further, it is of great interest to study the nuclear localization of EOMT found in this study.

Lastly, on the topic of DNA barcoding, it was possible to differentiate *O. tenuiflorum* from surrogates in commercial samples, in an optimized method given by the specificity of EOMT. However, the primers here used could be further adapted to other techniques such as qPCR and Next Generation Sequencing, in order to obtain a more sensitive assay that can perform better in mixture samples.

---

## 7 References

- Abudayyak, M., Jannuzzi, A.T., Özhan, G. and Alpertunga, B. (2015) Investigation on the toxic potential of *Tribulus terrestris* in vitro. *Pharm. Biol.*, **53**, 469–476.
- Achnine, L., Blancaflor, E.B., Rasmussen, S. and Dixon, R.A. (2004) Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. *Plant Cell*, **16**, 3098–3109.
- Al-Malahmeh, A.J., Al-ajlouni, A.M., Wesseling, S., Vervoort, J. and Rietjens, I.M.C.M. (2017) Determination and risk assessment of naturally occurring genotoxic and carcinogenic alkenylbenzenes in basil-containing sauce of pesto. *Toxicol. Reports*, **4**, 1–8. Available at: <https://doi.org/10.1016/j.toxrep.2016.11.002>.
- An, J., Moon, J.C., Kim, J.H., Kim, G.S. and Jang, C.S. (2019) Development of DNA-based species-specific real-time PCR markers for four berry fruits and their application in commercial berry fruit foods. *Appl. Biol. Chem.*, **62**, 0–6. Available at: <https://doi.org/10.1186/s13765-019-0413-9>.
- Ang, L.H., Chattopadhyay, S., Wei, N., Oyama, T., Okada, K., Batschauer, A. and Deng, X.W. (1998) Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of *Arabidopsis* development. *Mol. Cell*, **1**, 213–222.
- Balaji, S. and Chempakam, B. (2010) Toxicity prediction of compounds from turmeric (*Curcuma longa* L). *Food Chem. Toxicol.*, **48**, 2951–2959. Available at: <http://dx.doi.org/10.1016/j.fct.2010.07.032>.
- Barbosa, C., Nogueira, S., Gadanho, M. and Chaves, S. (2019) Study on commercial spice and herb products using next-generation sequencing (NGS). *J. AOAC Int.*, **102**, 369–375.
- Benitez, N.P., Meléndez León, E.M. and Stashenko, E.E. (2009) Eugenol and methyl eugenol chemotypes of essential oil of species *Ocimum gratissimum* L. and *Ocimum campechianum* mill. from Colombia. *J. Chromatogr. Sci.*, **47**, 800–803.
- Bhatnagar, A., Singh, S., Khurana, J.P. and Burman, N. (2020) HY5-COP1: the central module of light signaling pathway. *J. Plant Biochem. Biotechnol.*, **29**, 590–610. Available at: <https://doi.org/10.1007/s13562-020-00623-3>.
- Biała, W. and Jasiński, M. (2018) The phenylpropanoid case – It is transport that matters. *Front. Plant Sci.*, **871**, 1–8.
- Biever, J.J. and Gardner, G. (2016) The relationship between multiple UV-B perception mechanisms and DNA repair pathways in plants. *Environ. Exp. Bot.*, **124**, 89–99. Available at: <http://dx.doi.org/10.1016/j.envexpbot.2015.12.010>.
- Boer, H.J. de, Ichim, M.C. and Newmaster, S.G. (2015) DNA barcoding and pharmacovigilance of herbal medicines. *Drug Saf.*, **38**, 611–620.
- Brahmavarchas (2007) *Ayurveda ka praan: vanaaushidhii vigyan*, Tapobhumi, Mathura: yug nirma yojnaa press.

- Breeze, R. (2017) Explaining superfoods: Exploring metaphor scenarios in media science reports. *Iberica*, **2017**, 67–88.
- Brown, B.A., Cloix, C., Jiang, G.H., Kaiserli, E., Herzyk, P., Kliebenstein, D.J. and Jenkins, G.I. (2005) A UV-B-specific signaling component orchestrates plant UV protection. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 18225–18230.
- Buchanan, B.B., Gruissem, W. and Jones, R.L. eds. (2015) *Biochemistry and molecular biology of plants* 2nd ed., Hoboken, NJ: Wiley-Blackwell.
- Butterworth, M., Davis, G., Bishop, K., Reyna, L. and Rhodes, A. (2020) What Is a Superfood Anyway? Six Key Ingredients for Making a Food “Super.” *Gastronomica*, **20**, 46–58.
- Carović-Stanko, K., Liber, Z., Besendorfer, V., Javornik, B., Bohanec, B., Kolak, I. and Satovic, Z. (2010) Genetic relations among basil taxa (*Ocimum* L.) based on molecular markers, nuclear DNA content, and chromosome number. *Plant Syst. Evol.*, **285**, 13–22.
- Casati, P., Morrow, D.J., Fernandes, J.F. and Walbot, V. (2011) UV-B signaling in maize: Transcriptomic and metabolomic studies at different irradiation times. *Plant Signal. Behav.*, **6**, 1926–1931.
- CBOL Plant Working Group (2009) A DNA barcode for land plants. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 12794–12797. Available at: [www.pnas.org/cgi/doi/10.1073/pnas.0905845106](http://www.pnas.org/cgi/doi/10.1073/pnas.0905845106).
- Charen, E. and Harbord, N. (2020) Toxicity of Herbs, Vitamins, and Supplements. *Adv. Chronic Kidney Dis.*, **27**, 67–71. Available at: <http://dx.doi.org/10.1053/j.ackd.2019.08.003>.
- Chorev, M. and Carmel, L. (2012) The function of introns. *Front. Genet.*, **3**, 1–15.
- Chowdhury, T., Mandal, A., Roy, S.C. and Sarker, D. De (2017) Diversity of the genus *Ocimum* (*Lamiaceae*) through morpho-molecular (RAPD) and chemical (GC–MS) analysis. *J. Genet. Eng. Biotechnol.*, **15**, 275–286. Available at: <http://dx.doi.org/10.1016/j.jgeb.2016.12.004>.
- Christina, V.L.P. and Annamalai, A. (2014) Nucleotide based validation of *Ocimum* species by evaluating three candidate barcodes of the chloroplast region. *Mol. Ecol. Resour.*, **14**, 60–68.
- Contreras, R.A., Pizarro, M., Köhler, H., Zamora, P. and Zúñiga, G.E. (2019) UV-B shock induces photoprotective flavonoids but not antioxidant activity in Antarctic *Colobanthus quitensis* (Kunth) Bartl. *Environ. Exp. Bot.*, **159**, 179–190. Available at: <https://doi.org/10.1016/j.envexpbot.2018.12.022>.
- Corazza, O., Martinotti, G., Santacroce, R., Chillemi, E., Giannantonio, M. Di, Schifano, F. and Celtek, S. (2014) Sexual enhancement products for sale online: Raising awareness of the psychoactive effects of yohimbine, maca, horny goat weed, and ginkgo biloba. *Biomed Res. Int.*, **2014**.
- Coutinho Moraes, D.F., Still, D.W., Lum, M.R. and Hirsch, A.M. (2015) DNA-Based Authentication of Botanicals and Plant-Derived Dietary Supplements: Where Have We Been and Where Are We Going? *Planta Med.*, **81**, 687–695.

- Ćwiela-Drabek, M., Piekut, A., Szymala, I., Oleksiuk, K., Razzaghi, M., Osmala, W., Jabłońska, K. and Dziubanek, G. (2020) Health risks from consumption of medicinal plant dietary supplements. *Food Sci. Nutr.*, **8**, 3535–3544.
- Dastmalchi, M., Bernards, M.A. and Dhaubhadel, S. (2016) Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. *Plant J.*, **85**, 689–706.
- Davey, M.P., Susanti, N.I., Wargent, J.J., Findlay, J.E., Paul Quick, W., Paul, N.D. and Jenkins, G.I. (2012) The UV-B photoreceptor UVR8 promotes photosynthetic efficiency in *Arabidopsis thaliana* exposed to elevated levels of UV-B. *Photosynth. Res.*, **114**, 121–131.
- Doerr, I. (2018) *Basilikum-Früchte: Entwicklung diagnostischer Methoden zur Authentifizierung von Chia (Salvia hispanica) Surrogaten und Prüfung auf Estragol und Methyleugenol*. Karlsruhe Institute of Technology.
- Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C. and Lepiniec, L. (2010) MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.*, **15**, 573–581.
- EC-SCF (2001) *Opinion of the scientific committee on food on methyleugenol (4-allyl-1,2-dimethoxybenzene)*.
- EC (2017) Commission implementing Regulation (EU) 2017/2470 of 20 December 2017 establishing the Union list of novel foods in accordance with Regulation (EU) 2015/2283 of the European Parliament and of the Council on novel foods. *Off. J. Eur. Union*, **351**, 1–188. Available at: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R2470&from=EN>.
- EC (2008) Regulation (EC) No 1334/2008 of the European Parliament and of the Council. *Off. J. Eur. Union*, 2008, **L 354/34**, 34–50.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
- EMA (2005) Public statement on the use of herbal medicinal products containing methyleugenol. *Emea/Hmpc/138363/2005 Comm.* Available at: [https://www.ema.europa.eu/en/documents/scientific-guideline/public-statement-use-herbal-medicinal-products-containing-methyleugenol\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/public-statement-use-herbal-medicinal-products-containing-methyleugenol_en.pdf).
- Emiliani, G., Fondi, M., Fani, R. and Gribaldo, S. (2009) A horizontal gene transfer at the origin of phenylpropanoid metabolism: a key adaptation of plants to land. *Biol. Direct*, **4**, 1–12.
- Fang, W., Cheng, H., Duan, Y., Jiang, X. and Li, X. (2012) Genetic diversity and relationship of clonal tea (*Camellia sinensis*) cultivars in China as revealed by SSR markers. *Plant Syst. Evol.*, **298**, 469–483.
- Figueiredo, P.L.B., Silva, S.G., Nascimento, L.D., Ramos, A.R., Setzer, W.N., Silva, J.K.R. Da and Andrade, E.H.A. (2018) Seasonal study of methyleugenol chemotype of *Ocimum campechianum* essential oil and its fungicidal and antioxidant activities. *Nat. Prod. Commun.*, **13**, 1055–1058.

- Frohnmeier, H., Loyall, L., Blatt, M.R. and Grabov, A. (1999) Millisecond UV-B irradiation evokes prolonged elevation of cytosolic-free Ca<sup>2+</sup> and stimulates gene expression in transgenic parsley cell cultures. *Plant J.*, **20**, 109–117.
- Galimberti, A., Mattia, F. De, Losa, A., Bruni, I., Federici, S., Casiraghi, M., Martellos, S. and Labra, M. (2013) DNA barcoding as a new tool for food traceability. *Food Res. Int.*, **50**, 55–63. Available at: <http://dx.doi.org/10.1016/j.foodres.2012.09.036>.
- Gális, I., Šimek, P., Narisawa, T., Sasaki, M., Horiguchi, T., Fukuda, H. and Matsuoka, K. (2006) A novel R2R3 MYB transcription factor NtMYBJS1 is a methyl jasmonate-dependent regulator of phenylpropanoid-conjugate biosynthesis in tobacco. *Plant J.*, **46**, 573–592.
- Gang, D.R. (2005) Evolution of flavors and scents. *Annu. Rev. Plant Biol.*, **56**, 301–325.
- Gang, D.R., Lavid, N., Zubieta, C., Chen, F., Beuerle, T., Lewinsohn, E., Noel, J.P. and Pichersky, E. (2002) Characterization of phenylpropene *O*-methyltransferases from Sweet Basil: facile change of substrate specificity and convergent evolution within a plant *O*-methyltransferase family. *Plant Cell*, **14**, 505–519. Available at: [www.plantcell.org/cgi/doi/10.1105/tpc.010327](http://www.plantcell.org/cgi/doi/10.1105/tpc.010327).
- Gang, D.R., Simon, J., Lewinsohn, E. and Pichersky, E. (2002) Peltate glandular trichomes of *Ocimum basilicum* L. (sweet basil) contain high levels of enzymes involved in the biosynthesis of phenylpropenes. *J. Herbs, Spices Med. Plants*, **9**, 189–195.
- Gang, D.R., Wang, J., Dudareva, N., Nam, K.H., Simon, J.E., Lewinsohn, E. and Pichersky, E. (2001) An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiol.*, **125**, 539–555.
- GEP (2015) Glossary for understanding eukaryotic genes. Genomic Education Partnership (GEP) - The University of Alabama, Washington University in St. Louis. Available at: [https://thegep.org/lessons/mlaakso-reference-ueg\\_glossary\\_beginning\\_students/](https://thegep.org/lessons/mlaakso-reference-ueg_glossary_beginning_students/).
- Gong, L., Qiu, X.H., Huang, J., Xu, W., Bai, J.Q., Zhang, J., Su, H., Xu, C.M. and Huang, Z.H. (2018) Constructing a DNA barcode reference library for southern herbs in China: A resource for authentication of southern Chinese medicine. *PLoS One*, **13**, 1–12.
- Goswami, H. and Ram, H. (2017) Ancient food habits dictate that food can be medicine but medicine cannot be “food”!! *Medicines*, **4**, 82.
- Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.*, 95–98.
- Haynes, E., Jimenez, E., Pardo, M.A. and Helyar, S.J. (2019) The future of NGS (Next Generation Sequencing) analysis in testing food authenticity. *Food Control*, **101**, 134–143. Available at: <https://doi.org/10.1016/j.foodcont.2019.02.010>.
- Hebert, P.D.N., Cywinska, A., Ball, S.L. and DeWaard, J.R. (2003) Biological identifications through DNA barcodes. *Proc. R. Soc. B Biol. Sci.*, **270**, 313–321.
- Hollingsworth, P.M. (2011) Refining the DNA barcode for land plants. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 19451–19452.

- Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., et al. (2009) A DNA barcode for land plants. *Proc. Natl. Acad. Sci. U. S. A.*, **106**, 12794–12797. Available at: <http://www.pnas.org/content/106/31/12794.abstract>.
- Hollingsworth, P.M., Graham, S.W. and Little, D.P. (2011) Choosing and using a plant DNA barcode. *PLoS One*, **6**.
- Horn, T., Völker, J., Rühle, M., Häser, A., Jürges, G. and Nick, P. (2014) Genetic authentication by RFLP versus ARMS? The case of Moldavian dragonhead (*Dracocephalum moldavica* L.). *Eur. Food Res. Technol.*, **238**, 93–104.
- Hrazdina, G. and Jensen, R. a (1992) Spatial organization of enzymes in plant metabolic pathways. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **43**, 241–267.
- Ibrahim, R.K., Bruneau, A. and Bantignies, B. (1998) Plant *O*-methyltransferases: molecular analysis, common signature and classification. *Plant Mol. Biol.*, **36**, 1–10.
- Ivanova, N. V., Kuzmina, M.L., Braukmann, T.W.A., Borisenko, A. V. and Zakharov, E. V. (2016) Authentication of herbal supplements using next-generation sequencing. *PLoS One*, **11**, 1–24.
- Jeurissen, S.M.F. (2007) *Bioactivation and genotoxicity of the herbal constituents safrole, estragole and methyleugenol*. Wageningen University.
- Jeurissen, S.M.F., Bogaards, J.J.P., Awad, H.M., et al. (2006) Human cytochrome P450 enzymes of importances for bioactivation of the proximate carcinogen 1'-hydroxymethyleugenol. *Chem. Res. Toxicol.*, **19**, 111–1116. Available at: <https://doi.org/10.1021/tx050267h>.
- Jirovetz, L., Buchbauer, G., Shafi, M.P. and Kaniampady, M.M. (2003) Chemotaxonomical analysis of the essential oil aroma compounds of four different *Ocimum* species from southern India. *Eur. Food Res. Technol.*, **217**, 120–124.
- Johnson, C.B., Kirby, J., Naxakis, G. and Pearson, S. (1999) Substantial UV-B-mediated induction of essential oils in sweet basil (*Ocimum basilicum* L.). *Phytochemistry*, **51**, 507–510.
- Joshi, R. (2013) Chemical composition, in vitro antimicrobial and antioxidant activities of the essential oils of *Ocimum gratissimum*, *O. sanctum* and their major constituents. *Indian J. Pharm. Sci.*, **75**, 457–462.
- Jürges, G., Beyerle, K., Tossenberger, M., Häser, A. and Nick, P. (2009) Development and validation of microscopical diagnostics for “Tulsi” (*Ocimum tenuiflorum* L.) in ayurvedic preparations. *Eur. Food Res. Technol.*, **229**, 99–106.
- Jürges, G., Sahi, V., Rios Rodriguez, D., Reich, E., Bhamra, S., Howard, C., Slater, A. and Nick, P. (2018) Product authenticity versus globalisation—The Tulsi case. *PLoS One*, **13**, 1–22.
- Kapteyn, J., Qualley, A. V., Xie, Z., Fridman, E., Dudareva, N. and Gang, D.R. (2007) Evolution of cinnamate/p-coumarate carboxyl methyltransferases and their role in the biosynthesis of methylcinnamate. *Plant Cell*, **19**, 3212–3229.

- Karimi, M., Inzé, D. and Depicker, A. (2002) GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.*, **7**, 193–195. Available at: file:///C:/Users/Risa/Downloads/1-s2.0-S1360138502022513-main.pdf.
- Khakdan, F., Nasiri, J., Ranjbar, M. and Alizadeh, H. (2017) Water deficit stress fluctuates expression profiles of 4Cl, C3H, COMT, CVOMT and EOMT genes involved in the biosynthetic pathway of volatile phenylpropanoids alongside accumulation of methylchavicol and methyleugenol in different Iranian cultivars of basil. *J. Plant Physiol.*, **218**, 74–83. Available at: <http://dx.doi.org/10.1016/j.jplph.2017.07.012>.
- Khalki, L., M'hamed, S.B., Bennis, M., Chait, A. and Sokar, Z. (2010) Evaluation of the developmental toxicity of the aqueous extract from *Trigonella foenum-graecum* (L.) in mice. *J. Ethnopharmacol.*, **131**, 321–325.
- Kim, H.J., Chen, F., Wang, X. and Rajapakse, N.C. (2006) Effect of methyl jasmonate on secondary metabolites of sweet basil (*Ocimum basilicum* L.). *J. Agric. Food Chem.*, **54**, 2327–2332.
- Koeduka, T., Baiga, T.J., Noel, J.P. and Pichersky, E. (2009) Biosynthesis of t-anethole in anise: Characterization of t-anol/isoegenol synthase and an *O*-methyltransferase specific for a C7-C8 propenyl side chain. *Plant Physiol.*, **149**, 384–394.
- Kreuzer, M., Howard, C., Adhikari, B., Pendry, C.A. and Hawkins, J.A. (2019) Phylogenomic approaches to DNA barcoding of herbal medicines: Developing clade-specific diagnostic characters for Berberis. *Front. Plant Sci.*, **10**.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. (2018) MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.*, **35**, 1547–1549.
- Kumari, R. and Agrawal, S.B. (2011) Comparative analysis of essential oil composition and oil containing glands in *Ocimum sanctum* L. (Holy basil) under ambient and supplemental level of UV-B through gas chromatography-mass spectrometry and scanning electron microscopy. *Acta Physiol. Plant.*, **33**, 1093–1101.
- Lawal, O.A., Ogunwande, I.A., Omikorede, O.E., Owolabi, M.S., Olorunsola, F.F., Sanni, A.A., Amisu, K.O. and Opoku, A.R. (2014) Chemical composition and antimicrobial activity of essential oil of *Ocimum kilimandscharicum* (R. Br.) Guerke: A new chemotype. *Am. J. Essent. Oil Nat. Prod.*, **2**, 41–46.
- Lecourieux, D., Mazars, C., Pauly, N., Ranjeva, R. and Pugin, A. (2002) Analysis and effects of cytosolic free calcium increases in response to elicitors in *Nicotiana plumbaginifolia* cells. *Plant Cell*, **14**, 2627–2641.
- Lecourieux, D., Raneva, R. and Pugin, A. (2006) Calcium in plant defence-signalling pathways. *New Phytol.*, **171**, 249–269.
- Li, Z., Wang, X., Chen, F. and Kim, H.J. (2007) Chemical changes and overexpressed genes in sweet basil (*Ocimum basilicum* L.) upon methyl jasmonate treatment. *J. Agric. Food Chem.*, **55**, 706–713.
- Liang, T., Yang, Y. and Liu, H. (2019) Signal transduction mediated by the plant UV-B photoreceptor UVR8. *New Phytol.*, **221**, 1247–1252.

- Lopez-Vizcón, C. and Ortega, F. (2012) Detection of mislabelling in the fresh potato retail market employing microsatellite markers. *Food Control*, **26**, 575–579. Available at: <http://dx.doi.org/10.1016/j.foodcont.2012.02.020>.
- Loyer, J. (2016) *The social lives of superfoods*. University of Adelaide.
- Malav, P., Pandey, A., Bhatt, K.C., Gopala Krishnan, S. and Bisht, I.S. (2015) Morphological variability in holy basil (*Ocimum tenuiflorum* L.) from India. *Genet. Resour. Crop Evol.*, **62**, 1245–1256. Available at: <http://dx.doi.org/10.1007/s10722-015-0227-5>.
- Maurya, S., Chandra, M., Yadav, R.K., et al. (2019) Interspecies comparative features of trichomes in *Ocimum* reveal insights for biosynthesis of specialized essential oil metabolites. *Protoplasma*, **256**, 893–907.
- Milan, E., Mandoulakani, B.A. and Kheradmand, F. (2017) The effect of methyl jasmonate on the expression of phenylalanine ammonia lyase and eugenol-o-methyl transferase genes in basil. *Philipp. Agric. Sci.*, **100**, 163–167.
- Muhlemann, J.K., Woodworth, B.D., Morgan, J.A. and Dudareva, N. (2014) The monoglignol pathway contributes to the biosynthesis of volatile phenylpropenes in flowers. *New Phytol.*, **204**, 661–670.
- Müller-Xing, R., Xing, Q. and Goodrich, J. (2014) Footprints of the sun: memory of UV and light stress in plants. *Front. Plant Sci.*, **5**, 1–12.
- Nagata, T., Nemoto, Y. and Hasezawa, S. (1992) Tobacco BY-2 cell Line as the “HeLa” cell in the cell biology of higher plants. *Int. Rev. Cytol.*, **132**, 1–30.
- Oravecz, A., Baumann, A., Máté, Z., et al. (2006) Constitutively photomorphogenic1 is required for the UV-B response in *Arabidopsis*. *Plant Cell*, **18**, 1975–1990.
- Panchy, N., Lehti-Shiu, M. and Shiu, S.H. (2016) Evolution of gene duplication in plants. *Plant Physiol.*, **171**, 2294–2316.
- Pandey, A., Misra, P., Bhambhani, S., Bhatia, C. and Trivedi, P.K. (2014) Expression of *Arabidopsis* MYB transcription factor, AtMYB111, in tobacco requires light to modulate flavonol content. *Sci. Rep.*, **4**.
- Patel, H.K., Fougat, R.S., Kumar, S., Mistry, J.G. and Kumar, M. (2015) Detection of genetic variation in *Ocimum* species using RAPD and ISSR markers. *3 Biotech*, **5**, 697–707. Available at: <http://dx.doi.org/10.1007/s13205-014-0269-y>.
- Phua, D.H., Zosel, A. and Heard, K. (2009) Dietary supplements and herbal medicine toxicities-when to anticipate them and how to manage them. *Int. J. Emerg. Med.*, **2**, 69–76.
- Pullaiah, T., Krishnamurthy, K.V. and Bahadur, B. eds. (2017) *Ethnobotany of India - Volume 2: Western Ghats and west coast of peninsular India*, Waretown, NJ: Apple Academic Press Inc.
- Pyne, R.M., Honig, J.A., Vaiciunas, J., Wyenandt, C.A. and Simon, J.E. (2018) Population structure, genetic diversity and downy mildew resistance among *Ocimum* species germplasm. *BMC Plant Biol.*, **18**, 1–15.

- Raclariu, A.C., Paltinean, R., Vlase, L., Labarre, A., Manzanilla, V., Ichim, M.C., Crisan, G., Brysting, A.K. and Boer, H. De (2017) Comparative authentication of *Hypericum perforatum* herbal products using DNA metabarcoding, TLC and HPLC-MS. *Sci. Rep.*, **7**, 8–10.
- Raina, A.P. and Misra, R.C. (2018) Chemo-divergence in essential oil composition among germplasm collection of five *Ocimum* species from eastern coastal plains of India. *J. Essent. Oil Res.*, **30**, 47–55. Available at: <http://doi.org/10.1080/10412905.2017.1371087>.
- Rajalakshmi, T.R., Aravindh Babu, N., Shanmugam, K.T. and Masthan, K.M.K. (2015) DNA adducts-chemical addons. *J. Pharm. Bioallied Sci.*, **7**, S197-9.
- Rajeswara Rao, B.R., Kothari, S.K., Rajput, D.K., Patel, R.P. and Darokar, M.P. (2011) Chemical and biological diversity in fourteen selections of four *Ocimum* species. *Nat. Prod. Commun.*, **6**, 1705–1710.
- Ramakrishnan, A.P., Ritland, C.E., Blas Sevillano, R.H. and Riseman, A. (2015) Review of potato molecular markers to enhance trait selection. *Am. J. Potato Res.*, **92**, 455–472.
- Rastogi, S., Kumar, R., Chanotiya, C.S., Shanker, K., Gupta, M.M., Nagegowda, D.A. and Shasany, A.K. (2013) 4-Coumarate: CoA ligase partitions metabolites for eugenol biosynthesis. *Plant Cell Physiol.*, **54**, 1238–1252.
- Rastogi, S., Meena, S., Bhattacharya, A., et al. (2014) De novo sequencing and comparative analysis of holy and sweet basil transcriptomes. *BMC Genomics*, **15**, 1–18.
- Ravi, D., Siril, E.A. and Nair, B.R. (2021) SCAR marker development for the identification of elite germplasm of *Moringa oleifera* Lam.-A never die plant. *Plant Mol. Biol. Report*. Available at: <https://doi.org/10.1007/s11105-021-01300-y>.
- Rawat, R., Tiwari, V., Singh, R. and Bisht, I.S. (2017) Assessment of the Essential Oil Composition in *Ocimum* species of Uttarakhand. *J. Essent. Oil-Bearing Plants*, **20**, 1331–1341.
- Reddy, V.A., Li, C., Nadimuthu, K., Tjhang, J.G., Jang, I.C. and Rajani, S. (2021) Sweet basil has distinct synthases for eugenol biosynthesis in glandular trichomes and roots with different regulatory mechanisms. *Int. J. Mol. Sci.*, **22**, 1–22.
- Renu, I.K., Haque, I., Kumar, M., Poddar, R., Bandopadhyay, R., Rai, A. and Mukhopadhyay, K. (2014) Characterization and functional analysis of eugenol *O*-methyltransferase gene reveal metabolite shifts, chemotype specific differential expression and developmental regulation in *Ocimum tenuiflorum* L. *Mol. Biol. Rep.*, **41**, 1857–1870.
- Ríos-Rodríguez, D., Sahi, V.P. and Nick, P. (2021) Authentication of holy basil using markers relating to a toxicology - relevant compound. *Eur. Food Res. Technol.* Available at: <https://doi.org/10.1007/s00217-021-03812-z>.
- Rizzini, L., Favory, J.J., Cloix, C., et al. (2011) Perception of UV-B by the *Arabidopsis* UVR8 protein. *Science (80- )*, **332**, 103–106.
- Robin, A.Y., Giustini, C., Graindorge, M., Matringe, M. and Dumas, R. (2016) Crystal structure of norcoclaurine-6-*O*-methyltransferase, a key rate-limiting step in the synthesis of benzyloquinoline alkaloids. *Plant J.*, **87**, 641–653.

- Runyoro, D., Ngassapa, O., Vagionas, K., Aligiannis, N., Graikou, K. and Chinou, I. (2010) Chemical composition and antimicrobial activity of the essential oils of four *Ocimum* species growing in Tanzania. *Food Chem.*, **119**, 311–316. Available at: <http://dx.doi.org/10.1016/j.foodchem.2009.06.028>.
- Ryan, M., Lazar, I., Nadasdy, G.M., Nadasdy, T. and Satoskar, A.A. (2015) Acute kidney injury and hyperbilirubinemia in a young male after ingestion of *Tribulus terrestris*. *Clin. Nephrol.*, **83**, 177–183.
- Sang, T., Crawford, D.J. and Stuessy, T.F. (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *Am. J. Bot.*, **84**, 1120–1136.
- Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*, **9**, 671–675.
- Schwarzerová, K., Bellinvia, E., Martinek, J., et al. (2019) Tubulin is actively exported from the nucleus through the Exportin1/CRM1 pathway. *Sci. Rep.*, **9**, 1–13.
- Sgamma, T., Lockie-Williams, C., Kreuzer, M., Williams, S., Scheyhing, U., Koch, E., Slater, A. and Howard, C. (2017) DNA barcoding for industrial quality assurance. *Planta Med.*, **83**, 1117–1129.
- Shaw, P.C., Wang, J. and But, P.-H. (2002) *Authentication of Chinese Medicinal Materials by DNA Technology* P. C. Shaw, J. Wang, and P. -H. But, eds., WORLD SCIENTIFIC. Available at: <https://doi.org/10.1142/4700>.
- Siddiqui, S., Ahmed, N., Goswami, M., Chakrabarty, A. and Chowdhury, G. (2021) DNA damage by withanone as a potential cause of liver toxicity observed for herbal products of *Withania somnifera* (Ashwagandha). *Curr. Res. Toxicol.*, **2**, 72–81. Available at: <https://doi.org/10.1016/j.crttox.2021.02.002>.
- Singh, P., Kalunke, R.M. and Giri, A.P. (2015) Towards comprehension of complex chemical evolution and diversification of terpene and phenylpropanoid pathways in *Ocimum* species. *RSC Adv.*, **5**, 106886–106904. Available at: <http://dx.doi.org/10.1039/C5RA16637C>.
- Solecki, R.S. (1975) Shanidar IV, a Neanderthal flower burial in northern Iraq. *Science*, **180**, 880–881.
- Spagnolo, A., Mondello, V., Larignon, P., Villaume, S., Rabenoelina, F., Clément, C. and Fontaine, F. (2017) Defense responses in grapevine (Cv. Mourvèdre) after inoculation with the Botryosphaeria dieback pathogens *Neofusicoccum parvum* and *Diplodia seriata* and their relationship with flowering. *Int. J. Mol. Sci.*, **18**, 1–12.
- Spink, J. and Moyer, D.C. (2011) Defining the public health threat of food fraud. *J. Food Sci.*, **76**, 157–163.
- Stoeckle, M.Y., Gamble, C.C., Kirpekar, R., Young, G., Ahmed, S. and Little, D.P. (2011) Commercial teas highlight plant DNA barcode identification successes and obstacles. *Sci. Rep.*, **1**, 1–7.
- Stratmann, J. (2003) Ultraviolet-B radiation co-opts defense signaling pathways. *Trends Plant Sci.*, **8**, 526–533.

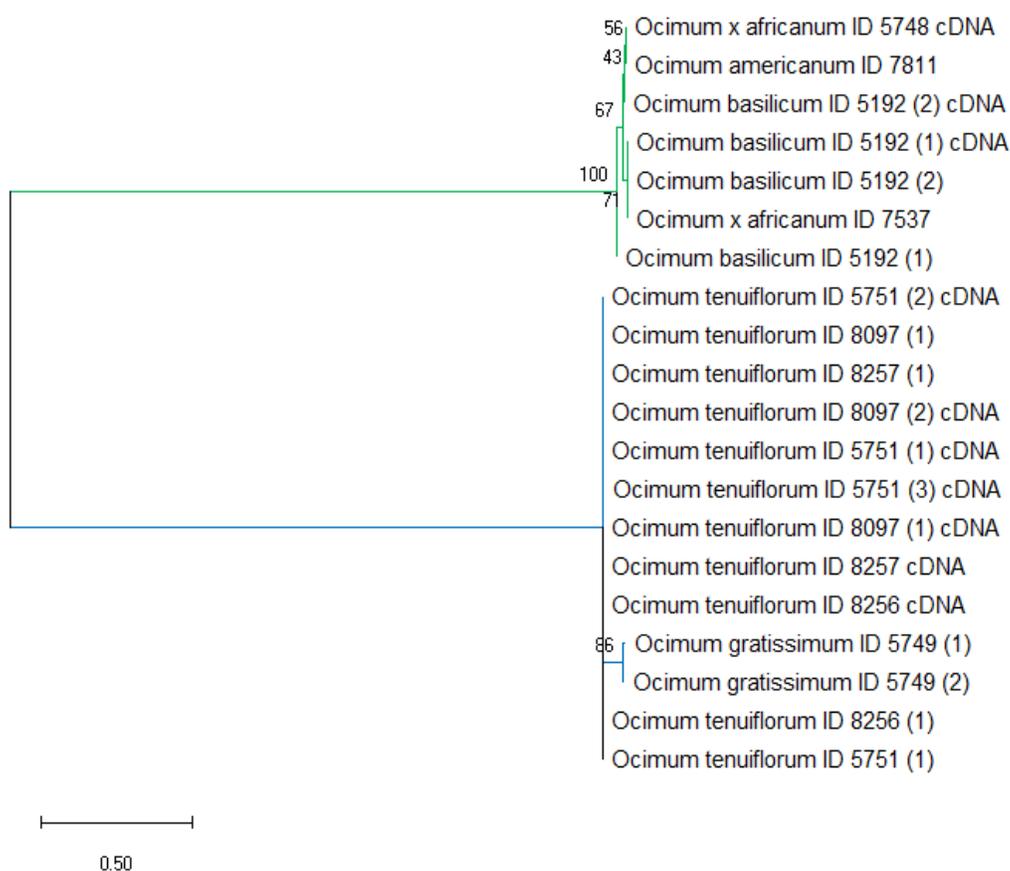
- Sunitha, S., Loyola, R., Alcalde, J.A., Arce-Johnson, P., Matus, J.T. and Rock, C.D. (2019) The role of UV-B light on small RNA activity during grapevine berry development. *G3 Genes, Genomes, Genet.*, **9**, 769–787.
- Tamura, K. and Nei, M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.*, **10**, 512–526.
- Tan, K.H. and Nishida, R. (2012) Methyl eugenol: Its occurrence, distribution, and role in nature, especially in relation to insect behavior and pollination. *J. Insect Sci.*, **12**, 1–74.
- Tate, J.A. and Simpson, B.B. (2003) Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Syst. Bot.*, **28**, 723–737.
- Tichy, H.V., Bruhs, A. and Palisch, A. (2020) Development of real-time polymerase chain reaction systems for the detection of so-called “superfoods” Chia and Quinoa in commercial food products. *J. Agric. Food Chem.*, **68**, 14334–14342.
- Torricelli, R., Silveri, D.D., Ferradini, N., Venora, G., Veronesi, F. and Russi, L. (2012) Characterization of the lentil landrace Santo Stefano di Sessanio from Abruzzo, Italy. *Genet. Resour. Crop Evol.*, **59**, 261–276.
- Tossi, V.E., Regalado, J.J., Iannicelli, J., Laino, L.E., Burrieza, H.P., Escandón, A.S. and Pitta-Álvarez, S.I. (2019) Beyond arabidopsis: differential UV-B response mediated by UVR8 in diverse species. *Front. Plant Sci.*, **10**, 1–17.
- Tuinen, A. Van, Peters, A.H.L.J., Kendrick, R.E., Zeevaert, J.A.D. and Koornneef, M. (1999) Characterisation of the procer mutant of tomato and the interaction of gibberellins with end-of-day far-red light treatments. *Physiol. Plant.*, **106**, 121–128.
- U.S. Food and Drug Administration (2019) Statement from FDA commissioner Scott Gottlieb, M.D., on the Agency’s new efforts to strengthen regulation of dietary supplements by modernizing and reforming FDA’s oversight. **February**, 1–6. Available at: <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm631065.htm>.
- Ulm, R., Baumann, A., Oravec, A., Máté, Z., Ádám, É., Oakeley, E.J., Schäfer, E. and Nagy, F. (2004) Genome-wide analysis of gene expression reveals function of the bZIP transcription factor HY5 in the UV-B response of *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 1397–1402.
- Upadhyay, A.K., Chacko, A.R., Gandhimathi, A., et al. (2015) Genome sequencing of herb Tulsi (*Ocimum tenuiflorum*) unravels key genes behind its strong medicinal properties. *BMC Plant Biol.*, **15**, 1–20. Available at: <http://dx.doi.org/10.1186/s12870-015-0562-x>.
- Vanhaelewyn, L., Straeten, D. Van Der, Coninck, B. De and Vandebussche, F. (2020) Ultraviolet radiation from a plant perspective: the plant-microorganism context. *Front. Plant Sci.*, **11**, 1–18.
- Vogt, T. (2010) Phenylpropanoid biosynthesis. *Mol. Plant*, **3**, 2–20. Available at: <http://dx.doi.org/10.1093/mp/ssp106>.
- Vyas, P., Haque, I., Kumar, M. and Mukhopadhyay, K. (2014) Photocontrol of differential gene expression and alterations in foliar anthocyanin accumulation: A comparative study using red and green forma *Ocimum tenuiflorum*. *Acta Physiol. Plant.*, **36**, 2091–2102.

- Wang, W., Wang, K., Liu, F. and Fang, J. (2012) An efficient identification of 68 apple cultivars using a cultivar identification diagram (CID) strategy and RAPD markers. *Korean J. Hort. Sci. Technol.*, **30**, 549–556.
- Wetters, S., Horn, T. and Nick, P. (2018) Goji who? Morphological and DNA based authentication of a “superfood.” *Front. Plant Sci.*, **871**, 1–14.
- WHO (2019) *WHO Global report on traditional and complementary medicine 2019*, Available at: <https://apps.who.int/iris/bitstream/handle/10665/312342/9789241515436-eng.pdf?ua=1>.
- Wu, K., Yang, M., Liu, H., Tao, Y., Mei, J. and Zhao, Y. (2014) Genetic analysis and molecular characterization of Chinese sesame (*Sesamum indicum* L.) cultivars using Insertion-Deletion (InDel) and Simple Sequence Repeat (SSR) markers. *BMC Genet.*, **15**, 1–15.
- Xie, Z., Kapteyn, J. and Gang, D.R. (2008) A systems biology investigation of the MEP/terpenoid and shikimate/phenylpropanoid pathways points to multiple levels of metabolic control in sweet basil glandular trichomes. *Plant J.*, **54**, 349–361.
- Yang, Yanjun, Guo, J., Cheng, J., et al. (2020) Identification of UV-B radiation responsive microRNAs and their target genes in chrysanthemum (*Chrysanthemum morifolium* Ramat) using high-throughput sequencing. *Ind. Crops Prod.*, **151**, 112484. Available at: <https://doi.org/10.1016/j.indcrop.2020.112484>.
- Yang, Yu, Zhang, L., Chen, P., Liang, T., Li, X. and Liu, H. (2020) UV-B photoreceptor UVR8 interacts with MYB73/MYB77 to regulate auxin responses and lateral root development. *EMBO J.*, **39**, 1–15.
- Yin, R. and Ulm, R. (2017) How plants cope with UV-B: from perception to response. *Curr. Opin. Plant Biol.*, **37**, 42–48. Available at: <http://dx.doi.org/10.1016/j.pbi.2017.03.013>.
- Yokawa, K., Kagenishi, T. and Baluška, F. (2016) UV-B induced generation of reactive oxygen species promotes formation of BFA-induced compartments in cells of *Arabidopsis* root apices. *Front. Plant Sci.*, **6**, 1–10.
- Zhou, W., Shi, M., Deng, C., Lu, S., Huang, F., Wang, Y. and Kai, G. (2021) The methyl jasmonate-responsive transcription factor SmMYB1 promotes phenolic acid biosynthesis in *Salvia miltiorrhiza*. *Hortic. Res.*, **8**. Available at: <http://dx.doi.org/10.1038/s41438-020-00443-5>.
- Złotek, U., Michalak-Majewska, M. and Szymanowska, U. (2016) Effect of jasmonic acid elicitation on the yield, chemical composition, and antioxidant and anti-inflammatory properties of essential oil of lettuce leaf basil (*Ocimum basilicum* L.). *Food Chem.*, **213**, 1–7.
- Zoghbi, M. das G.B., Oliveira, J., Andrade, E.H.A., Trigo, J.R., Fonseca, R.C.M. and Rocha, A.E.S. (2007) Variation in volatiles of *Ocimum campechianum* Mill. and *Ocimum gratissimum* L. cultivated in the north of Brazil. *J. Essent. Oil-Bearing Plants*, **10**, 229–240.

## Appendix

**Table A.1.** Reference plant material used in this research; ID of the voucher specimen cultivated in the Botanical Garden of the KIT; scientific name; Genbank accession numbers of the sequences from the plastidic *psbA-trnH* region, and the partial sequences of the enzyme eugenol *O*-methyltransferase (EOMT).

KIT Accession ID	Species	<i>psbA-trnH</i>	EOMT
5192	<i>O. basilicum</i> L.	MF784535	-
7811	<i>O. americanum</i>	MF784536	-
7537	<i>O. x africanum</i> Lour.	MF784538	-
5748	<i>O. x africanum</i> Lour.	MF784537	-
5751	<i>O. tenuiflorum</i> L.	MF784540	MW582310
8097	<i>O. tenuiflorum</i> L.	MF784541	-
8099	<i>O. tenuiflorum</i> L.	MF784542	-
8256	<i>O. tenuiflorum</i> L.	MF784543	-
8257	<i>O. tenuiflorum</i> L.	MF784544	-
8258	<i>Ocimum</i> sp.	MF784545	-
9056	<i>O. basilicum</i> L.	MW582309	MW582311
5749	<i>O. gratissimum</i>	MF784560	-
7564	<i>O. campechianum</i> Mill.	MF784557	-
7809	<i>O. kilimandscharicum</i>	-	-
7810	<i>O. kilimandscharicum</i>	MF784539	-
5391	<i>M. spicata</i> var. <i>crispa</i>	MH753570	-
5393	<i>M. x piperita</i>	MH753571	-

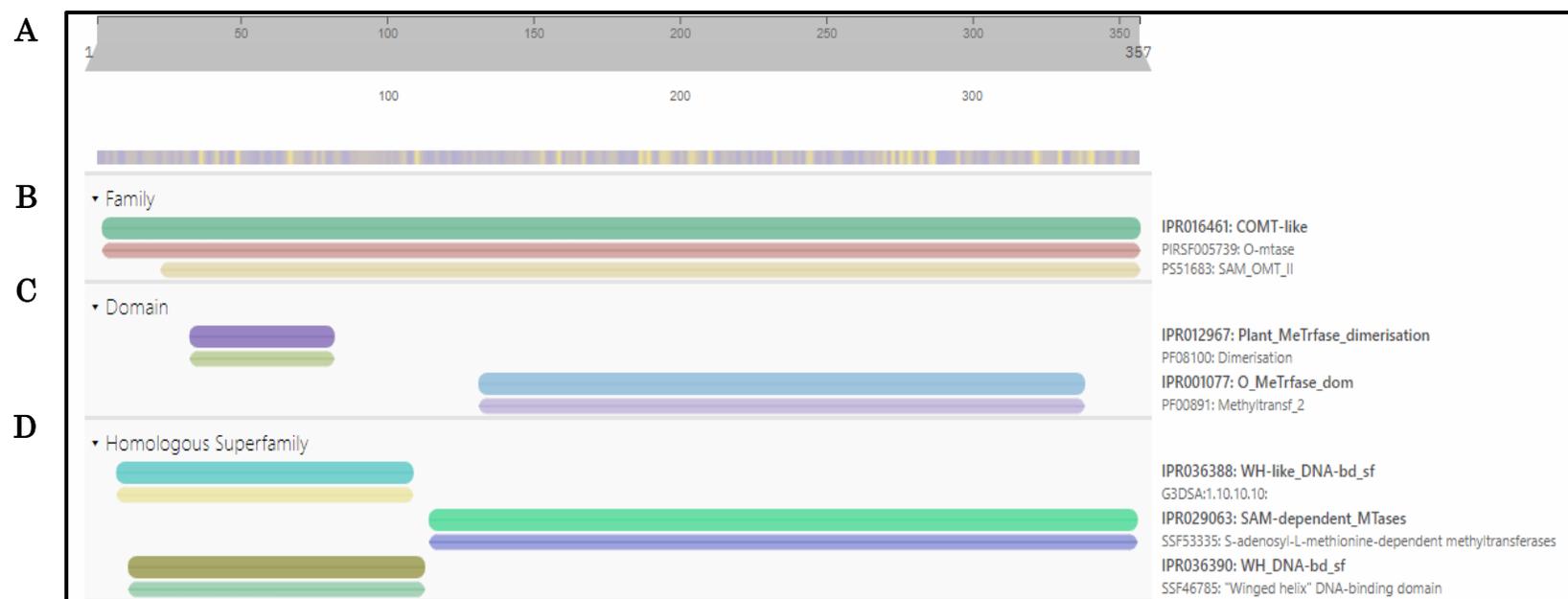


**Figure A.1** Phylogeny of *Ocimum* species using the nucleotide sequence from the eugenol *O*-methyltransferase enzyme from exons (gDNA) and cDNA. Blue branches indicate the Tulsi clade (*O. tenuiflorum* and *O. gratissimum*), and green branches indicate the Basil clade (*O. basilicum*, *O. americanum*, *O. africanum* and *O. kilimandscharicum*). Neighbour-joining distant tree, bootstrap values (based on 1000 replicates) next to the branches. For details on the accessions refer to **Table 2.1**

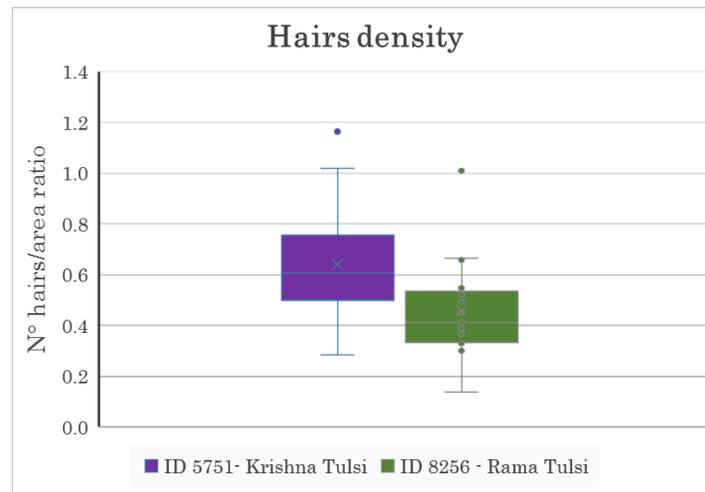


**Table A.2.** Motifs found in different amino acid sequences of eugenol *O*-methyltransferase from *Ocimum* species. Amino acid residue positions are in reference to EOMT positions from *O. tenuiflorum*. Motifs are compared to *O*-methyltransferase superfamily suggested by Ibrahim et al. (1998). Blue coloured *Ocimum* species are accessions from the Tulsi clade (see **Figure 3.1**), green coloured *Ocimum* species are accessions from the Basil clade (see **Figure 3.1**). Blocks of yellow amino acid residues indicate the motifs I to V from the described for OMT by Ibrahim et al. (1998). Amino acid residues marked in orange tones indicate differences among sequences. Amino acid residue in position 261 coloured in purple indicates mutagenic amino acid.

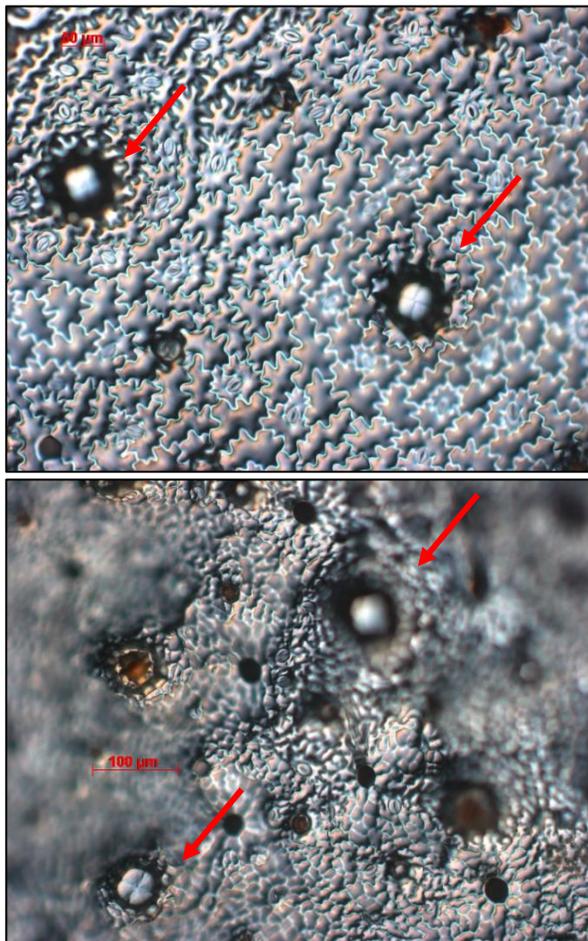
<i>Ocimum</i> sp. / aa position in <i>Ocimum tenuiflorum</i>	199	200	201	202	203	204	205	206	207	208	216	217	220	222	223	224	225	226	227	228	229	230	239	241	242	243	244	245	246	247	248	249	257	261	267	272	285	286	287	288	289	299	320	321	322	323	324	325		
<i>O. tenuiflorum</i> ID 5751 (1) cDNA	L	V	D	V	G	G	G	D	G	T	T	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 5751 (2) cDNA	L	V	D	V	G	G	G	D	G	T	A	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 5751 (3) cDNA	L	A*	D	V	G	G	G	D	G	T	T	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 8097 (1) cDNA	L	V	D	V	G	G	G	D	G	T	T	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 8097 (2) cDNA	L	V	D	V	G	G	G	D	G	T	T	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 8257 cDNA	L	V	D	V	G	G	G	D	G	T	A	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 8256 cDNA	L	V	D	V	G	G	G	D	G	T	A	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 5751	L	V	D	V	G	G	G	D	G	T	T	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 8097	L	V	D	V	G	G	G	D	G	T	T	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 8257	L	V	D	V	G	G	G	D	G	T	A	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. tenuiflorum</i> ID 8256	L	V	D	V	G	G	G	D	G	T	A	V	I	C	T	V	I	D	L	P	H	V	S	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	Q	N	A	K	E	R	T		
<i>O. gratissimum</i> ID 5749 (1)	L	V	D	V	G	G	G	N	G	T	A	V	M	C	T	V	I	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	H	N	A	K	E	R	T		
<i>O. gratissimum</i> ID 5749 (2)	L	V	D	V	G	G	G	N	G	T	A	V	M	C	T	V	I	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	R	V	S	N	L	G	G	K	V	I	H	N	A	K	E	R	T		
<i>O. basilicum</i> ID 5192 (1) cDNA	L	V	D	V	G	G	G	N	G	T	A	M	M	C	T	V	L	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	Q	I	S	D	I	G	G	K	V	I	H	N	A	K	E	R	T		
<i>O. basilicum</i> ID 5192 (2) cDNA	L	V	D	V	G	G	G	N	G	S	A	V	M	C	T	V	L	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	Q	I	S	D	L	G	G	K	V	I	H	N	A	K	E	R	T		
<i>O. x africanum</i> ID 5748 cDNA	L	V	D	V	G	G	G	N	G	S	A	V	M	C	T	V	L	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	Q	I	S	D	L	G	G	K	V	I	H	N	A	K	E	R	T		
<i>O. basilicum</i> ID 5192 (1)	L	V	D	V	G	G	G	N	G	S	A	V	M	C	T	V	L	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	Q	I	S	D	I	G	G	K	V	I	H	N	A	K	E	R	T		
<i>O. basilicum</i> ID 5192 (2)	L	V	D	V	G	G	G	N	G	T	A	M	M	C	T	V	L	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	Q	I	S	D	I	G	G	K	V	I	H	N	A	K	E	R	T		
<i>O. americanum</i> ID 7811	L	V	D	V	G	G	G	N	G	S	A	V	M	C	T	V	L	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	Q	I	S	D	L	G	G	K	V	I	H	N	A	K	E	R	T		
<i>O. x africanum</i> ID 7537	L	V	D	V	G	G	G	N	G	T	A	M	M	C	T	V	L	D	L	P	H	V	R	S	Y	I	G	G	D	M	F	Q	I	S	D	I	G	G	K	V	I	H	N	A	K	E	R	T		
OMT Motifs in <i>Ocimum</i> sp.	I										II										III										IV										V									
	L	V*	D	V	G	G	G	x	G	C	T	V	x	D	L	P	H	V	S	Y	I	G	G	D	M	F	G	G	K	V	I	N	A	K	E	R	T													
OMT Motifs by Ibrahim et al (1998)	L V D V G G G x G										G I N F D L P H V										E H V G G D M F										N G K V I										G G K E R T									



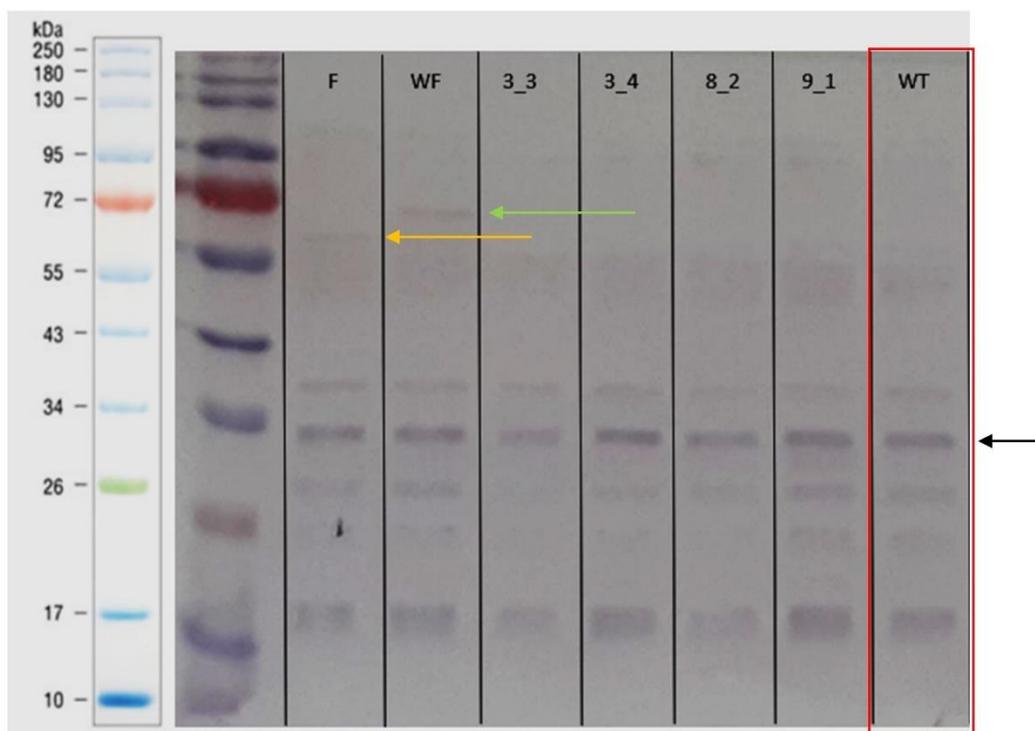
**Figure A.3.** Summary of results matches for the analysis run using the amino acidic sequence of EOMT from *O. tenuiflorum* in UniProt ([www.uniprot.org](http://www.uniprot.org)). **a** Length of the enzyme: 357 aa residues. **b** Family similarity with other O-methyltransferase. **c** Domains matches: dimerization domain and O-methyltransferase domain. **d** Homologous superfamilies of proteins such as S-adenosyl methionine (SAM) dependent methyltransferases.



**Figure A.4.** Hairs density analysis results for two *Ocimum tenuiflorum* chemotypes, Krishna Tulsi (ID 5751) and Rama Tulsi (ID 8256), showing that Krishna has higher number of hairs per area than Rama.



**Figure A.5.** Images of adaxial side of *O. tenuiflorum* leaves showing peltate glands (red arrows). Krishna Tulsi (upper) and Rama Tulsi (lower)



**Figure A.7.** Western blot analysis results from transformed BY-2 cells expressing the enzyme eugenol *O*-methyltransferase fused with green fluorescence protein (GFP). Green arrow shows a band at about 70kDa representing the EOMT-*W* (N-terminally fused GFP to EOMT). Orange arrow shows a band at about 70kDa representing the EOMT-*F* (C-terminally fused GFP to EOMT). Red box indicates unspecific pattern in western blot for BY-2 wild type cells in this experiment. Black arrow shows the GFP band. For details on the BY-2 transformed cells see section 3.3.