



HAL
open science

Sclareol biosynthesis in clary sage and its regulation

Camille Chalvin

► **To cite this version:**

Camille Chalvin. Sclareol biosynthesis in clary sage and its regulation. Plant breeding. Université Paris Saclay (COMUE), 2019. English. NNT : 2019SACLS194 . tel-02426233

HAL Id: tel-02426233

<https://theses.hal.science/tel-02426233>

Submitted on 2 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sclareol biosynthesis in clary sage and its regulation

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris-Sud

École doctorale n°567
Sciences du Végétal : du gène à l'écosystème (SDV)
Spécialité de doctorat: Biologie

Thèse présentée et soutenue à Gif-sur-Yvette, le 12/07/2019, par

Camille Chalvin

Composition du Jury :

Marianne Delarue Professeure, Université Paris-Sud (IPS2)	Présidente
Sylvie Baudino-Caissard Professeure, Université Jean Monnet, Saint-Etienne (LBVpam)	Rapporteur
Andréa Hemmerlin Chargée de recherche, CNRS (IBMP, Strasbourg)	Rapporteur
Johannes Panten Responsable R&D, Symrise AG	Examineur
Françoise Granet Manager d'expertise, Michelin	Examineur
Michel Dron Professeur émérite, Université Paris-Sud (IPS2)	Directeur de thèse
Adnane Boualem Directeur de recherche, INRA (IPS2)	Invité
Stéphanie Drevensek Chargée de projet, Adeprina (IPS2)	Invité

Table of contents

Table of contents.....	1
Résumé en français.....	7
1. Introduction.....	7
2. Résultats.....	7
2.1. Localisation de la biosynthèse du sclaréol.....	7
2.2. Identification de l'origine métabolique du sclaréol par marquage isotopique.....	8
2.3. La biosynthèse du sclaréol est-elle régulée par les jasmonates ?.....	9
2.4. Etude de la diversité naturelle de la sauge sclarée.....	9
3. Revue de la littérature.....	10
4. Perspectives.....	11
Introduction.....	13
1. Ambergris and sclareol.....	13
1.1. The search for substitutes to ambergris.....	13
1.2. Interest in increasing sclareol global production.....	14
2. Clary sage.....	15
2.1. Botanical description of clary sage.....	15
2.2. Sclareol biological function in clary sage.....	16
2.3. Clary sage commercial use.....	17
2.4. First advances in clary sage genetics and biotechnology.....	19
2.5. Clary sage breeding in France.....	20
3. Terpene biosynthesis and function in plants.....	21
3.1. Specialized metabolism and terpenoids.....	21
3.2. The diverse biological functions of terpenes.....	22
3.3. Industrial exploitation of terpenes.....	24
3.4. Terpene biosynthesis.....	25
3.5. Terpene biosynthesis regulation.....	32
4. Glandular trichomes.....	35
4.1. Capitata and peltate glandular trichomes of the Lamiaceae.....	35
4.2. Clary sage glandular trichomes.....	37
4.3. Genetic control of glandular trichome development.....	38

5.	Terpene production by clary sage	39
5.1.	Essential oil production in aerial parts.....	39
5.2.	Sclareol biosynthesis and secretion	40
5.3.	Terpene production in roots.....	42
6.	Aim of the project.....	43
Chapter 1	45
Mass spectrometry imaging to localize sclareol biosynthesis	45
1.	Introduction.....	45
2.	Materials and methods	47
2.1.	Plant material	47
2.2.	Scanning electron microscopy	47
2.3.	Metabolite localization by LDI-FT-ICR.....	47
2.4.	Sclareol and linalyl acetate quantification by GC-MS	48
3.	Results.....	49
3.1.	Different accumulation patterns for sclareol and linalyl acetate.....	49
3.2.	Distinct roles of different trichome types in sclareol production	52
3.3.	A clary sage glabrous mutant produces less sclareol and linalyl acetate	56
4.	Discussion.....	58
4.1.	Sclareol is mainly produced by glandular trichomes	58
4.2.	Some sclareol could also be produced by other epidermal cells	60
4.3.	Engineering higher glandular trichome density to increase sclareol yield	61
Chapter 2	63
A retro-biosynthetic approach to decipher sclareol metabolic origin	63
1.	Introduction.....	63
1.1.	Metabolic crosstalk between MVA and MEP pathways	63
1.2.	Towards clary sage metabolic engineering.....	64
2.	Materials and methods	66
2.1.	Plant material	66
2.2.	Treatment with MVA or MEP pathway-specific inhibitors.....	66
2.3.	Sclareol and linalyl acetate quantification by GC-MS	67
2.4.	Isotope labeling and analysis by ¹³ C-NMR or GC-MS	67
3.	Results.....	71
3.1.	Sclareol and linalyl acetate are highly abundant terpenes of clary sage calyx surface	71
3.2.	Sclareol and linalyl acetate are MEP-derived terpenes	73

3.3.	Comparison of results obtained for sclareol and linalyl acetate	79
3.4.	The case of the acetate group of linalyl acetate	80
3.5.	The mixed origin of the sesquiterpene β -caryophyllene	82
4.	Discussion.....	84
4.1.	Biosynthetic origin of mono-, sesqui- and diterpenes in clary sage and other plants.....	84
4.2.	Metabolic engineering of terpene production in plants	85
4.3.	Perspectives for clary sage metabolic engineering	88
Chapter 3	93
Is sclareol biosynthesis regulated by jasmonates?		93
1.	Introduction.....	93
1.1.	Regulation of plant specialized metabolism by jasmonates	93
1.2.	How to use jasmonate signaling to enhance the production of compounds of interest ..	94
1.3.	Regulation of clary sage specialized metabolism by jasmonates: state of the art	95
2.	Materials and methods	97
2.1.	Calyx development stages	97
2.2.	MeJA treatment by feeding	97
2.3.	MeJA treatment by spraying.....	97
2.4.	Sclareol and linalyl acetate quantification by GC-MS	98
3.	Results.....	99
3.1.	Regulation of sclareol biosynthesis during calyx development	99
3.2.	MeJA treatment on flowers by feeding	101
3.3.	MeJA treatment on flowers by spraying.....	105
4.	Discussion.....	105
4.1.	MeJA apparently does not impact sclareol production, but further protocol optimization is needed	105
4.2.	Other potential ways to manipulate sclareol biosynthesis regulation.....	107
Chapter 4	109
Clary sage natural diversity		109
1.	Introduction.....	109
1.1.	The use of natural genetic diversity for breeding	109
1.2.	Impact of environmental conditions on terpene production and genetic diversity	110
2.	Materials and methods	111
2.1.	Plant material	111
2.2.	Genetic diversity analysis	113

2.3.	Field trial.....	113
2.4.	Sclareol and linalyl acetate quantification by GC-MS	114
3.	Results.....	115
3.1.	Croatian populations show genetic polymorphism in MEP pathway genes	115
3.2.	Choice of the populations to be analyzed in the field.....	117
3.3.	Comparison of flowering date and metabolite content of the different populations	119
4.	Discussion.....	122
4.1.	Complex relationships between genetic diversity, chemical diversity and environmental conditions.....	122
4.2.	Perspectives for the study of clary sage natural diversity.....	124
4.3.	Vatican White plants flower earlier and produce less specialized metabolites	125
Conclusion		127
1.	Context and objectives of the project	127
2.	Perspectives for research on sclareol production in clary sage	127
3.	Perspectives for clary sage genetic improvement	129
Annex.....		131
Review: Genetic control of glandular trichome development.....		131
1.	Introduction.....	131
1.1.	Why study the genetic control of glandular trichome development?	131
1.2.	Which species could serve as model(s)?.....	132
2.	Description of glandular trichome morphology and development	133
3.	Genes controlling glandular trichome initiation	135
3.1.	Transcription factors.....	135
3.2.	Cyclins.....	137
3.3.	Regulatory complexes	139
3.4.	Genes involved in hormonal signaling	140
4.	Genes controlling later steps of glandular trichome development	140
4.1.	Cytoskeleton regulators.....	140
4.2.	Cuticle deposition regulators.....	141
5.	Evolution of glandular trichome development regulators.....	142
5.1.	Evidence supporting the conservation of glandular trichome development regulators in the Solanaceae	142
5.2.	Some regulators appear to be conserved in distant plant families, while the others are likely to have evolved independently	142
5.3.	Different scenarios for the evolution of glandular trichome development regulators ...	143

6. Perspectives	144
References	147

Résumé en français

1. Introduction

Le sclaréol est un terpène produit par les organes floraux de la sauge sclarée (*Salvia sclarea*), en particulier par les calices. Il est utilisé pour l'hémisynthèse de l'ambroxide, un composé très apprécié en parfumerie pour son odeur ambrée et ses propriétés fixatrices augmentant la ténacité des parfums. Le sclaréol utilisé pour produire de l'ambroxide est extrait à partir de sauge sclarée cultivée. En réponse au développement de l'utilisation de l'ambroxide en parfumerie fonctionnelle (détergents, shampoings...), la demande mondiale de sclaréol est actuellement en hausse, stimulant la recherche de moyens permettant d'en accroître la production.

Le projet PISTILL, porté par l'équipe « Développement des fleurs et des carpelles » de l'Institut de Sciences des Plantes de Paris-Saclay, a pour objectif d'augmenter le rendement de la production de sclaréol via l'amélioration génétique ciblée de la sauge sclarée. Le travail présenté dans ce manuscrit visait à compléter les connaissances actuelles sur la biosynthèse du sclaréol et sa régulation chez la sauge sclarée, avec un double objectif: d'une part, orienter le projet PISTILL vers des stratégies prometteuses pour l'augmentation du rendement en sclaréol, et d'autre part, améliorer notre compréhension générale de la production des terpènes chez les plantes. Nous avons entrepris de préciser l'origine cellulaire et métabolique du sclaréol dans les calices de sauge sclarée, et d'identifier les mécanismes de régulation contrôlant son accumulation. Nous avons également commencé à étudier comment la diversité naturelle de la sauge sclarée pourrait être exploitée. Cette étude ouvre des perspectives intéressantes pour la recherche fondamentale sur la production des terpènes chez les plantes, tout en mettant en évidence des pistes prometteuses pour l'amélioration génétique ciblée des performances de la sauge sclarée.

2. Résultats

2.1. Localisation de la biosynthèse du sclaréol

L'épiderme des calices de sauge sclarée est caractérisé par une forte densité de trichomes glandulaires. Chez de nombreuses espèces de plantes, ces structures multicellulaires sont

spécialisées dans la production de terpènes (Lange and Turner, 2013). La présence de sclaréol a été détectée dans les différents types de trichomes glandulaires de la sauge sclarée (peltés et capités) (Schmiderer et al., 2008). Cependant, d'autres cellules de l'épiderme du calice pourraient également sécréter du sclaréol (Caissard et al., 2012). Dans le premier chapitre de ce manuscrit, nous présentons plusieurs éléments appuyant l'hypothèse selon laquelle le sclaréol est principalement sécrété par les trichomes glandulaires. La répartition spatiale du sclaréol à la surface des calices a été analysée par imagerie par spectrométrie de masse, et les résultats indiquent que le sclaréol est particulièrement abondant au niveau des trichomes glandulaires capités. De plus, des mutants de sauge sclarée dépourvus de ces trichomes (obtenus via une mutagenèse à l'EMS) produisent moins de sclaréol. Dans le futur, l'identification de la mutation (ou des mutations) à l'origine de ce phénotype améliorerait notre compréhension du développement des trichomes glandulaires et de la production du sclaréol chez la sauge sclarée.

2.2. Identification de l'origine métabolique du sclaréol par marquage isotopique

La biosynthèse des terpènes débute par la condensation d'un nombre variable d'unités à cinq carbones : l'isopentényl diphosphate (IPP) et son isomère le diméthylallyl diphosphate (DMAPP). Chez les plantes, l'IPP et le DMAPP sont produits par deux voies métaboliques différentes : la voie du mévalonate dans le cytoplasme (voie MVA) et la voie du méthylérythritol-phosphate dans les chloroplastes (voie MEP). L'activité de ces deux voies alimente deux pools d'IPP et de DMAPP physiquement séparés au sein de la cellule. Selon la localisation subcellulaire des terpènes-synthases impliquées, la production d'un terpène donné ne dépend généralement que de l'un de ces deux pools (Lipko and Swiezewska, 2016). Lorsqu'elles sont exprimées de manière transitoire dans des feuilles de tabac (*Nicotiana benthamiana*), les terpènes-synthases impliquées dans la biosynthèse du sclaréol sont adressées aux chloroplastes (Caniard et al., 2012), le compartiment où la voie MEP est localisée. Cependant, l'origine métabolique du sclaréol n'avait pas encore été étudiée. Les approches développées dans le deuxième chapitre de ce manuscrit mettent en évidence les contributions respectives des voies MVA et MEP à la biosynthèse de trois terpènes dans les calices de sauge sclarée : le sclaréol, l'acétate de linalyle et le β -caryophyllène. Nous avons

réalisé un marquage isotopique en fournissant du glucose-1-¹³C à des tiges d'inflorescence coupées. L'analyse des résultats par RMN du ¹³C indique que le sclaréol et l'acétate de linalyle proviennent tous deux de la voie MEP, comme la plupart des mono- et diterpènes (Reed and Osbourn, 2018). En revanche, le β-caryophyllène semble être d'origine mixte, alors que la plupart des sesquiterpènes proviennent de la voie MVA (Reed and Osbourn, 2018) ; mais ce résultat nécessite d'être confirmé.

2.3. La biosynthèse du sclaréol est-elle régulée par les jasmonates ?

La biosynthèse des terpènes est stimulée en réponse à des signaux développementaux ou environnementaux via l'activation de voies de signalisation spécifiques. Le renforcement de l'activité de ces voies de signalisation représente une autre possibilité d'accroître la production d'un terpène d'intérêt (Lu et al., 2016). Cette stratégie nécessite l'identification préalable des voies de signalisation stimulant la biosynthèse du terpène d'intérêt. Étant donné que la production de nombreux terpènes est activée par les jasmonates (Colinas and Goossens, 2018; Wasternack and Strnad, 2019), le troisième chapitre de ce manuscrit examine l'impact de ces phytohormones sur la production de sclaréol dans les calices de sauge sclarée. Du méthyljasmonate en solution a été fourni à des tiges d'inflorescence coupées ou appliqué sur plante entière par spray. Cependant, les expériences réalisées n'ont pas permis de démontrer une implication des jasmonates dans la régulation de la production de sclaréol. Deux possibilités restent plausibles: soit les jasmonates ne sont pas impliqués dans la régulation de la biosynthèse de sclaréol, soit le protocole expérimental doit encore être optimisé.

2.4. Etude de la diversité naturelle de la sauge sclarée

Les conditions environnementales et les polymorphismes génétiques sont responsables de variations de la teneur en terpènes entre différentes populations de la même espèce végétale. La pression de sélection exercée par les conditions de croissance peut conduire à la sélection de modifications génétiques modulant les niveaux de terpènes et rendant la plante plus adaptée à son environnement (Figueiredo et al., 2008; Moore et al., 2014). La sauge sclarée est présente à l'état sauvage dans divers types d'environnements du bassin méditerranéen et d'Asie occidentale (Wagner et al., 2012; Nasermodeli and Rowshan,

2013). Les populations de sauge sclarée cultivées ne représentent probablement qu'un échantillon de la diversité naturelle de la sauge sclarée en termes de polymorphismes génétiques et de teneur en sclaréol. Le quatrième chapitre de ce manuscrit explore la diversité génétique et phénotypique de populations sauvages croates originaires de trois types d'environnements différents: côte, montagne et plateau. Bien que des polymorphismes génétiques soient détectables dans les gènes de biosynthèse du sclaréol, les populations croates cultivées dans le même environnement ne diffèrent pas significativement en termes de teneur en sclaréol et de date de floraison. Les conditions climatiques croates n'ont apparemment pas entraîné de différenciation génétique affectant la production de sclaréol et la floraison. Cependant, la différence de teneur en sclaréol observée entre les populations de référence « Vatican White » et « Milly » mérite une étude plus approfondie. En effet, l'identification de la cause de cette différence pourrait mener à l'identification de mécanismes de régulation de la production de sclaréol.

3. Revue de la littérature

Les trichomes glandulaires des plantes sont des structures épidermiques dédiées à la sécrétion de divers métabolites spécialisés. Ces métabolites contribuent à l'adaptation des plantes à leur environnement et beaucoup d'entre eux possèdent des propriétés exploitées par différents secteurs de l'industrie (parfumerie, pharmacie...). L'identification du réseau de gènes contrôlant le développement des trichomes glandulaires permettrait de mieux comprendre comment les plantes produisent ces métabolites spécialisés. Nos connaissances sur ce processus développemental sont encore limitées, mais des gènes contrôlant l'initiation et la morphogenèse du trichome glandulaire ont récemment été identifiés. La revue de la littérature présentée dans ce manuscrit a pour objectif de synthétiser ces découvertes récentes. D'après notre analyse, les facteurs de transcription appartenant aux sous-familles R2R3-MYB et HD-ZIP IV semblent jouer un rôle essentiel dans l'initiation du trichome glandulaire chez *Artemisia annua* et la tomate. Dans notre revue, nous nous concentrons sur les résultats obtenus chez ces deux espèces et proposons de premiers modèles de régulation génétique intégrant ces données.

4. Perspectives

Plusieurs stratégies visant à accroître la production de sclaréol chez la sauge sclarée peuvent être envisagées à partir des résultats décrits dans ce manuscrit. Le sclaréol étant dérivé de la voie MEP et principalement produit dans les trichomes glandulaires, l'augmentation de la densité des trichomes glandulaires ou la stimulation de la voie MEP pourraient permettre d'augmenter le contenu en sclaréol. De plus, étant donné que l'acétate de linalyle est également dérivé de la voie MEP et produit dans les trichomes glandulaires, bloquer la biosynthèse de l'acétate de linalyle pourrait augmenter la production de sclaréol en réduisant la compétition pour le substrat. La voie du MEP et la biosynthèse de l'acétate de linalyle pourraient être manipulées au moyen de modifications génétiques ciblant les enzymes impliquées dans ces voies. De même, la densité des trichomes glandulaires pourrait être augmentée via la manipulation des régulateurs de l'initiation des trichomes glandulaires. Par ailleurs, les programmes de sélection de la sauge sclarée pourraient être enrichis par une exploitation efficace de la diversité naturelle de la sauge sclarée. Les populations croates étudiées dans le quatrième chapitre de ce manuscrit représentent une ressource génétique précieuse, car nous avons montré qu'elles portent de nouveaux allèles par rapport aux populations de référence.

La faisabilité de ces approches dépendra de la mise au point d'outils dédiés à l'amélioration génétique de la sauge sclarée. Le séquençage du génome de la sauge sclarée et le développement d'outils biotechnologiques spécifiques à cette espèce seraient très utiles pour les projets de recherche fondamentale et appliquée visant à comprendre et à améliorer cette plante à parfum.

Introduction

1. Ambergris and sclareol

1.1. The search for substitutes to ambergris

Ambergris is a waxy substance secreted by the digestive tract of male sperm whales (*Physeter macrocephalus*) upon mucosa injuries caused by squid beaks (**Figure 1**). It is then expelled by the animal as viscous lumps which float on the sea and blocks of 10 g to 1 kg can eventually be found on beaches (Caissard et al., 2012). After several years of aging and exposure to sunlight, air and water, ambergris acquires a musky and sweet earthy smell (Barrero et al., 1993). It has been used for centuries in perfume manufacture for its distinct odor and it also possesses fixative properties which increase the tenacity of perfume compositions (Barrero et al., 1993; Caniard et al., 2012; Caissard et al., 2012). Ambergris properties are mainly due to one of its constitutive components, ambroxide (Ambrox®), a product of ambrein autoxidation occurring during ambergris aging (Barrero et al., 1996; Leffingwell and Leffingwell, 2015) (**Figure 1**). The decrease in sperm whale populations increased ambergris scarcity and prompted the search for synthetic alternative routes to ambroxide. Total synthesis protocols have been elaborated, as well as hemisynthesis strategies starting from various naturally occurring terpenes (Barrero et al., 1993, 1996).

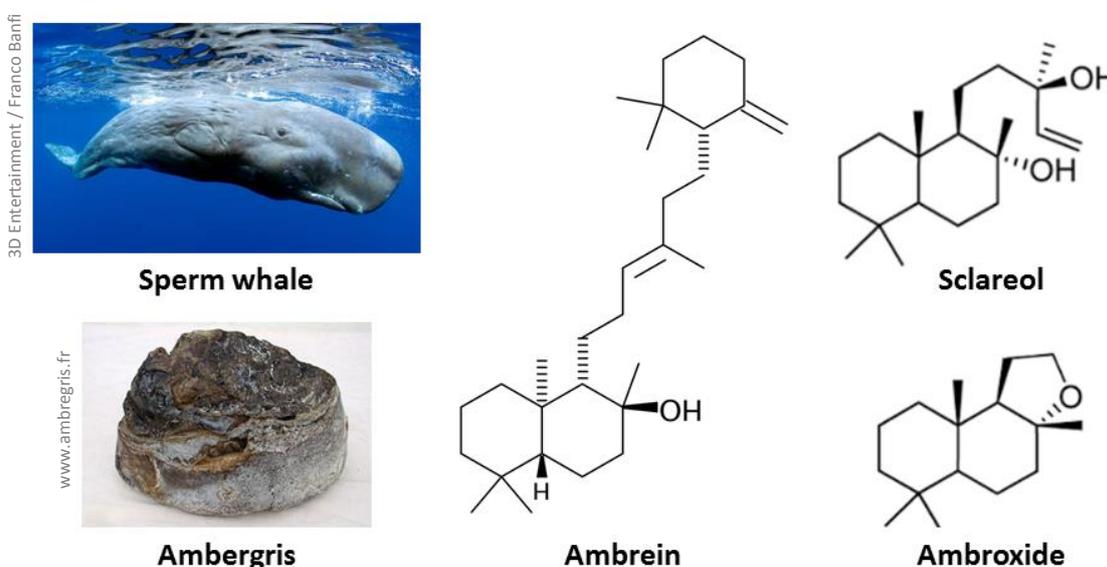


Figure 1: Chemical formulas of ambrein, ambroxide and sclareol.

Sclareol is a natural diterpene which was originally characterized in clary sage (*Salvia sclarea*). It is present in at least five other species of the genus *Salvia* (Çulhaoglu et al., 2013; Moridi Farimani and Miran, 2014; Rapposelli et al., 2015; Park et al., 2016; Craft et al., 2017; Mofidi Tabatabaei et al., 2017) and has been reported in species belonging to other plant families, namely *Cistus creticus* (Cistaceae), *Nicotiana glutinosa* (Solanaceae), *Cleome spinosa* (Brassicaceae) (Caniard et al., 2012; Mastino et al., 2018) and two species of the genus *Pseudognaphalium* (Asteraceae) (Mendoza et al., 2002). Due to its sweet and balsamic scent, sclareol is employed as flavoring in food and tobacco industry and as perfume ingredient in cosmetic industry (Ample Market Research, 2019). Moreover, its labdane-type hydrocarbon skeleton and its hydroxyl groups are particularly attractive for organic synthesis of labdane-derived compounds like ambroxide (Caniard et al., 2012) (**Figure 1**). Today, hemisynthesis from sclareol is the most commonly used synthetic route to ambroxide. The optimization of this route is still an object of research and a one-pot synthesis of ambrox starting from sclareol has been recently described (Yang et al., 2016). Sclareol-derived ambroxide is widely used in perfume manufacture as a substitute to ambergris (Caniard et al., 2012).

1.2. Interest in increasing sclareol global production

The sclareol market is mainly driven by the growing demand of cosmetic industry, which already accounts for around 45.8% of the global consumption of sclareol (Ample Market Research, 2019). Sclareol-derived ambroxide has been long employed in the formulation of high-end perfumes. However, it is more and more found among the ingredients of functional products like soap, detergent or beauty care products, where its fixative properties are also particularly desirable (source: Symrise). The large size of functional perfumery market compared to luxury perfumery drives the current increase in sclareol demand. Estimated to be worth 62 million US\$ in 2019, the worldwide market for sclareol is expected to grow at a compound annual growth rate of roughly 5.5% over the next five years, and may reach 86 million US\$ in 2024 (Ample Market Research, 2019).

Today, sclareol is mainly produced through extraction from cultivated clary sage (Caniard et al., 2012; Leffingwell and Leffingwell, 2015). Clary sage produces relatively high amounts of sclareol compared to other plants, but the yield is submitted to important variations which

remain unexplained (Caniard et al., 2012; Caissard et al., 2012). The raising global demand for sclareol prompted various initiatives aiming at increasing sclareol production and securing the supply. Research programs have been conducted during the past decade to develop biotechnological production of sclareol using bioengineered yeast or moss cell cultures (Schalk, 2009; Simonsen et al., 2014). Since growing clary sage and extracting sclareol is a relatively easy and well established process, the genetic improvement of clary sage is another promising strategy which could rapidly enhance sclareol production. In France, the Interprofessional Technical Institute of Perfume, Medicinal, Aromatic and Industrial Plants (ITEIPMAI) is currently conducting a clary sage breeding program and has already obtained several improved varieties (see page 20). The development of new clary sage varieties with enhanced sclareol yield is the purpose of the PISTILL project driven by the Flower and Carpel Development Group at the Institute of Plant Sciences Paris-Saclay. The work presented here is intended to improve our understanding of sclareol biosynthesis and of mechanisms regulating sclareol production in clary sage. This knowledge will ultimately guide plant breeding approaches developed in the context of the PISTILL project.

2. Clary sage

2.1. Botanical description of clary sage

Salvia is the largest genus of the Lamiaceae family of plants with almost 1000 species described (Hao et al., 2015; Foutami et al., 2018). It is widely spread around the world, as *Salvia* species are found in both temperate and subtropical regions (Ozdemir and Senel, 1999). Clary sage (*Salvia sclarea* L.) is a biennial or short-living perennial herbaceous plant species of the genus *Salvia* (Lattoo et al., 2006; Wagner et al., 2012; Nasermodeli and Rowshan, 2013). It is a native species of the Mediterranean basin and Western Asia (Wagner et al., 2012; Nasermodeli and Rowshan, 2013) and has become an invasive species in North America, where it is known as European sage (Foutami et al., 2018). The plant generally grows in dry and rocky places (Ozdemir and Senel, 1999; Wagner et al., 2012) (**Figure 2**), preferably sunny (Kumar et al., 2013). It develops a rosette which survives the winter and then flowers in June-July (Wagner et al., 2012). White, pink or purple flowers are grouped verticillately in inflorescences of the heterothetic compound raceme type (Ozdemir and Senel, 1999). One plant generally develops several floral scapes, each of them measuring

between 50 and 150 cm (Lattoo et al., 2006; Nasermoadeli and Rowshan, 2013) (**Figure 2**). Insect-driven cross-pollination is important for clary sage reproduction, albeit not indispensable. Indeed, in the genus *Salvia*, the absence of a genetic incompatibility system enables autogamy (Claßen-Bockhoff et al., 2004). Cross-pollination is mainly performed by carpenter bees (*Xylocopa violacea*), which are strongly attracted by clary sage flowers (Kugler, 1972).



Figure 2: Wild clary sage growing in the region of Makarska, Croatia (July, 2016).

2.2. Sclareol biological function in clary sage

The biological function of sclareol in clary sage remains elusive. Sclareol is mainly produced in clary sage flowers, especially calyces (Caissard et al., 2012) (**Figure 3**). Calyces are tubular structures formed by the fusion of sepals and are generally considered to play a role in the protection of flower reproductive organs (Legrand et al., 2010). Some insects are known to pierce holes in the calyx to access nectar without pollination benefit for the plant, a behavior known as nectar robbing (Irwin et al., 2010). Sclareol has been shown to form epicuticular crystals at the surface of the calyx (Caissard et al., 2012). These crystals may protect the flower against florivory and nectar robbing, and could also reduce fungal spores adhesion (Legrand et al., 2010; Caissard et al., 2012).

The biological function of sclareol has been investigated in other plant species (Seo et al., 2012; Fujimoto et al., 2015). Notably, sclareol has been reported to display antimicrobial activity against different species of plant pathogenic bacteria and fungi (Seo et al., 2012; Fujimoto et al., 2015; Chain et al., 2015). Exogenous application of sclareol on tobacco, tomato and *Arabidopsis thaliana* plants was shown to induce resistance to bacterial wilt disease (Seo et al., 2012). In tomato and *Arabidopsis thaliana*, sclareol could also inhibit the penetration of the root-knot nematode in the roots, thereby conferring resistance to this pathogen (Fujimoto et al., 2015). Taken together, these results suggest that sclareol may be involved in the defense of clary sage flowers against pathogenic microorganisms.

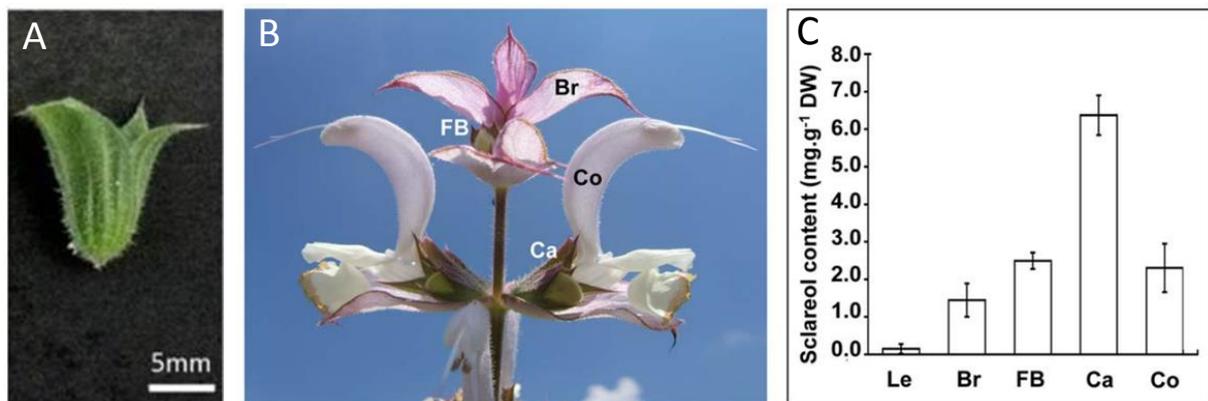


Figure 3: Sclareol is mainly produced in clary sage calyces.

A: Clary sage calyx. **B:** Apex of a clary sage inflorescence (Caissard et al., 2012). **C:** Sclareol quantification in different flower organs of clary sage (Caissard et al., 2012). Legends: **Br**, bract; **FB**, flower bud; **Co**, corolla; **Ca**, calyx; **Le**, leaf.

2.3. Clary sage commercial use

Clary sage has been traditionally used as a medicinal plant for its antispasmodic, carminative and estrogenic properties (Lattoo et al., 2006). Notably, essential oil extracted from clary sage has been reported to have numerous biological effects, including anti-inflammatory, antioxidant, antimicrobial and cytotoxic activity (Kuźma et al., 2009b; Sharopov and Setzer, 2012; Kumar et al., 2013). The individual contribution of each of its components to essential oil medicinal properties is not known and synergy, antagonism or additive effects may be at play (Kuźma et al., 2009b). In particular, sclareol has been shown to display anti-bacterial and anti-proliferative activity (Mendoza et al., 2002; Sashidhara et al., 2007).

Today, clary sage is exploited at industrial scale mainly for its aromatic properties. It is commercially cultivated in different European countries, particularly in France, Hungary and Bulgaria, but also in North America and China (Lawrence, 1994; Lattoo et al., 2006; Caniard et al., 2012). Clary sage essential oil is used as a flavor for the elaboration of diverse processed food products (Lattoo et al., 2006; Wagner et al., 2012; Sharopov and Setzer, 2012). Most importantly, essential oil and sclareol extracted from clary sage are employed in perfume manufacture. Essential oil is used directly as a perfume component for its tenacious, herbaceous, sweaty and amber odor (Laville et al., 2012; Kumar et al., 2013), while sclareol is a common starting material for the hemisynthesis of the perfume component ambroxide (Yang et al., 2016). When grown for commercial use in flavor and fragrance industry, clary sage inflorescences are dried and submitted to steam distillation in order to collect essential oil. Sclareol is then obtained from remaining plant material by solid/liquid extraction with an organic solvent (Caissard et al., 2012; Laville et al., 2012; Leffingwell and Leffingwell, 2015) (**Figure 4**). In general, essential oil yield is between 0.1% and 0.3% (Wagner et al., 2012) and sclareol yield is between 0.5% and 1.5% (Caissard et al., 2012; ITEIPMAI, 7èmes Rencontres du Végétal, 2013).

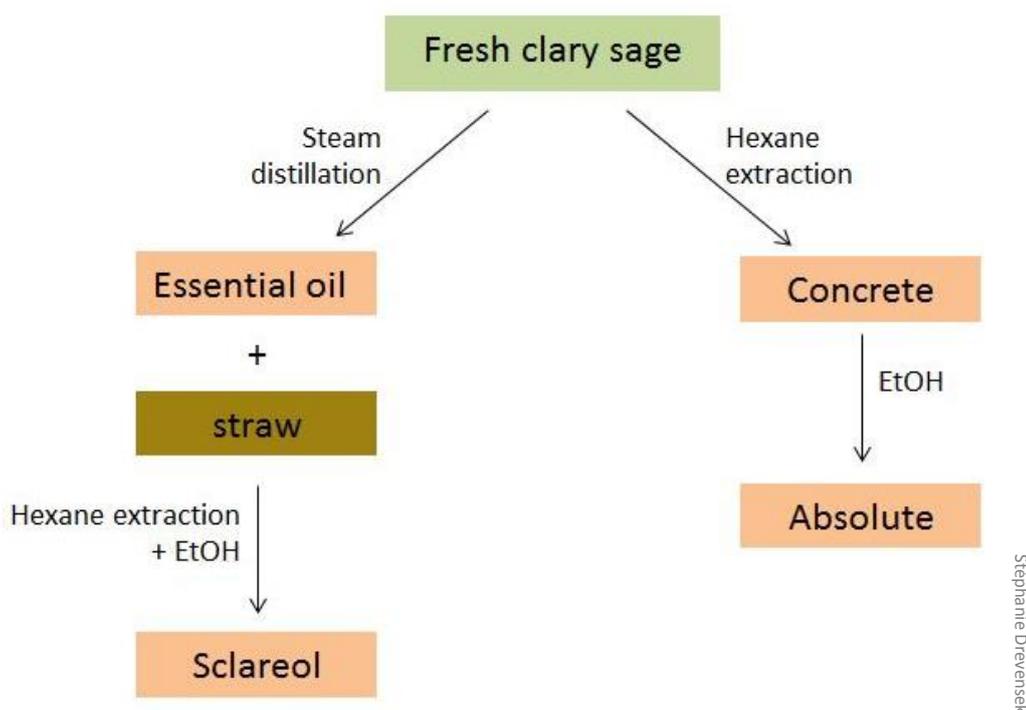


Figure 4: Exploitation of clary sage in perfume industry.

2.4. First advances in clary sage genetics and biotechnology

The identification of genes involved in specific biological processes is greatly facilitated when genomic data are available in the species of interest or in related species. The genome of clary sage is not sequenced yet, but genome sequences of *Salvia miltiorrhiza* and *Salvia splendens* have been recently published (Xu et al., 2016; Dong et al., 2018). The study of clary sage chromosomic material shows that it is a diploid species with a karyotype of $2n=22$ (Ozdemir and Senel, 1999). Genome size was estimated around 600 Mb (Drevensek et al., unpublished data). Two sets of clary sage transcriptomic data are currently available. The first clary sage transcriptome was obtained from calyces and allowed the identification of putative candidate genes for different steps of terpenoid biosynthesis (Legrand et al., 2010). Another transcriptome has been generated from leaves and roots treated with the phytohormone methyljasmonate (Hao et al., 2015).

Biotechnological tools enabling the modification of gene expression level are highly valuable for the functional characterization of genes of interest. In many plants, the development of roots named “hairy roots” can be induced using the bacterium *Agrobacterium rhizogenes*, which is able to transfer human-engineered genetic material into plant genome. Hairy roots can then be separated from aerial parts and cultivated in a bioreactor (Ron et al., 2014). Several studies show that transgenic hairy roots can be generated from clary sage roots (Kuźma et al., 2009a; Vaccaro et al., 2014, 2017; Alfieri et al., 2018) (**Figure 5**). Stable lines of clary sage calli and cell cultures can also be obtained from seeds or stem explants, and were shown to produce sclareol (Banthorpe et al., 1990). Moreover, protocols enabling *in vitro* regeneration of whole clary sage plants from fragments of somatic tissue are available, starting from immature zygotic embryo cotyledons (Liu et al., 2000) or shoot tips (Kuźma et al., 2009b). Essential oil composition of clary sage plants regenerated from shoot tips was comparable to plants reproduced from seeds, but the yield was twice lower (Kuźma et al., 2009b). Therefore, the metabolic activity of regenerated clary sage plants may be different compared to plants reproduced from seeds, but further investigation is needed to confirm this hypothesis.



Figure 5: Clary sage hairy root culture (Alfieri et al., 2018).

2.5. Clary sage breeding in France

The growing demand for sclareol encouraged the development of clary sage breeding programs in countries where clary sage is cultivated, including Hungary, Bulgaria, Moldavia, Romania and France (source: Community Plant Variety Office). In France, the Interprofessional Technical Institute of Perfume, Medicinal, Aromatic and Industrial Plants (ITEIPMAI) has obtained several improved clary sage varieties through recurrent selection targeting sclareol yield and flowering time. Recurrent selection is a plant breeding approach involving successive cycles of three steps: intermating between best-performing populations, evaluation of traits of interest, and selection of best-performing individuals. Through this method, a clary sage variety with high sclareol yield was obtained: “Scalia”. Unfortunately, this variety displayed a late-flowering phenotype. This phenotype can be detrimental to sclareol yield because late-flowering plants are potentially more impacted by summer drought (Otto et al., 2017). Further recurrent selection cycles led to the development of the “Toscalia” variety, which is characterized by a high sclareol yield without any late-flowering phenotype. A third variety, “Claryssima”, is currently being developed using the same approach (source: ITEIPMAI).

3. Terpene biosynthesis and function in plants

3.1. Specialized metabolism and terpenoids

Throughout their lifetime, plants produce a wide variety of chemical compounds which perform different biological functions. Plant compounds have been historically divided into two classes according to their biological role. Metabolites which are essential to sustain normal growth are called primary metabolites. For example, this category comprises structural compounds like cell wall carbohydrates or membrane lipids, and compounds involved in energy metabolism or photosynthesis like organic acids or coenzymes. In addition to primary metabolites, plants produce various other compounds which are not indispensable for plant growth and development, but are involved in plant adaptation to biotic and abiotic environmental factors (Figueiredo et al., 2008; Tholl, 2015; Chezem and Clay, 2016). These compounds are therefore critical for plant survival in a fluctuating environment and are known as secondary or specialized metabolites. Unlike primary metabolites, which are generally present in all plant lineages, specialized metabolites are highly diverse between plant species. In a single plant species, the amount of specialized metabolites produced can also vary drastically according to environmental conditions.

Plant metabolites can also be classified according to early steps of their biosynthesis pathway. Terpenoids represent the largest class of plant metabolites with at least 50,000 compounds described (Liao et al., 2016). Some terpenoids play critical roles in plant growth and development. They are therefore considered as primary metabolites, like for example phytohormones of the gibberellin family and photosynthetic pigments of the carotenoid family. However, most plant terpenoids are specialized metabolites involved in diverse biological functions related to plant interaction with its environment (Lipko and Swiezewska, 2016).

The biosynthesis of all terpenoids starts with the condensation of the same five-carbon building blocks: isopentenyl pyrophosphate (IPP) and its isomer dimethylallyl pyrophosphate (DMAPP) (Tholl, 2015). The number of five-carbon units condensed defines subclasses among terpenoids: for instance, compounds derived from the condensation of 2, 3 or 4 five-carbon units belong to monoterpene (10C), sesquiterpene (15C) and diterpene (20C) subclasses, respectively. Sclareol belongs to the diterpene subclass.

3.2. The diverse biological functions of terpenes

Numerous terpenes are indispensable for plant cell function and correct plant growth and development. Primary metabolites belonging to the terpenoid family include carotenoids, which are essential for light harvesting during photosynthesis, and phytosterols, which are structural elements of plant plasma membranes (**Figure 6**). They also include important signaling metabolites, as several classes of important plant hormones are terpenoids: gibberellins, abscisic acid, strigolactones and brassinosteroids. Moreover, some protein post-translational modifications are related to terpenoid metabolism: protein prenylation consists in the transfer of a prenyl moiety to a cysteine, and the terpenoid dolichol is involved in protein glycosylation (Tholl, 2015).

Some terpenoids are involved in plant response to abiotic stress, although this is not their most common function. For example, in poplar, isoprene emissions provide protection against thermal stress and ozone-induced oxidative stress (Behnke et al., 2007, 2009). By contrast, a large number of plant terpenoids are key elements of plant constitutive and induced chemical defense systems against pathogenic microorganisms (fungi, bacteria) and insect herbivores (**Figure 6**). They can act as direct defense by displaying toxicity to plant pathogens or by acting as herbivore repellents (Moore et al., 2014). For example, numerous rice diterpenoid phytoalexins such as oryzalexins show antibiotic activity against the fungal pathogen *Magnaporthe oryzae* (Schmelz et al., 2014), and 17-hydroxygeranylinalool diterpenoid glycosides produced by *Nicotiana attenuata* work as antifeedants against tobacco hornworm (Jassbi et al., 2008). Terpenoids can also act as indirect defense via the attraction of natural enemies of plant pathogens, i.e. predators or parasites (Moore et al., 2014). Such indirect action is illustrated by the response of maize roots to feeding by *Diabrotica virgifera* larvae. Damaged maize roots produce β -caryophyllene, which in turn strongly attracts an entomopathogenic nematode of *Diabrotica virgifera* larvae (Rasmann et al., 2005).

In addition to defense against pathogenic microorganisms and herbivores, terpenoids play other roles in plant-insect interactions. Notably, volatile scent terpenoids emitted by flowers are critical for pollinator attraction. The monoterpene 1,8-cineole is frequently present in the scent of orchids and strongly attract euglossine bees, a group of bees specifically pollinating orchids (Dodson et al., 1969). Plant-plant interactions can also be mediated by

terpenes. Some terpenoids allow parasitic plants to detect their host, induce defense responses in neighboring plants, or display toxicity to plant competitors (Tholl, 2015). For example, momilactone diterpenes produced by rice roots have a negative effect on the growth of barnyard grass, a plant competitor of rice (Xu et al., 2012).

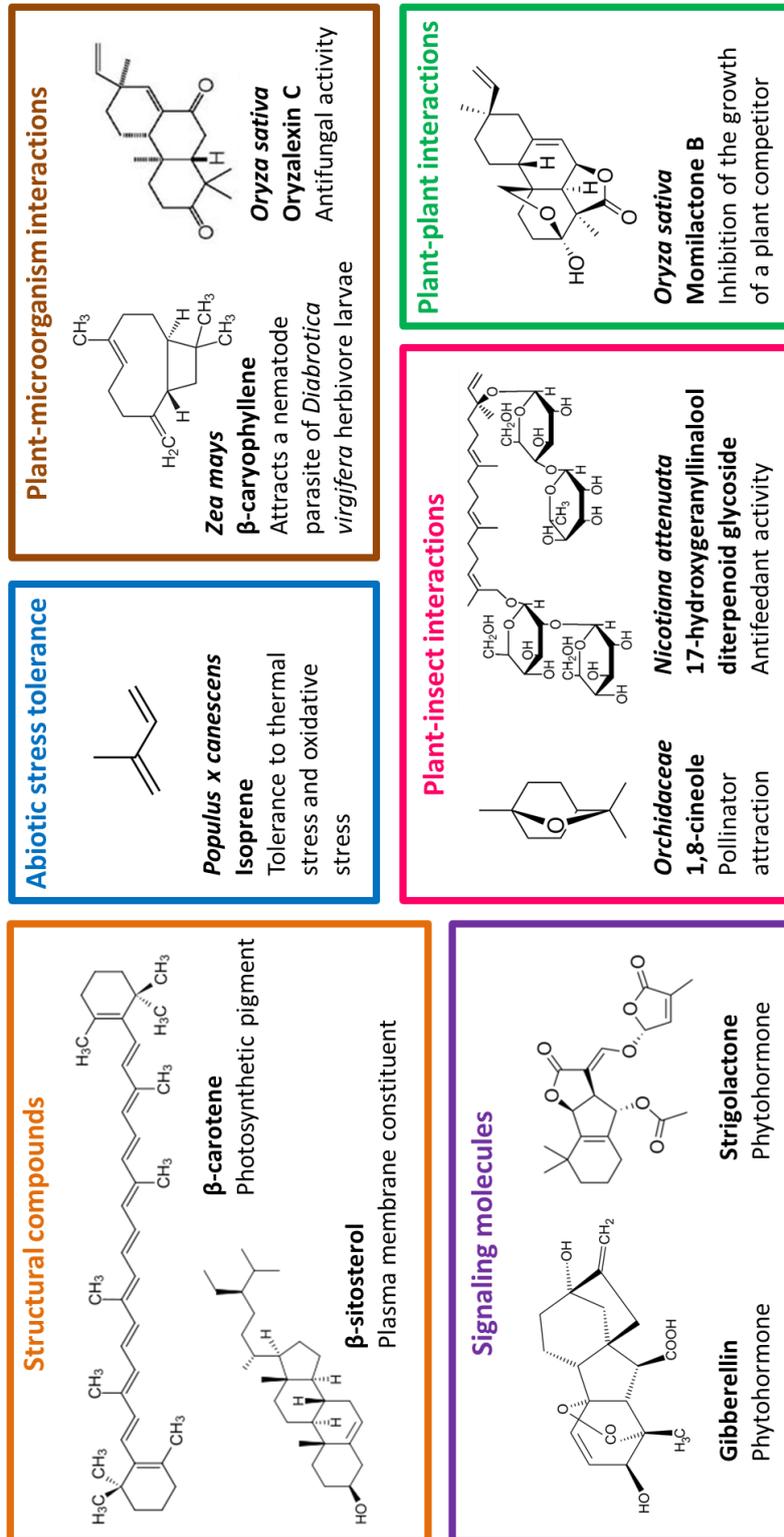


Figure 6: Examples illustrating the diverse biological functions of terpenes in plants.

Camille Chalvin

3.3. Industrial exploitation of terpenes

Because of their various ecological functions, terpenes display distinct properties and biological activities which are valuable for industrial applications. In particular, a number of plant-derived terpenoids have been used for a long time as scents or flavors, and as active pharmacological ingredients (Tetali, 2018). Essential oils, which are mainly composed of terpenoids, are important natural ingredients of flavor and perfume industry, and their components are also individually employed for specific uses. For example, the sesquiterpene nootkatone is the main molecule responsible for the distinct aroma of grapefruit, and the monoterpenes limonene, linalool and 1,8-cineole are common ingredients of lemon or lime flavor (Hausch et al., 2015; Tetali, 2018). Pharmacologically active molecules are found in all subclasses of terpenes, including monoterpenes and diterpenes (**Figure 7**). The most famous plant diterpene possessing medicinal properties may be the cancer drug paclitaxel (Taxol®), initially isolated from yew (Mafu and Zerbe, 2018). The diterpene ingenol mebutate from *Euphorbia peplus* also displays antitumor effects and has been recently approved in clinical trials for skin cancer treatment (Mafu and Zerbe, 2018). In addition to cancer therapy, terpenes are used in many other medicinal applications; for example, the diterpene forskolin is a cardioprotective compound found in the roots of *Coleus forskohlii* (Mafu and Zerbe, 2018), and the sesquiterpene artemisinin produced by *Artemisia annua* is a reference for the treatment of malaria (Tetali, 2018).

In addition to scents and medicines, other plant-derived terpenoids are used in food, cosmetic and chemical industries (Tetali, 2018). Some terpenoids have important nutritional value, as illustrated by the role of carotenoids which represent an essential source of vitamin A for humans (Rodriguez-Concepcion et al., 2018). Carotenoids are also employed as antioxidants in cosmetic industry, notably in the elaboration of skin care products (Sathasivam and Ki, 2018). Terpenoid properties are useful in other sectors of chemical industry, like for instance polymer chemistry in the case of polyisoprene constituents of natural rubber produced by *Hevea brasiliensis*. Moreover, due to their low hygroscopy, high energy density, and good fluidity at low temperatures, biotechnologically-produced sesquiterpenes and diterpenes have been suggested as potential renewable alternatives for transportation fuels (Tippmann et al., 2013).

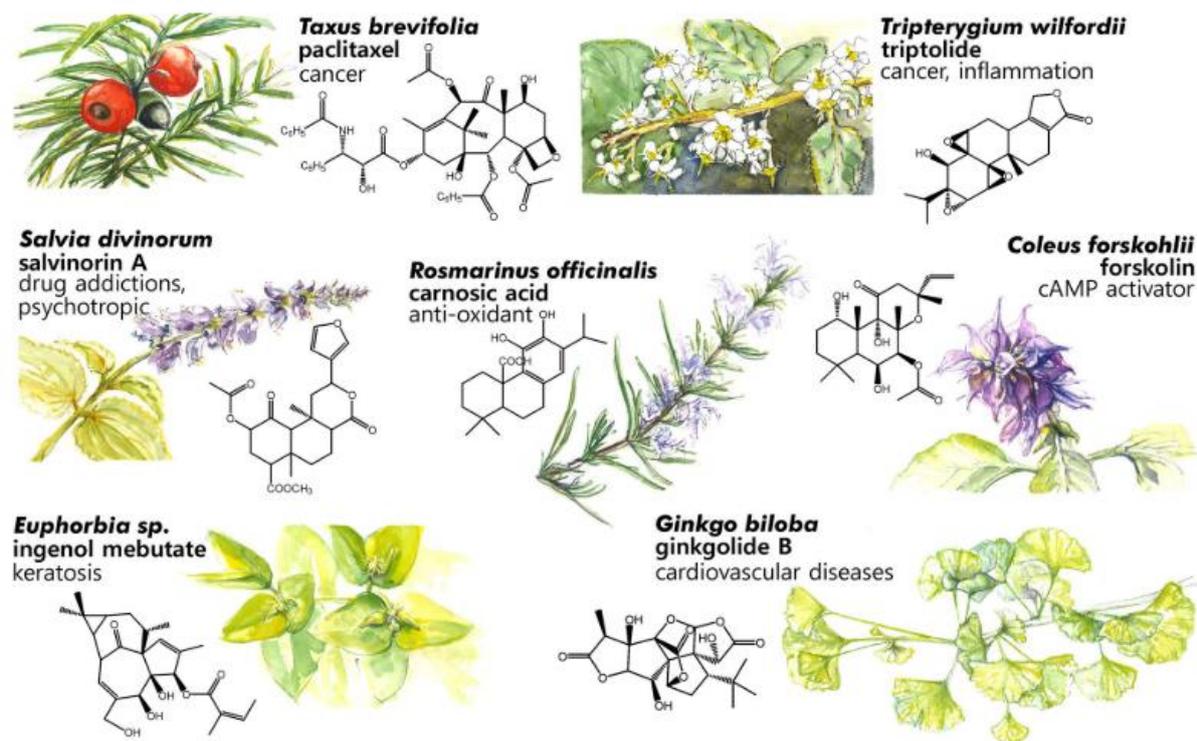


Figure 7: Pharmacologically active diterpenoids and their respective source plants (Mafu and Zerbe, 2018).

3.4. Terpene biosynthesis

The first step of terpenoid biosynthesis corresponds to the production of the five-carbon building blocks IPP and DMAPP. In plants, IPP and DMAPP are produced by two distinct metabolic pathways: the mevalonate pathway (MVA pathway) and the methylerythritol-phosphate pathway (MEP pathway). Then, linear intermediate compounds called prenyl intermediates are formed by condensation of IPP and DMAPP units (Tholl, 2015). Prenyl intermediates are finally converted into functional terpenes through a number of molecular modifications, for example cyclization and addition of functional groups like hydroxyl groups.

3.4.1. Production of five-carbon building blocks by MEP and MVA pathways

IPP and DMAPP are the universal five-carbon building blocks from which all terpenes are derived. The co-existence of two metabolic pathways for the production of IPP and DMAPP is an atypical feature of plants: most bacteria use only the MEP pathway, and animals, fungi and archae only have the MVA pathway (Hemmerlin et al., 2012; Lipko and Swiezewska,

2016). In plants, MVA and MEP pathways are compartmentalized: the MVA pathway is localized in the cytosol, whereas the MEP pathway is localized in plastids. It is now admitted that plants inherited the MEP pathway from a photosynthetic prokaryote through endosymbiosis (Vranová et al., 2013).

In plants, the MVA pathway involves 6 enzymatic steps (Tholl, 2015) (**Figure 8**). It starts with the condensation of 2 molecules of acetyl-coenzyme A (acetyl-CoA) by the acetoacetyl-CoA thiolase (AACT), which leads to the formation of acetoacetyl-CoA. The HMG-CoA synthase (HMGS) then catalyzes the condensation of acetoacetyl-CoA with another molecule of acetyl-CoA to generate hydroxymethylglutaryl-CoA (HMG-CoA). The third step of the pathway is a rate-limiting step where HMG-CoA is reduced to mevalonate (MVA) by the key enzyme HMG-CoA reductase (HMGR). MVA is then phosphorylated twice through the consecutive action of the mevalonate kinase (MK) and the phosphomevalonate kinase (PMK). The last step of the MVA pathway corresponds to the conversion of mevalonate-diphosphate into IPP, catalyzed by the mevalonate-diphosphate decarboxylase (MPDC).

The MEP pathway found in plants consists of 7 enzymatic steps (Tholl, 2015) (**Figure 8**). The first step of the pathway is a rate-limiting step where the DXP-synthase (DXS) catalyzes the formation of 1-deoxy-D-xylulose-5-phosphate (DXP) from glyceraldehyde-3-phosphate and pyruvate. Plant DXS enzymes are divided in 2 classes which correspond to 2 distinct phylogenetic clades: DXS1 and DXS2 enzymes. DXS1 enzymes are expressed in photosynthetic and floral tissues and are thought to be mostly involved in primary metabolism, whereas DXS2 enzymes have more restricted spatiotemporal expression patterns and are generally involved in specialized metabolism (Tholl, 2015). The second step of the MEP pathway corresponds to the conversion of DXP into 2-methyl-D-erythritol-4-phosphate (MEP) through the action of the DXP-reductoisomerase (DXR). The MEP-cytidyltransferase (MCT) then couples MEP to CDP to form 4-diphosphocytidyl-2-methyl-D-erythritol (CDP-ME), which is subsequently phosphorylated by the CDP-ME kinase (CMK). The obtained 4-diphosphocytidyl-2-methyl-D-erythritol-2-phosphate (CDP-ME₂P) is cyclized by the MEcPP synthase (MDS) to form 2-methyl-D-erythritol-2,4-cyclodiphosphate (MEcPP). MEcPP is reduced into 4-hydroxy-3-methylbut-2-enyl-diphosphate (HMBPP) by the HMBPP-synthase (HDS) and HMBPP is finally converted into a mixture of IPP and DMAPP by the HMBPP-reductase (HDR) with a ratio of 5 to 6:1.

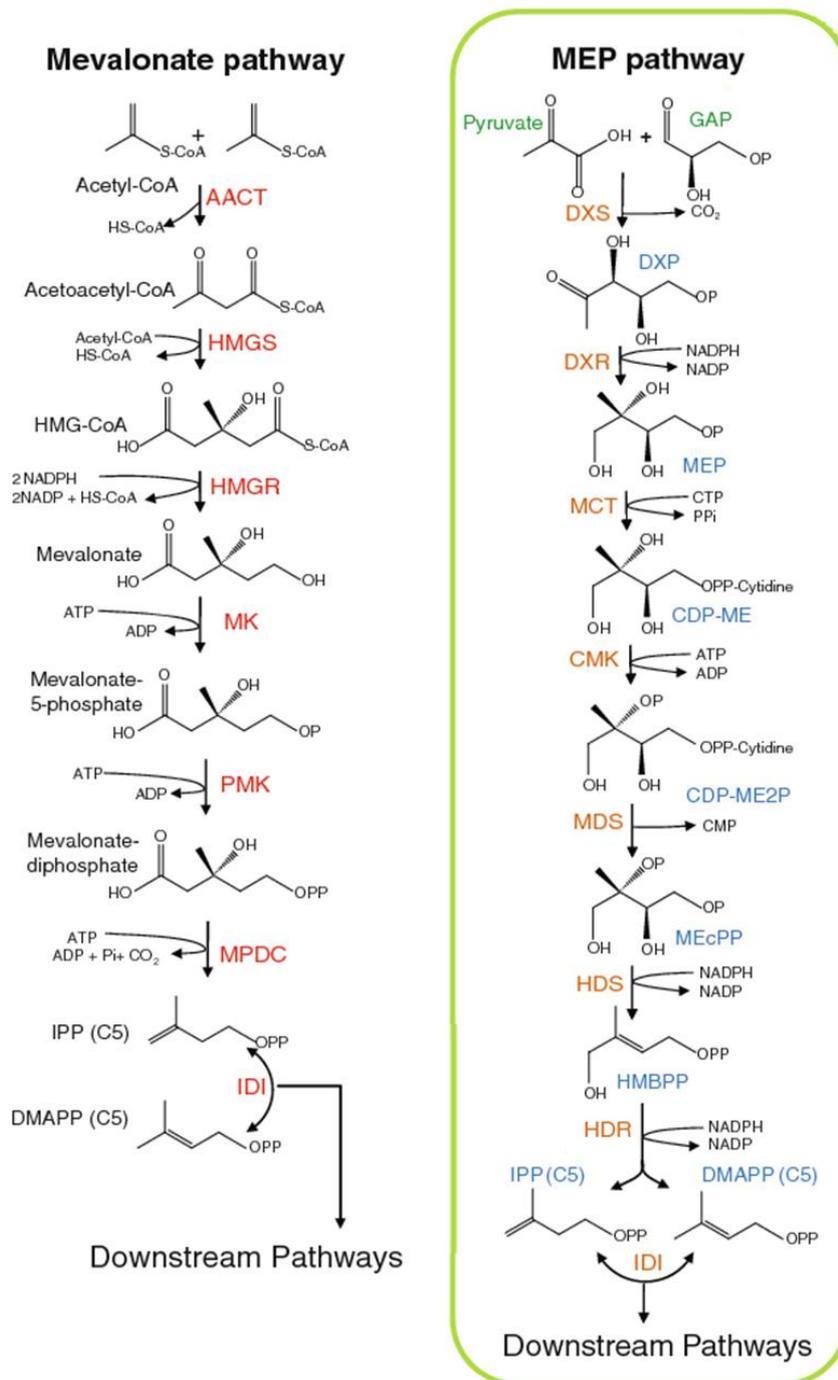


Figure 8: Production of five-carbon building blocks by MEP and MVA pathways (adapted from Tholl, 2015).

AACT, acetoacetyl-CoA thiolase; **CDP-ME**, 4-diphosphocytidyl-2-methyl-D-erythritol; **CDP-ME2P**, 4-diphosphocytidyl-2-methyl-D-erythritol-2-phosphate; **CMK**, CDP-ME kinase; **CoA**, coenzyme A; **DMAPP**, dimethylallyl pyrophosphate; **DXP**, 1-deoxy-D-xylulose-5-phosphate; **DXR**, DXP-reductoisomerase; **DXS**, DXP-synthase; **GAP**, glyceraldehyde-3-phosphate; **HDR**, HMBPP-reductase; **HDS**, HMBPP-synthase; **HMBPP**, 4-hydroxy-3-methylbut-2-enyl-diphosphate; **HMG-CoA**, hydroxymethylglutaryl-CoA; **HMGR**, HMG-CoA reductase; **HMGS**, HMG-CoA synthase; **IDI**, IPP isomerase; **IPP**, isopentenyl pyrophosphate; **MCT**, MEP-cytidyltransferase; **MDS**, MEcPP synthase; **MEcPP**, 2-methyl-D-erythritol-2,4-cyclodiphosphate; **MEP**, 2-methyl-D-erythritol-4-phosphate; **MK**, mevalonate kinase; **MPDC**, mevalonate-diphosphate decarboxylase; **MVA**, mevalonate; **PMK**, phospho-mevalonate kinase.

3.4.2. Condensation of five-carbon building blocks by prenyltransferases

The second step of terpenoid biosynthesis corresponds to the condensation of IPP and DMAPP units to form prenyl diphosphate intermediates, which are the precursors of all terpenes. These condensation reactions are catalyzed by enzymes called prenyl diphosphate synthases or prenyltransferases. First, an IPP molecule is condensed with a DMAPP molecule to produce a ten-carbon intermediate, and further addition of IPP units leads to the formation of prenyl diphosphate intermediates of various lengths (Kharel and Koyama, 2003). Prenyltransferase activity requires an appropriate ratio of IPP and its more reactive isomer DMAPP. Conversion of IPP into DMAPP is performed by the IPP isomerase (IDI). IPP isomerization is more important in the cytosol than in plastids because the MVA pathway produces only IPP, whereas the MEP pathway produces both five-carbon units (Tholl, 2015). Prenyltransferases are divided in two classes according to the stereochemistry of double bonds in the product of the condensation reaction: *cis*-prenyltransferases and *trans*-prenyltransferases produce all-*cis* and all-*trans* prenyl diphosphate intermediates, respectively (Kharel and Koyama, 2003). Condensation of 2, 3 or 4 five-carbon units in *trans* orientation leads to the formation of geranyl pyrophosphate (GPP, 10 carbons), (*E,E*)-farnesyl pyrophosphate (FPP, 15 carbons) and geranylgeranyl pyrophosphate (GGPP, 20 carbons), respectively (**Figure 9**). These *trans*-prenyl diphosphate intermediates are the main precursors of monoterpene, sesquiterpene and diterpene biosynthesis, respectively. They are more abundant than their all-*cis* isomers and constitute central metabolic branching points for the biosynthesis of terpenoids involved in primary and specialized metabolism (Tholl, 2015). For these reasons, most research studies conducted so far have been focused on *trans*-prenyltransferases.

GGPP, FPP and GPP are produced by geranylgeranyl pyrophosphate synthases (GGPP-synthases), farnesyl pyrophosphate synthases (FPP-synthases) and geranyl pyrophosphate synthases (GPP-synthases), respectively (**Figure 9**). Most GGPP and GPP molecules are produced in plastids, whereas FPP synthesis essentially occurs in the cytosol (Lu et al., 2016). GGPP-synthase and FPP-synthase genes are generally present in multiple copies in plant genomes, GGPP-synthase gene families being larger than FPP-synthase gene families (Tholl, 2015). In angiosperms, GGPP-synthases and FPP-synthases usually work as homodimers to produce respectively GGPP and FPP as major or unique product (Tholl, 2015). By contrast,

several GPP-synthases have been shown to work as heterodimers including a large catalytic subunit and a small regulatory subunit, like those described in *Antirrhinum majus*, *Clarkia breweri* and *Humulus lupulus* (Tholl et al., 2004; Wang and Dixon, 2009). The large subunit is closely related to GGPP-synthases and displays GGPP-synthase activity *in vitro*, whereas the small subunit shows only limited sequence similarity with GGPP-synthases (around 20%) and does not show any catalytic activity. It is admitted that *in vivo*, the small subunit binds to the large subunit and modifies its activity to make it produce GPP instead of GGPP (Burke and Croteau, 2002; Orlova et al., 2009). Homodimeric and heterotetrameric GPP-synthases have also been characterized (Burke et al., 1999; Burke and Croteau, 2002; Hsiao et al., 2008). Notably, the first GPP-synthase discovered in plants was identified in the Lamiaceae *Mentha piperita* and works as a heterotetramer with two large catalytic subunits and two small regulatory subunits similar to subunits forming heterodimeric GPP-synthases (Burke et al., 1999; Burke and Croteau, 2002). Some plant species have both heteromeric and homomeric GPP-synthases, for example *Catharanthus roseus* (Rai et al., 2013).

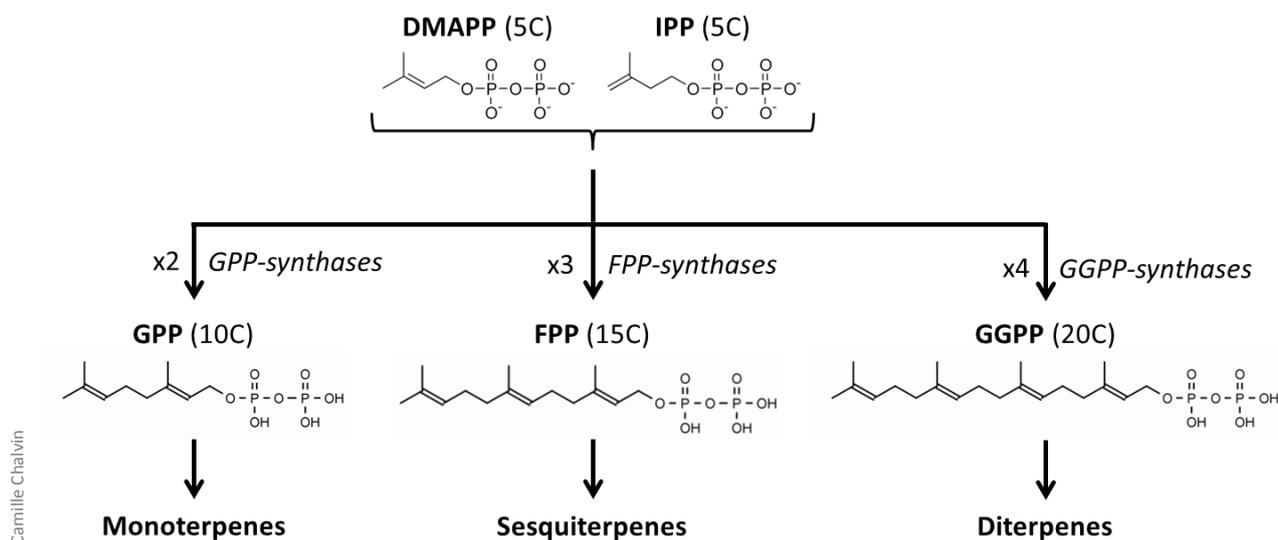


Figure 9: The role of prenyltransferases in terpene biosynthesis.

DMAPP, dimethylallyl pyrophosphate; **FPP**, farnesyl pyrophosphate; **GPP**, geranyl pyrophosphate; **GGPP**, geranylgeranyl pyrophosphate; **IPP**, isopentenyl pyrophosphate.

3.4.3. Terpene synthases

In the last step of terpenoid biosynthesis, prenyl diphosphate intermediates undergo various molecular modifications which eventually lead to functional terpenes. Possible modifications include cyclization or cleavage of the hydrocarbon skeleton, and functionalization reactions such as hydroxylation, peroxidation, methylation, acylation or glycosylation (Tholl, 2015). This step involves enzymes of the terpene synthase (TPS) superfamily and often requires other types of enzymes, like for example cytochrome P450 monooxygenases (Zerbe and Bohlmann, 2015; Bathe and Tissier, 2019).

Plant genomes generally contain numerous putatively functional *TPS* genes. Only 2 were found in the genome of the moss *Physcomitrella patens*, but genomes of vascular plants generally have more than 20 *TPS* genes, up to 113 in *Eucalyptus grandis* (Chen et al., 2011). The large number of *TPS* genes found in plant genomes, the promiscuous activity of these enzymes and their ability to acquire new catalytic properties by minor structural changes are believed to be responsible for the huge diversity of plant terpenoids (Tholl, 2015). According to phylogenetic studies, *TPS* genes are divided into 8 subfamilies: *TPSa* to *TPSh*. *TPS* enzymes can also be classified in 2 classes according to the mechanism of the enzymatic reaction they catalyze (Chen et al., 2011). This enzymatic reaction always involves the formation of a carbocation within the prenyl diphosphate substrate, which opens up the possibility of intramolecular rearrangements or attacks by nucleophilic compounds such as water. Class I *TPS* enzymes ionize their substrate by removing the diphosphate group, whereas class II *TPS* do it through a protonation reaction (Chen et al., 2011). *In vitro*, *TPS* enzymes are sometimes able to use prenyl diphosphate substrates of different lengths, but *in vivo* their activity is determined by the substrate pool present in the compartment where they are localized. Given the respective proportions of GPP, FPP and GGPP molecules in plastids and in the cytosol, plastidial *TPS* enzymes mainly produce monoterpenes and diterpenes, whereas cytosolic *TPS* enzymes mainly produce sesquiterpenes (Tholl, 2015). However, exceptions to this rule have been reported; for instance, plastidial sesquiterpene synthases have been characterized in tomato (Sallaud et al., 2009).

3.4.4. Focus on diterpene biosynthesis

Diterpene biosynthesis usually occurs in chloroplasts, where most GGPP precursors are produced (Tholl, 2015). The various diterpene biosynthesis pathways are believed to have evolved from gibberellin biosynthesis (Peters, 2010). A majority of diterpenes, including sclareol, are labdane-type diterpenes. Formation of diterpenes with a hydrocarbon skeleton of the labdane type usually involves the cyclization of GGPP through the consecutive action of a class II diterpene synthase (**Figure 10**) and a class I diterpene synthase (Peters, 2010). The obtained bicyclic skeleton can be further modified, for example by the addition of hydroxyl groups mediated by diterpene synthases or cytochrome P450 monooxygenases (Zerbe and Bohlmann, 2015; Bathe and Tissier, 2019). Another cyclization reaction leads to tricyclic abietane diterpenes (Peters, 2010) (**Figure 10**).

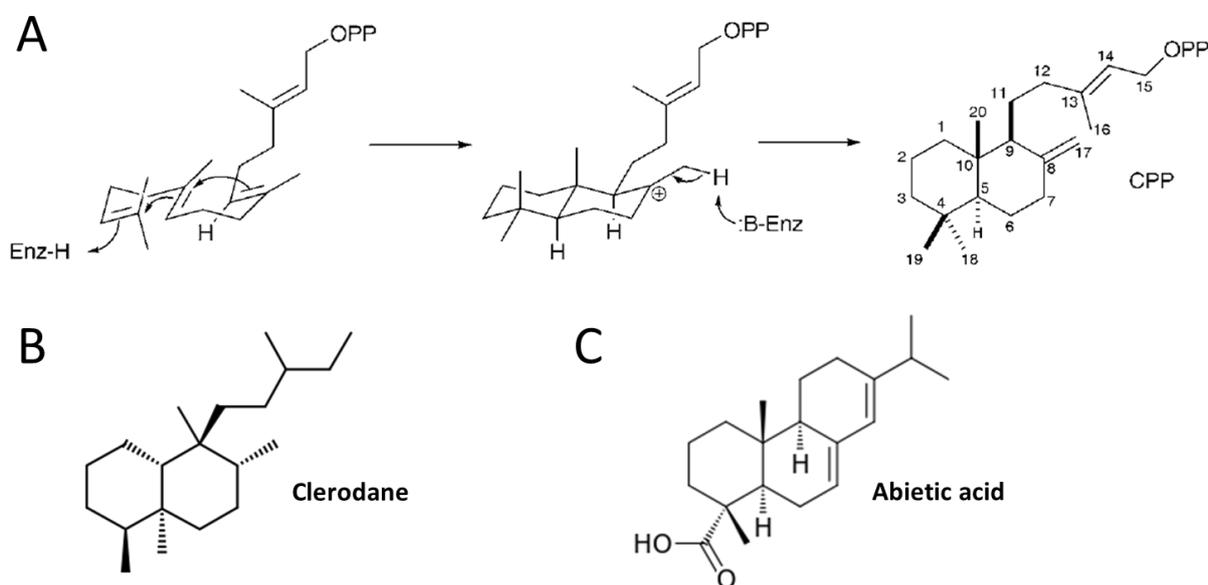


Figure 10: Examples of labdane-related diterpenes.

A: Example of enzymatic reaction generating a labdane-type hydrocarbon skeleton (Peters, 2010). This protonation-initiated cyclization of GGPP is catalyzed by a class II diterpene synthase (**Enz**). Copalyl diphosphate (**CPP**) is an intermediate in the biosynthesis of many diterpenes (Peters, 2010).

B: Chemical formula of a bicyclic labdane-related diterpene, clerodane.

C: Chemical formula of a tricyclic labdane-related diterpene, abietic acid.

3.5. Terpene biosynthesis regulation

3.5.1. Factors affecting terpene content

Specialized metabolite content varies at all scales, between species and populations, but also between individuals and between the different parts of the same plant (Moore et al., 2014). In a given species, populations displaying distinct metabolite profiles are referred to as chemotypes (Moore et al., 2014). Physiological variations, environmental conditions and genetic polymorphisms, along with the interactions of these three factors, are known to determine specialized metabolite content in plants (Figueiredo et al., 2008; Moore et al., 2014).

In a single plant, terpenoid production depends on the organ considered, on the developmental stage of this organ, and also on circadian and seasonal cycles (Moore et al., 2014). Important qualitative and quantitative variations are observed between the different organs of the plant. For example, the flowers of *Lavandula pinnata* produce more monoterpenes and less sesquiterpenes than leaves and stems. These differences are related to the important role played by volatile monoterpenoids in guiding pollinators to the flowers (Figueiredo et al., 2008). In *Achillea millefolium*, the amount of volatile terpenoids produced by flowers differs according to their developmental stage. The amount of sesquiterpene chamazulene present in the essential oil decreases with flower maturation, while the proportion of the monoterpenes 1,8-cineole and camphor increases (Figueiredo et al., 2008). Diurnal variations of light level impact MEP and MVA pathways in the opposite way: the MEP pathway is stimulated by exposure to light, whereas the MVA pathway is upregulated in the dark (Tholl, 2015). Seasonal variations in terpenoid content can also be observed in some species. In sea fennel (*Crithmum maritimum*), proportions of sabinene and γ -terpinene in the essential oil vary in opposite ways: sabinene is more abundant during the flowering season (between July and October), whereas γ -terpinene predominates during the rest of the year and peaks in April (Barroso et al., 1992) (**Figure 11**).

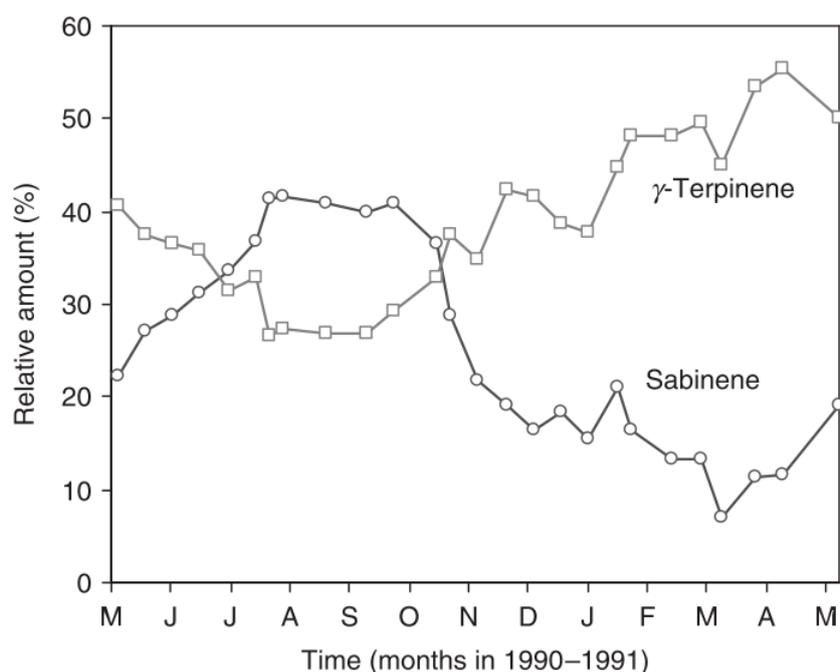


Figure 11: Seasonal variations of terpene production in *Crithmum maritimum* (Barroso et al., 1992).

Environmental conditions impacting specialized metabolite content include climatic factors (temperature, hygrometry), biotic factors (herbivores, microorganisms) and edaphic factors. Terpenoid production is often tightly related to weather conditions. Monoterpene emission by slash pine (*Pinus elliotii*) has been shown to increase with temperature (Tingey et al., 1980). Drought has contrasted effects on terpenoid production depending on the species considered: for example, hydric stress enhances terpenoid emissions in *Ocimum basilicum*, but decreases artemisinin production in *Artemisia annua* (Figueiredo et al., 2008). Since many terpenoids play a role in plant defense against pathogens, biotic factors significantly influence terpene levels in various plants. For instance, the sesquiterpene capsidiol accumulates in pepper stems upon infection with the oomycete *Phytophthora capsici* (Egea et al., 1996). Finally, soil physical and chemical structure is also considered as an important factor impacting specialized metabolite content in plants. Notably, the proportion of silt and sand as well as potassium oxide content influence the composition of *Salvia desoleana* essential oil (Rapposelli et al., 2015).

Genetic polymorphisms are responsible for variations in terpenoid content between plants belonging to the same species or to related species. In species of the genus *Mentha*, the dominant allele *C* has an important impact on essential oil composition: *CC* or *Cc* genotypes are associated with high carvone content, whereas the *cc* genotype promotes the

production of pulegone and menthol (Figueiredo et al., 2008). Genetic variations are shaped by the selective pressure exerted by environmental factors mentioned above. The action of natural selection is illustrated by the study of monoterpene production by *Pinus ponderosa* trees in the context of the development of bark beetles, which are the most destructive agent of conifer forest worldwide. Directional selection has been shown to take place for trees producing high amounts of limonene, a monoterpene toxic to the beetle (Trapp and Croteau, 2001).

3.5.2. Molecular mechanisms of terpene biosynthesis regulation

Phytohormone signaling pathways and master transcriptional regulators of plant growth and development control terpenoid biosynthesis in response to environmental and developmental signals. In *Arabidopsis thaliana*, the upregulation of the MEP pathway and carotenoid biosynthesis in light conditions was shown to be performed through the action of phytochrome interacting factors (Toledo-Ortiz et al., 2010) and the biosynthesis of floral sesquiterpenes is induced by master regulators of flower maturation in response to jasmonates, auxin and gibberellins (Tholl, 2015). Jasmonates stimulate terpenoid biosynthesis in many other plants (Goossens et al., 2016); for example, they induce artemisinin biosynthesis in *Artemisia annua* (Chen et al., 2017b).

Genes involved in the same terpene biosynthesis pathway are often organized as gene clusters in plant genomes, an organization which facilitates their transcriptional coregulation (Töpfer et al., 2017). Positive transcriptional regulators of terpene biosynthesis genes have been identified in several plant species and generally belong to WRKY, AP2/ERF, bHLH or basic leucine zipper (bZIP) transcription factor families (Lu et al., 2016). For example, AaWRKY1, AaERF1, AaERF2, AaORA1 and AabZIP1 transcription factors positively regulate the biosynthesis of the sesquiterpene artemisinin in *Artemisia annua* (Yu et al., 2011; Lu et al., 2013; Han et al., 2014; Zhang et al., 2015a), TcWRKY1 stimulates the production of the diterpene paclitaxel in *Taxus chinensis* (Li et al., 2013) and the bZIP transcription factor OsTGAP1 positively regulates the biosynthesis of diterpene phytoalexins in rice (Okada et al., 2009).

In plants, MVA and MEP pathways are compartmentalized. Therefore, they can be differentially regulated, and controlled regulatory crosstalks between terpenoid biosynthesis

pathways have been reported (Tholl, 2015). Notably, terpenoid biosynthesis has been shown to be regulated at the post-translational level by protein prenylation (Liao et al., 2016). For instance, the production of the MVA-derived sesquiterpene capsidiol in tobacco relies on MEP-dependent protein geranylgeranylation (Huchelmann et al., 2014) and in *Catharanthus roseus*, MVA-derived farnesylated proteins regulate the expression of genes involved in MEP-dependent monoterpenoid biosynthesis (Courdavault et al., 2005).

4. Glandular trichomes

4.1. Capitate and peltate glandular trichomes of the Lamiaceae

In Angiosperms, terpenoid production is often localized in specialized secretory structures. For example, in *Hevea brasiliensis*, polyisoprene components of natural rubber are produced in specialized ducts known as laticifers (Lange, 2015). Lamiaceae are generally characterized by the presence of epidermal secretory structures called glandular trichomes, which produce various compounds mainly involved in pollinator attraction or defense against herbivores (Werker, 1993). Terpenoids secreted by glandular trichomes are major components of the essential oil which can be extracted by hydrodistillation. Glandular trichomes of the Lamiaceae are multicellular and composed of 3 parts: a base, a stalk and a gland (or glandular head) (**Figure 12**). The gland is responsible for the secretion of specialized metabolites, the stalk is the structure bearing the gland, and the base connects the stalk to surrounding epidermal cells (Lange, 2015). Both vegetative and reproductive organs of the Lamiaceae harbor glandular trichomes already at early development stages (Werker, 1993).

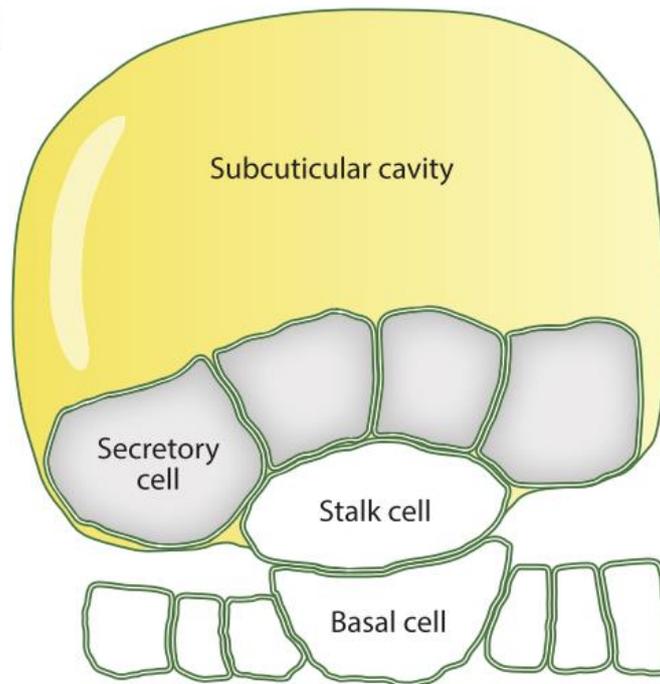


Figure 12: Structure of a peltate glandular trichome (Lange, 2015)

Glandular trichomes of the Lamiaceae show high morphological diversity (Werker, 1993; Gul et al., 2019). Lamiaceae generally display at least two types of glandular trichomes: capitate glandular trichomes and peltate glandular trichomes (Werker, 1993). This classification is based on their distinct morphology: capitate trichomes have a long stalk topped by a small spherical secretory head, while peltate trichomes are sessile trichomes with a very short stalk and a large flattened glandular head. In addition to these morphological differences, capitate and peltate trichomes also differ in the way they secrete chemicals. Both glandular trichome types accumulate secreted chemicals under an elevated cuticle. In peltate glandular trichomes, the cuticle is thick and ruptures only upon contact with an insect or as a consequence of another mechanical stimulation. By contrast, compounds secreted by capitate glandular trichomes are regularly released to the outside; they are thought to be either exuded through the cuticle, or released after spontaneous breakage of the cuticle (Werker, 1993; Figueiredo et al., 2008; Tissier, 2012).

4.2. Clary sage glandular trichomes

Like in most plants of the Lamiaceae family, both capitate and peltate glandular trichomes are found in clary sage, along with non-glandular trichomes (Werker et al., 1985; Schmiderer et al., 2008) (**Figure 13**). Several subtypes of capitate trichomes have been described according to their size: small capitate trichomes and large capitate trichomes are the most frequently observed, along with few intermediate trichomes. Small capitate heads measure between 10 and 20 μm , while large capitate heads measure between 50 and 100 μm (**Figure 13**). By contrast, the morphology of clary sage peltate trichomes is relatively uniform. Peltate glands have a diameter of around 80-90 μm (**Figure 13**).

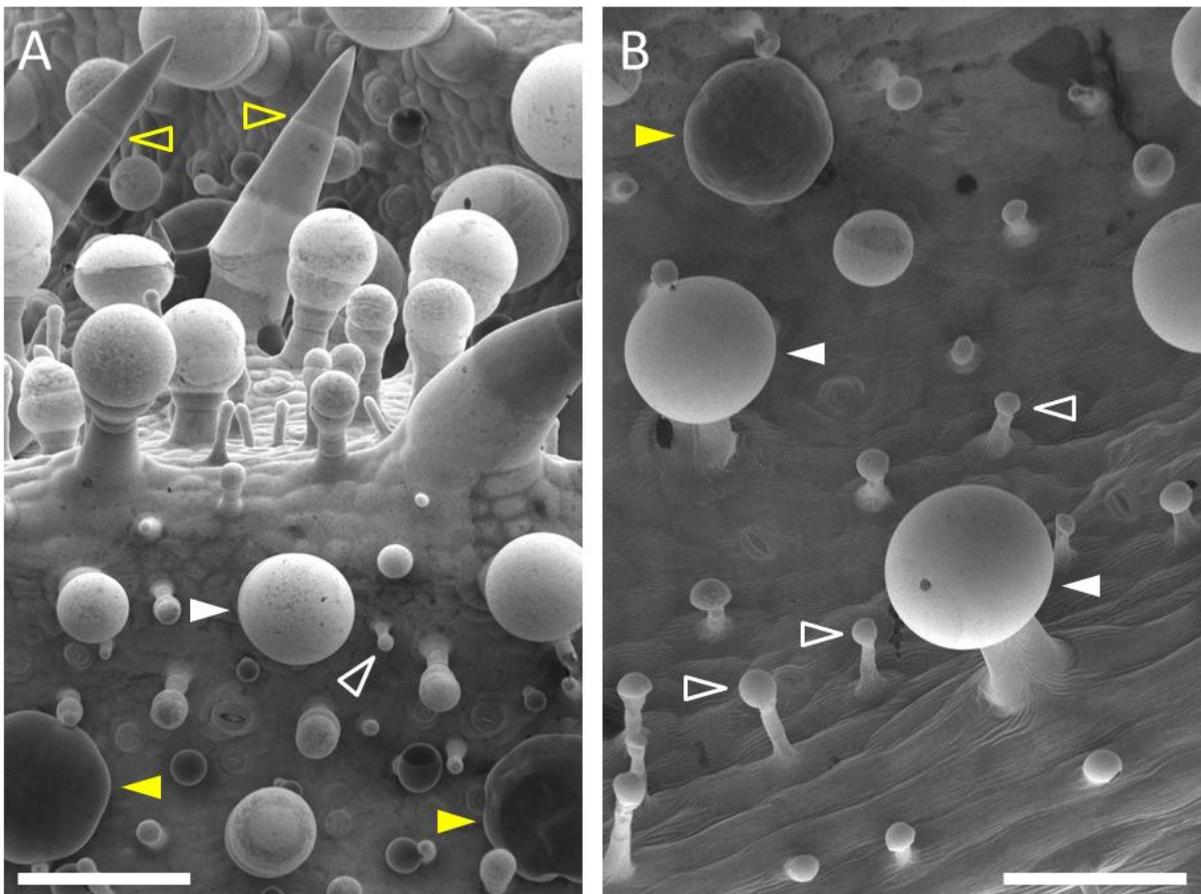


Figure 13: Glandular trichomes of clary sage calyces observed by scanning electron microscopy. **A:** Developing calyx (stage 1). **B:** Mature calyx (stage 5). Calyx development stages: see page 97. **Full white arrows** indicate large capitate glandular trichomes; **empty white arrows** indicate small capitate glandular trichomes; **full yellow arrows** indicate peltate trichomes; **empty yellow arrows** indicate non-glandular trichomes. Scale bars: 100 μm .

Capitate and peltate glandular trichomes are present on most aerial organs of clary sage, including leaves (mostly on the abaxial face), floral stems and flower organs (Ozdemir and Senel, 1999). Their density is particularly high on flower organs, namely the bract, the corolla and especially the calyx (Lattoo et al., 2006). Proportion of each glandular trichome type varies between the different organs. On calyces, large capitate glandular trichomes are more abundant than peltate trichomes, with a ratio of approximately 10:1 (Schmiderer et al., 2008). According to scanning electron microscopy observations of clary sage calyces at different developmental stages, glandular trichome density was reported to be maximum at bud stage and to gradually decrease during calyx development, probably because of pavement cell expansion (Lattoo et al., 2006). Glandular trichomes are not fully expanded at bud stage and glands progressively increase in size during calyx development (Ozdemir and Senel, 1999; Lattoo et al., 2006).

Single-gland analyses of glandular trichome content show that capitate and peltate glandular trichomes accumulate the same chemicals, but in different proportions: capitate glands tend to produce more monoterpenes and less sesquiterpenes than peltate glands (Schmiderer et al., 2008). Quantitative variations of terpene production between capitate and peltate glandular trichomes have also been evidenced in peppermint, another plant of the Lamiaceae family (Werker, 1993). Moreover, single-gland analyses performed on clary sage highlight important variations between glands of the same trichome type (Schmiderer et al., 2008). The proportion of compounds accumulated in glandular trichomes of the same type is different between leaves and floral organs. Notably, floral glandular trichomes accumulate more monoterpenes involved in pollinator attraction like linalool and linalyl acetate. In all organs of clary sage, both capitate and peltate glandular trichomes produce sclareol. Depending on the organ, sclareol represents in average 60.5-76.6% and 21.2-48.6% of capitate and peltate glandular head content, respectively (Schmiderer et al., 2008).

4.3. Genetic control of glandular trichome development

See review in the **Annex** (page 131).

5. Terpene production by clary sage

5.1. Essential oil production in aerial parts

So far, terpene production by clary sage and its regulation have mainly been investigated through the study of essential oil yield and composition. Although clary sage oil chemical composition shows important variations between studies, linalool and its acetylated derivative linalyl acetate are generally reported as being its main constituents (Lattoo et al., 2006; Schmiderer et al., 2008; Wagner et al., 2012; Sharopov and Setzer, 2012; Nasermodeli and Rowshan, 2013). The composition of clary sage essential oils produced in different countries of the Mediterranean basin have been compared and chemotypes have been identified according to the proportion of each compound in the essential oil (Sharopov and Setzer, 2012). Variations observed between clary sage chemotypes can probably be attributed to natural genetic diversity and to the influence of growth conditions. The impact of environmental conditions is illustrated by a study demonstrating that clary sage essential oil production is affected by shade level (Kumar et al., 2013). Shade level increases when the distance between plants decreases; in the study, it was artificially fixed using green shading nets. Essential oil yield was lower in shade conditions compared to open-sun conditions, and oil composition varied upon shade level: linalool content was highest in open-sun conditions and decreased with shade, whereas the opposite was observed for the sesquiterpene germacrene D (Kumar et al., 2013).

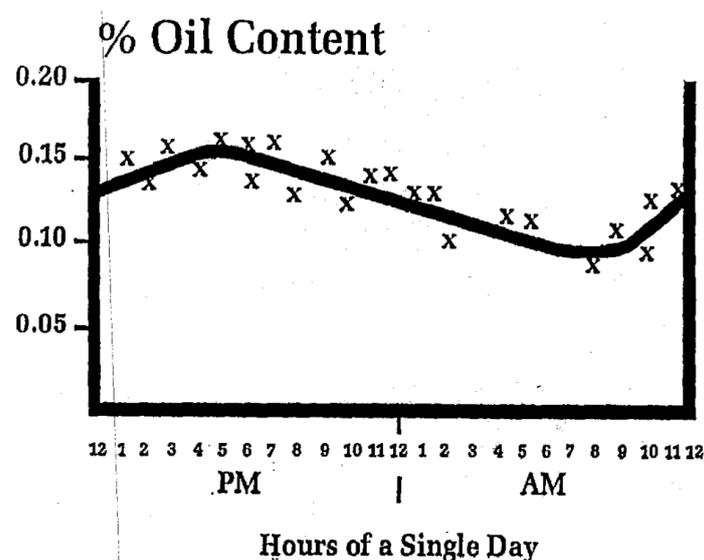


Figure 14: Diurnal variations in clary sage essential oil yield (Lawrence, 1994).

In a single clary sage plant, essential oil production is submitted to physiological variations. Notably, it was shown to vary during the day, being the lowest early in the morning, and reaching its maximum during the afternoon (Lawrence, 1994) (**Figure 14**). Moreover, essential oil yield depends on inflorescence maturity (Lattoo et al., 2006). Being initially low before flower opening, essential oil content of inflorescences gradually increases during inflorescence maturation and peaks at full bloom stage (all flowers open) before sharply declining at full maturity (withered flowers and brown seeds) (Lattoo et al., 2006). The proportion of linalyl acetate and other monoterpenes in the essential oil also reaches its maximum at full bloom stage and decreases at full maturity (Lattoo et al., 2006). By contrast, the proportion of sclareol and sesquiterpenes in the essential oil does not drop at full maturity (Lattoo et al., 2006). The sharper decline of monoterpene content may be due to higher evaporation of monoterpene compounds; indeed, these compounds are generally more volatile than sesquiterpenes and diterpenes (Lattoo et al., 2006). However, it is important to note that sclareol is poorly extracted by hydrodistillation compared to solid-liquid extraction with an organic solvent (Caissard et al., 2012). Its concentration in essential oil may not accurately reflect the overall sclareol content of the inflorescence.

5.2. Sclareol biosynthesis and secretion

Like all other diterpenes, sclareol is produced from GGPP and this conversion is performed through the action of diterpene synthases. Several enzymes involved in sclareol biosynthesis from GGPP have been cloned and characterized by three different research groups (Caniard et al., 2012; Schalk et al., 2012; Günnewich et al., 2013). A class II diterpene synthase was shown to catalyze *in vitro* the protonation-initiated bicyclization of GGPP along with the capture of a hydroxyl ion at position 8, resulting in a labdane-related intermediate: labda-13-en-8-ol diphosphate (LPP) (Caniard et al., 2012; Schalk et al., 2012; Günnewich et al., 2013) (**Figure 15**). Besides, a class I diterpene synthase was shown to produce sclareol from LPP *in vitro* after cleavage of the diphosphate group and an additional hydroxylation at position 13 (Caniard et al., 2012; Schalk et al., 2012) (**Figure 15**). These two enzymes were called labda-13-en-8-ol diphosphate-synthase (LPPS) and sclareol-synthase (SS), respectively (Caniard et al., 2012). When combined *in vitro* or co-expressed in yeast, they were responsible for the production of sclareol in two steps from GGPP (Caniard et al., 2012; Schalk et al., 2012). Both

enzymes possess an N-terminal signal sequence associated with chloroplast localization and were localized in chloroplasts according to transient expression experiments of GFP-fusion constructs in tobacco (Caniard et al., 2012). LPPS and SS are hydroxylating terpene synthases, a function which is quite rare among terpene synthases in angiosperms (Caniard et al., 2012). Hydroxyl groups on the carbon skeleton of terpenes are usually obtained after oxidation by cytochrome P450 monooxygenases (Bathe and Tissier, 2019).

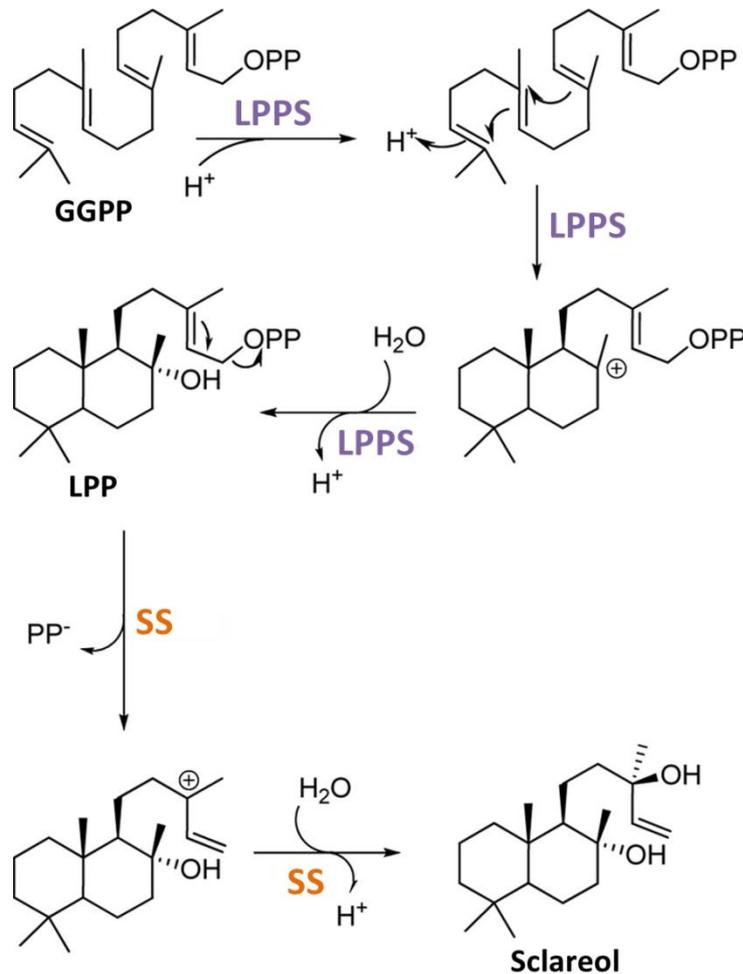


Figure 15: Molecular mechanism of sclareol biosynthesis (adapted from Schalk et al., 2012). **GGPP**, geranylgeranyl pyrophosphate; **LPP**, labda-13-en-8-ol diphosphate; **LPPS**, LPP-synthase; **SS**, sclareol-synthase.

Sclareol is known to accumulate on the epidermis of clary sage calyces (Caissard et al., 2012), but the identity of cells secreting sclareol is still discussed. Sclareol is generally thought to be produced by glandular trichomes like other essential oil components, but a study focused on sclareol localization at the surface of the calyx suggests that it may be secreted by other

epidermal cells (Caissard et al., 2012). Before accumulating at the surface of the epidermis, sclareol must be transported through the plasma membrane, the cell wall and the cuticle, but the molecular mechanism underlying sclareol secretion remains unknown. Passive transport, active transport together with cuticular components or active transport involving ABC transporters can be hypothesized (Caissard et al., 2012).

5.3. Terpene production in roots

In addition to terpenes produced in aerial parts, clary sage plants produce diterpenes in roots. The abietane-type diterpenes aethiopinone, 1-oxoaethiopinone, salvipisone and ferruginol are generally the most abundant ones (Kuźma et al., 2009a; Vaccaro et al., 2014, 2017; Alfieri et al., 2018) (**Figure 16**). These diterpenes display a range of interesting pharmacological activities, notably antibacterial, anti-proliferative, anti-inflammatory and antifungal properties (Alfieri et al., 2018). For example, salvipisone and aethiopinone display bactericidal activity against *Staphylococcus aureus* (Kuźma et al., 2009a), and aethiopinone is active against different solid tumor cell lines, particularly human melanoma cells (Vaccaro et al., 2014). Given their potential medical use, several studies have focused on enhancing their biosynthesis in clary sage hairy root cultures (Kuźma et al., 2009a; Vaccaro et al., 2014, 2017; Alfieri et al., 2018). However, the biological function of clary sage root diterpenes remains unknown.

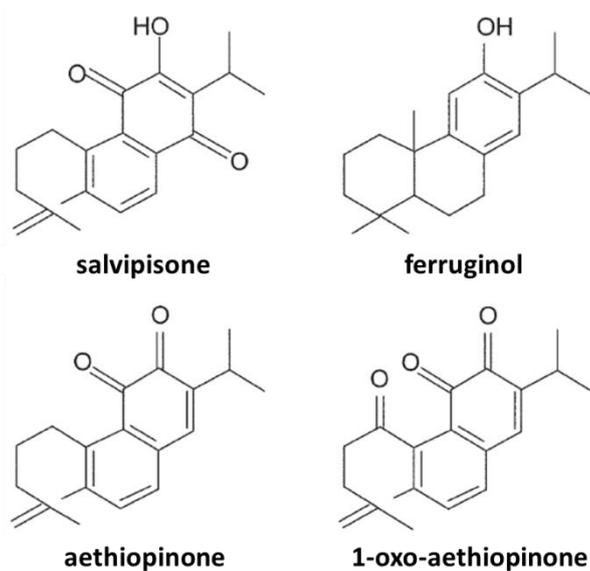


Figure 16: Chemical formulas of the main abietane diterpenes found in clary sage roots.

6. Aim of the project

Multiple strategies aiming at increasing terpenoid production in plants through targeted genetic improvement have been developed in diverse plant species like *Nicotiana benthamiana* or *Artemisia annua* (Xiao et al., 2016; Lu et al., 2016; Reed and Osbourn, 2018). The success of these targeted approaches depends on diverse parameters which strongly differ in each case, such as the process of accumulation of the terpenoid of interest, the activity of related metabolic pathways and the nature of regulation mechanisms. A better understanding of sclareol production in clary sage would be highly valuable to provide guidance for future breeding projects dedicated at improving clary sage performances.

When the metabolite of interest is secreted by glandular trichomes, manipulating plant development to increase glandular trichome density can significantly increase the production of this metabolite (Xiao et al., 2016). This strategy can be successful only if the production of the metabolite of interest is localized in glandular trichomes. The presence of sclareol has been reported in glandular trichomes of clary sage calyces (Schmiderer et al., 2008), but other epidermal cells may also secrete sclareol (Caissard et al., 2012). The first chapter of this manuscript provides new data about the respective contributions of glandular trichomes and other epidermal cells to sclareol accumulation at the surface of the calyx.

The biosynthesis of all terpenoids depends on prenyl substrate supply by MVA and/or MEP pathways. The pool of substrate available for the biosynthesis of a terpenoid of interest can be enlarged by increasing the activity of MVA and/or MEP pathways, or by blocking competing pathways using the same substrate (Lu et al., 2016; Reed and Osbourn, 2018). To design such modifications of metabolic fluxes, sufficient information about metabolic pathways connected to the pathway of interest must be gathered. Notably, it is important to know whether the terpenoid of interest derives from the MVA or from the MEP pathway, and to identify potential competing pathways. When transiently expressed in tobacco leaves, enzymes involved in sclareol biosynthesis are targeted to plastids (Caniard et al., 2012), the cellular compartment where the MEP pathway is localized. However, respective contributions of MVA and MEP pathways to sclareol biosynthesis have not been reported to date. The second chapter of this manuscript focuses on the metabolic origin of sclareol and another abundant terpene, linalyl acetate, in clary sage calyces.

Terpenoid biosynthesis is activated in response to developmental or environmental stimuli through the action of specific signaling pathways. Enhancing the activity of these signaling pathways represents another possibility to increase the production of a terpenoid of interest (Lu et al., 2016). This strategy requires the prior identification of signaling pathways stimulating the biosynthesis of the terpenoid of interest. Signaling mechanisms involved in sclareol biosynthesis regulation have not been explored so far. Since the production of many terpenoids is enhanced by jasmonates (Colinas and Goossens, 2018; Wasternack and Strnad, 2019), the third chapter of this manuscript investigates the impact of jasmonates on sclareol production in clary sage calyces.

Environmental conditions and genetic polymorphisms are responsible for variations in terpenoid content between different populations of the same plant species. The selective pressure exerted by growth conditions can lead to the selection of genetic modifications modulating terpene levels and making the plant more adapted to its environment (Figueiredo et al., 2008; Moore et al., 2014). Wild clary sage naturally grows in diverse environments in countries of the Mediterranean basin and Western Asia (Wagner et al., 2012; Nasermodeli and Rowshan, 2013). Cultivated clary sage populations probably represent only a subset of clary sage natural diversity in terms of genetic variation and sclareol content. The fourth chapter of this manuscript reports the study of genetic and phenotypic diversity among Croatian wild populations of clary sage. The exploration of this natural diversity may ultimately lead to the identification of new ways to increase clary sage productivity.

Chapter 1

Mass spectrometry imaging to localize sclareol biosynthesis

1. Introduction

Information about the precise localization of sclareol in clary sage has already led to improvements in clary sage cultivation practices. Since more than 90% of sclareol is contained in floral scapes and less than 10% in the rosette, the cutting height used for clary sage harvest was increased from 20 cm to 70 cm. This modification enhanced sclareol yield from 1.67% to 2.25% (source: Philippe Gallois, 7èmes Rencontres du Végétal, 2013). In the perspective of clary sage breeding, knowing the fact that sclareol is mainly present in floral parts suggests that increasing the reproductive/vegetative biomass ratio may be an efficient way to improve clary sage productivity. Therefore, obtaining more knowledge concerning the process of sclareol accumulation in clary sage is key to improve the yield of sclareol production and to design new clary sage breeding strategies. Today, the identity of cells involved in sclareol production and the mechanism of sclareol secretion remain elusive (Caissard et al., 2012). The purpose of this chapter is to improve our understanding of sclareol production by clary sage with new information about its localization in calyces.

Calyces have been previously shown to be the organs of clary sage with the highest sclareol content (Lawrence, 1994; Caissard et al., 2012). They are also characterized by a very high density of glandular trichomes (Lattoo et al., 2006), which are known to produce high amounts of terpenes in diverse plant species (Lange and Turner, 2013). Hence, sclareol is generally thought to be produced in these specialized epidermal structures. This hypothesis is reinforced by single gland GC-MS analyses of glandular trichome content performed by Schmiderer et al. (Schmiderer et al., 2008). Although gland content seemed to be highly variable from gland to gland, including between glands of the same type, subcuticular storage cavities of capitate and peltate glandular trichomes generally contained substantial amounts of sclareol. In average, the content of calyx capitate and peltate glands consisted of 61.1% and 48.6% sclareol, respectively, suggesting that sclareol is synthesized by both glandular trichome types (Schmiderer et al., 2008).

A few years later, a study analyzed sclareol localization at the surface of the calyx using a set of microscopy and analytical chemistry techniques (Caissard et al., 2012). This study highlighted the presence of abundant epicuticular crystals at the surface of the calyx (**Figure 17**) and identified these crystals as being composed of almost pure sclareol. The authors hypothesized that sclareol may be produced by all epidermal cells, like other cuticular secretions, and suggested that single gland analyses performed by Schmiderer et al. may have been biased by contaminations with sclareol originating from epidermal crystals (Caissard et al., 2012).

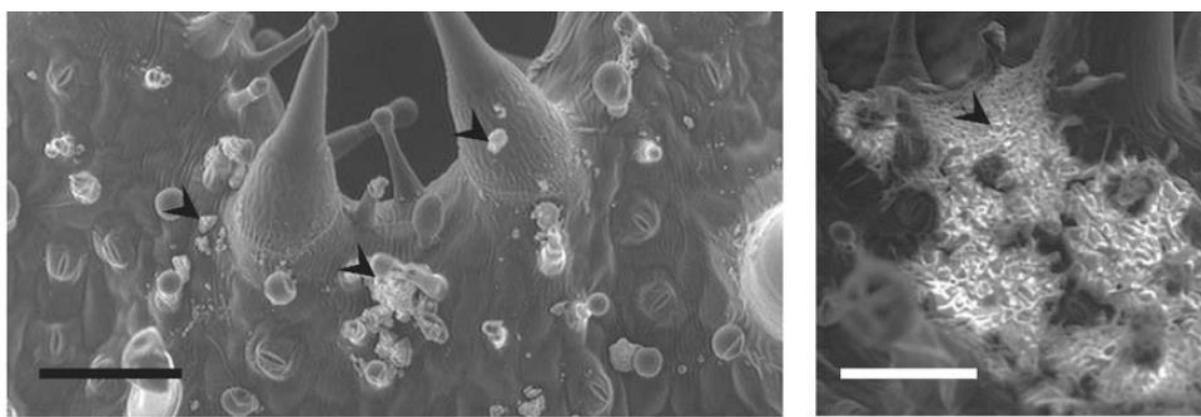


Figure 17: Epidermal crystals observed by Caissard et al. at the surface of clary sage calyces (Caissard et al., 2012).

Clary sage calyces were observed by environmental scanning electron microscopy. **Arrows** indicate sclareol crystals. Scale bars: 100 μ m.

To shed some light on the process of sclareol accumulation at the surface of the calyx, we analyzed sclareol localization with different analytical techniques than those used by Schmiderer et al. and Caissard et al. The presence of sclareol at the surface of the calyx was assessed using mass spectrometry imaging and its distribution was compared with that of linalyl acetate, a compound considered to be produced in glandular trichomes. Indeed, glandular trichomes are known to be specialized in the production of essential oil terpenes (Glas et al., 2012; Lange and Turner, 2013) and linalyl acetate is generally abundant in clary sage essential oil (Lattoo et al., 2006; Wagner et al., 2012; Sharopov and Setzer, 2012; Nasermodeli and Rowshan, 2013). This analysis was complemented by the characterization of a glabrous mutant line of clary sage.

2. Materials and methods

2.1. Plant material

Clary sage seeds from the “Vatican White” population were purchased from Jelitto (Germany). The “glabrous” clary sage mutant was generated after treatment of “Vatican White” seeds with ethyl methanesulfonate (EMS) to induce the formation of random point mutations in the genome (Drevensek et al., unpublished data). The first generation of mutants was grown in the field at the French National Conservatory of Perfume, Medicinal, Aromatic and Industrial Plants (CNPMAI) in Milly-la-Forêt, France. Seeds from this first generation were harvested in 2016 and sown in 2017 to obtain the second generation of mutants. Wild-type plants and mutants of the second generation were grown in pots for 1 month in growth chamber (16h of day; 27 °C during the day, 21 °C at night; 60% hygrometry) and then 6 months in the greenhouse (24 °C from 6 a.m. to 10 p.m., 20 °C from 10 p.m. to 6 a.m.; 60% hygrometry). Pots were then moved outdoor in December 2017 for natural vernalization and left outdoor until flowering in June-August 2018. Wild-type plants were moved back to the greenhouse in October 2018 before scanning electron microscopy and LDI-FT-ICR experiments.

2.2. Scanning electron microscopy

Mature calyces were collected from a wild-type “Vatican White” clary sage plant cultivated in the greenhouse (24 °C from 6 a.m. to 10 p.m., 20 °C from 10 p.m. to 6 a.m.; 60% hygrometry). A freshly cut calyx was introduced in the observation chamber of the scanning electron microscope (Helios NanoLab 660 instrument, FEI) without any sample preparation. Observations were performed with the voltage set to 0.35 kV.

2.3. Metabolite localization by LDI-FT-ICR

Calyx samples were collected from a wild-type “Vatican White” clary sage plant cultivated in the greenhouse (24 °C from 6 a.m. to 10 p.m., 20 °C from 10 p.m. to 6 a.m.; 60% hygrometry). The zone located at the junction of the upper and the lower lip of the calyx, measuring approximately 2 mm x 4 mm, was dissected with a scalpel and fixed onto a

conductive microscope slide with conductive double-sided tape. Samples were left on the bench for one night in a microscope slide box. Optical images showing top views of calyx samples were taken the day after sample preparation using a zoom stereomicroscope (AxioZoom V16, Zeiss). Samples were then analyzed the same day by mass spectrometry imaging. Laser-desorption ionization was performed using the MALDI source of the Solarix XR 9.4 Tesla Fourier-Transform Ion Cyclotron Resonance (FT-ICR) mass spectrometer (Bruker). Sampling was performed using the Smart Beam II laser at 25% of its power to avoid sample damage. The distance between 2 sampling points was set to 25 μm . Data analysis was performed with the SCiLS Lab 2019b Pro software (Bruker). Data were normalized with the total ion current (TIC). To improve data visualization, for each ion of interest, the 0.3% spots with the highest signal were all displayed with the maximum brightness level using the “hotspot removal” function of the software.

2.4. Sclareol and linalyl acetate quantification by GC-MS

Clary sage calyces were immersed in hexane without grinding to extract metabolites present at the surface. 1.5 mL of solvent was used to extract metabolites from 1 mature calyx. 10-undecen-1-ol was employed as internal standard and was added in the solvent before extraction at a concentration of 0.75 mM. After 2 h of agitation, 1 mL of extract was centrifuged at 13,000 rpm for 10 min to eliminate potential dust or debris. Samples were directly analyzed by gas chromatography coupled to mass spectrometry (GC-MS) using a 7890B/5977A GC-MS system from Agilent and according to a protocol adapted from Laville et al. (Laville et al., 2012). The system was equipped with a 10 m guard column and an Rxi-5Sil MS column (length, 30 m; inner diameter, 0.25 mm; film thickness, 0.25 μm) (Restek, Bellefonte, PA, USA). 1 μL of sample was injected with a split ratio of 30:1. Oven temperature was set to 110 $^{\circ}\text{C}$ for 2 min, then increased to 270 $^{\circ}\text{C}$ at a rate of 10 $^{\circ}\text{C}/\text{min}$, and finally set to 270 $^{\circ}\text{C}$ for 2 min. Other temperatures were set as follows: injector, 270 $^{\circ}\text{C}$; transfer line, 270 $^{\circ}\text{C}$; source, 230 $^{\circ}\text{C}$; quadrupole, 150 $^{\circ}\text{C}$. The carrier gas was helium at a constant flow of 1 mL/min. The quadrupole mass spectrometer was switched on after a solvent delay of 5 min and was programmed to scan from 35 u to 350 u. Data analysis was carried out using the MassHunter Quantitative Analysis software from Agilent. Sclareol and linalyl acetate were identified by comparison with analytical standards purchased from Sigma-Aldrich. Absolute

quantification of sclareol and linalyl acetate was performed using a calibration curve prepared with analytical standards.

3. Results

3.1. Different accumulation patterns for sclareol and linalyl acetate

The chemical composition of clary sage calyx surface was analyzed by mass spectrometry imaging. Ions were generated from calyx surface using the laser-desorption ionization (LDI) method. In most cases, this method requires the application of a matrix on the sample to facilitate compound ionization (matrix-assisted laser-desorption ionization, MALDI). In our case, compounds of interest were detectable without matrix and calyx samples were analyzed in a state as close as possible to their native state. Ion detection and quantification was performed by Fourier-transform ion cyclotronic resonance (FT-ICR), a high-resolution mass spectrometry technique allowing reliable identification of compounds in a complex mixture of chemicals such as plant epidermis surface.

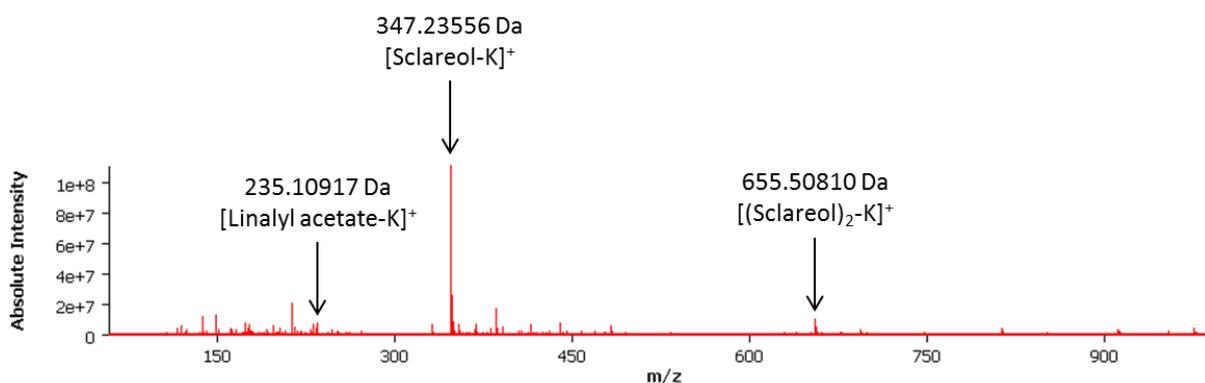


Figure 18: Average mass spectrum of the surface of a clary sage calyx.

The surface of a mature clary sage calyx was analyzed by mass spectrometry imaging (LDI-FT-ICR). This mass spectrum corresponds to the average of mass spectra obtained at each position of the calyx sample.

The most abundant ion on all calyx samples was an ion of exact mass equal to 347.23556 Da (**Figure 18**). Analysis of sclareol pure standard confirmed the identity of this ion as being an adduct between a molecule of sclareol and a potassium ion ($[\text{sclareol-K}]^+$). In addition to $[\text{sclareol-K}]^+$, we focused our analysis on two other ions of exact masses equal to 235.10917 Da and 655.50810 Da, respectively. Analysis of linalyl acetate pure standard indicates that

the ion of 235.10917 Da is an adduct between a molecule of linalyl acetate and a potassium ion ($[\text{linalyl acetate-K}]^+$). According to its exact mass, the ion of 655.50810 Da is an adduct between two molecules of sclareol and a potassium ion K^+ ($[(\text{sclareol})_2\text{-K}]^+$).

$[\text{linalyl acetate-K}]^+$ ions are mainly present in the form of spots which often coincide with a capitate glandular trichome on the optical image (**Figure 19a-b and 20a-b**, full white arrows). This observation suggests that linalyl acetate is produced by capitate glandular trichomes. This result is in accordance with the literature, which indicates that essential oil components are generally produced in glandular trichomes (Glas et al., 2012; Lange and Turner, 2013).

Surprisingly, $[\text{sclareol-K}]^+$ ion distribution differs from $[(\text{sclareol})_2\text{-K}]^+$ (**Figure 19c-d and 20c-d**). $[\text{Sclareol-K}]^+$ ions are found over most of the calyx surface, particularly on the veins bordering the analyzed zone. By contrast, $[(\text{sclareol})_2\text{-K}]^+$ ions are detected as spots which are often associated with a capitate glandular trichome, in a similar way than $[\text{linalyl acetate-K}]^+$ ions. This discrepancy between $[\text{sclareol-K}]^+$ and $[(\text{sclareol})_2\text{-K}]^+$ distributions could be explained by a higher sclareol/ K^+ ratio at the surface of capitate glandular trichomes compared to the rest of the epidermis. Indeed, $[(\text{sclareol})_2\text{-K}]^+$ ions may form upon laser impact only if many sclareol molecules are present around a single K^+ ion. The sclareol/ K^+ ratio could be higher at the surface of capitate glandular trichomes if K^+ ions are less abundant than on the rest of the epidermis, for reasons which remain to be elucidated. Conversely, the sclareol/ K^+ ratio could also be higher if sclareol is more abundant at the surface of capitate glandular trichomes than on the rest of the epidermis. This would be the case if capitate glandular trichomes produce more sclareol than other epidermal cells. Thus, the presence of $[(\text{sclareol})_2\text{-K}]^+$ spots associated with capitate glandular trichomes supports the hypothesis that sclareol is mainly produced by glandular trichomes. However, since $[\text{sclareol-K}]^+$ ions are detected all over calyx surface, the hypothesis that other epidermal cells contribute to sclareol production cannot be excluded.

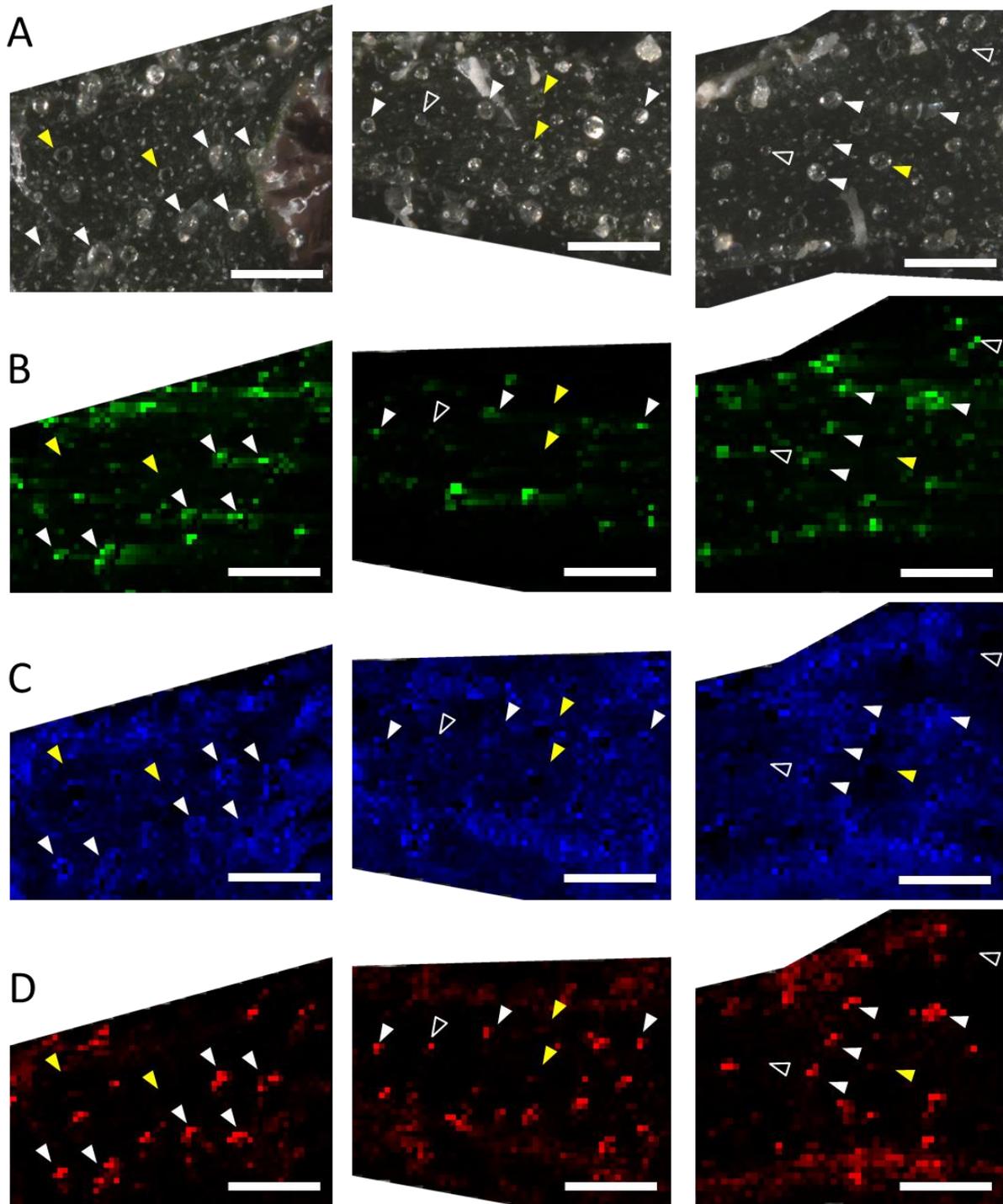


Figure 19: Localization of sclareol and linalyl acetate at the surface of a clary sage calyx.

The chemical composition of the epidermal surface of a clary sage calyx was analyzed through laser desorption-ionization followed by FT-ICR mass spectrometry (LDI-FT-ICR).

A: Top views of three clary sage calyx samples observed with a zoom stereomicroscope. **B-D:** Spatial distribution of [linalyl acetate-K]⁺ (**green**), [sclareol-K]⁺ (**blue**) and [(sclareol)₂-K]⁺ (**red**) ions at the surface of calyx samples shown in A. **White arrows** (full or empty) indicate capitate glandular trichomes. **Yellow arrows** indicate peltate glandular trichomes. Scale bars: 500 μm.

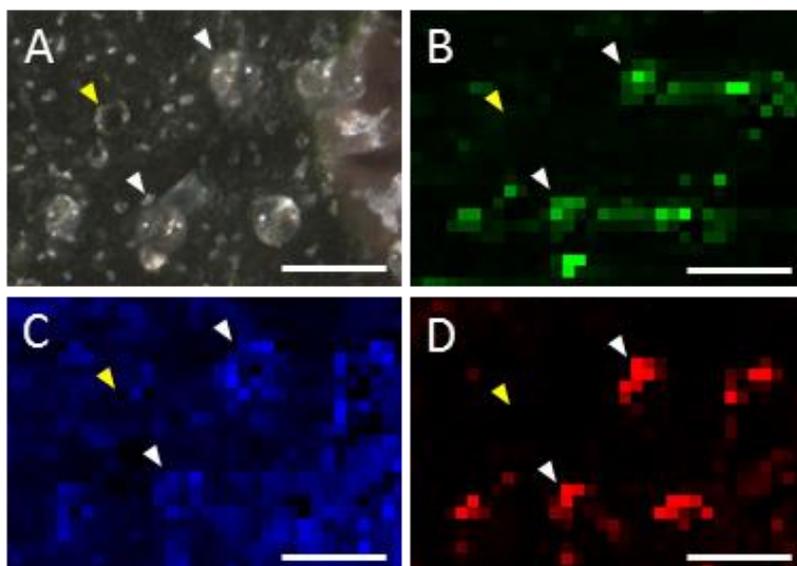


Figure 20: Localization of sclareol and linalyl acetate at the surface of a clary sage calyx (detail).

The chemical composition of the epidermal surface of a clary sage calyx was analyzed through laser desorption-ionization followed by FT-ICR mass spectrometry (LDI-FT-ICR).

A: Top view of a clary sage calyx sample observed with a zoom stereomicroscope. **B-D:** Spatial distribution of [linalyl acetate-K]⁺ (**green**), [sclareol-K]⁺ (**blue**) and [(sclareol)₂-K]⁺ (**red**) ions at the surface of the calyx sample shown in A. **White arrows** indicate capitulate glandular trichomes and **yellow arrows** indicate a peltate glandular trichome. Scale bars: 250 μm.

3.2. Distinct roles of different trichome types in sclareol production

Scanning electron microscopy observation of a freshly cut mature calyx suggests that capitulate glandular trichomes spontaneously secrete their content. Indeed, spills of electron-dense material were visible on many capitulate glandular trichome stalks (**Figure 21b**). The use of a pin to puncture a glandular head confirmed that this electron-dense material comes from capitulate glandular trichome heads (**Figure 21c-g**). If we hypothesize that sclareol is produced only by glandular trichomes, the spilling of capitulate glandular trichome content could explain why sclareol is present all over calyx epidermis while not being produced by pavement cells. The absence of [linalyl acetate-K]⁺ signal on calyx surface except on capitulate glandular trichomes could be due to the higher volatility of linalyl acetate, which probably evaporates quickly after its secretion by capitulate glandular trichomes.

Although [linalyl acetate-K]⁺ and [(sclareol)₂-K]⁺ distribution indicates that they are involved in linalyl acetate and sclareol production, capitulate glandular trichomes are not systematically associated with both [linalyl acetate-K]⁺ and [(sclareol)₂-K]⁺ signals. Some capitulate glandular trichomes visible on optical images are associated only with a [linalyl acetate-K]⁺ spot

without any $[(\text{sclareol})_2\text{-K}]^+$ signal, and conversely (**Figure 19 and 20**, empty white arrows). Metabolic activity may vary from gland to gland due to differential expression of biosynthetic enzymes. Consequently, each gland may contain specialized metabolites in different proportions, as suggested by single gland analyses performed by Schmiderer et al. (Schmiderer et al., 2008).

A large number of capitate glandular trichomes appear intact on scanning electron microscopy images, with no visible secretion along their stalks (**Figure 21a**). This observation may explain why some capitate glandular trichomes visible on optical images are not associated with any $[\text{sclareol-K}]^+$, $[\text{linalyl acetate-K}]^+$ or $[(\text{sclareol})_2\text{-K}]^+$ signal on mass spectrometry images (**Figure 19 and 20**). The impact of the laser used for compound ionization may not be sufficient to break the cuticle and access glandular head content. Similarly, no $[\text{sclareol-K}]^+$, $[\text{linalyl acetate-K}]^+$ or $[(\text{sclareol})_2\text{-K}]^+$ signal is visible at the position of most peltate glandular trichomes (**Figure 19 and 20**, yellow arrows), although these trichomes were previously shown to contain high amounts of sclareol and linalyl acetate (Schmiderer et al., 2008). The thick cuticle covering the storage cavity of peltate glandular trichomes (Tissier, 2012) may prevent their content from being ionized upon laser impact.

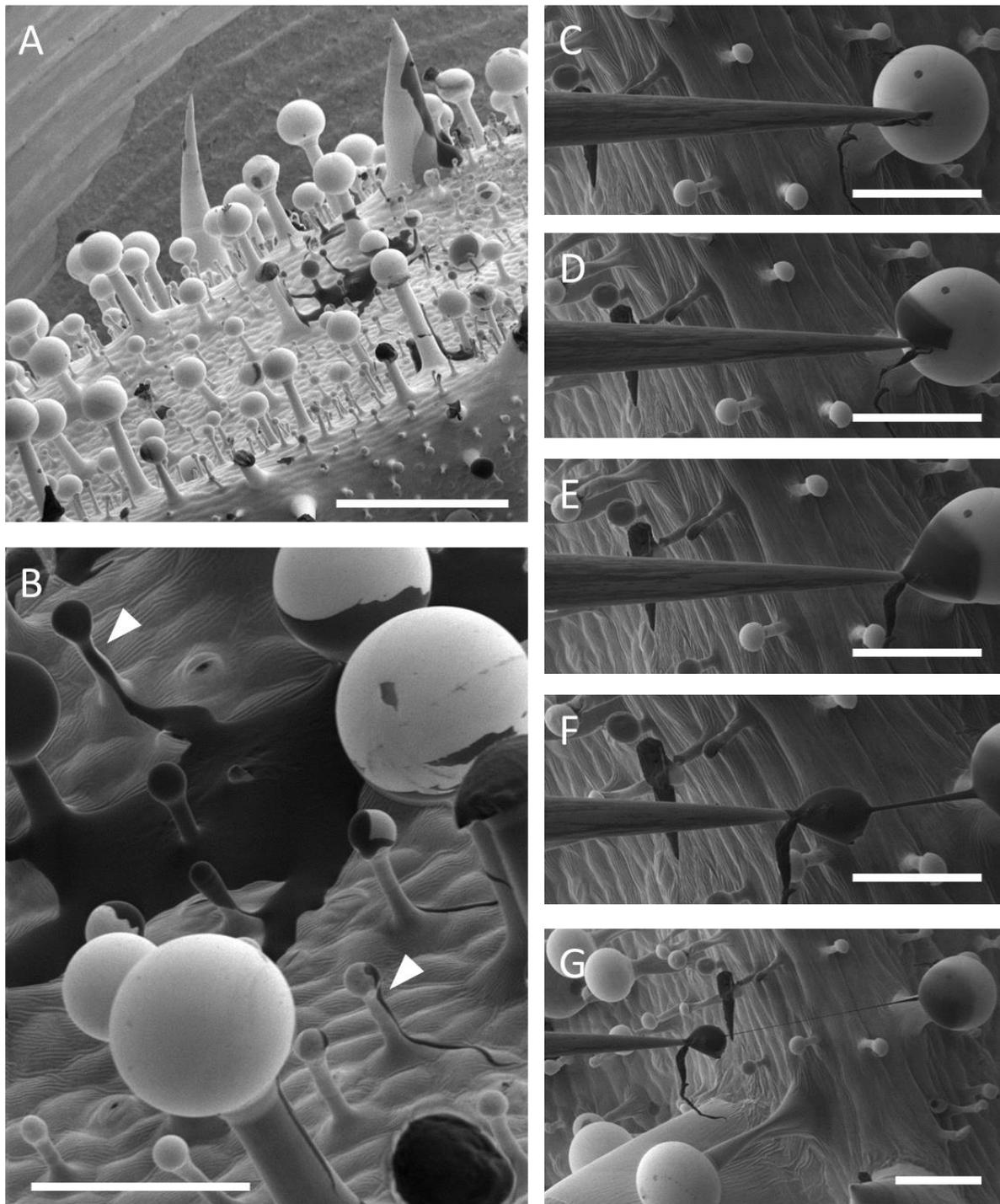


Figure 21: Epidermal surface of a clary sage calyx observed by scanning electron microscopy. Scale bars: A, 500 μm ; B-G, 100 μm . **White arrows** indicate glandular head content spilling out of the trichome.

Intriguingly, some $[\text{sclareol-K}]^+$ signal seems to be associated with non-glandular trichomes implanted on calyx veins and at the edge of the calyx (**Figure 22a-b**). Moreover, drops of electron-dense material can sometimes be seen dripping down non-glandular trichomes on scanning electron microscopy images (**Figure 22c-d**). These observations suggest that “non-glandular” trichomes may produce some sclareol and may even have a secretory activity, although this affirmation needs to be confirmed by further observations.

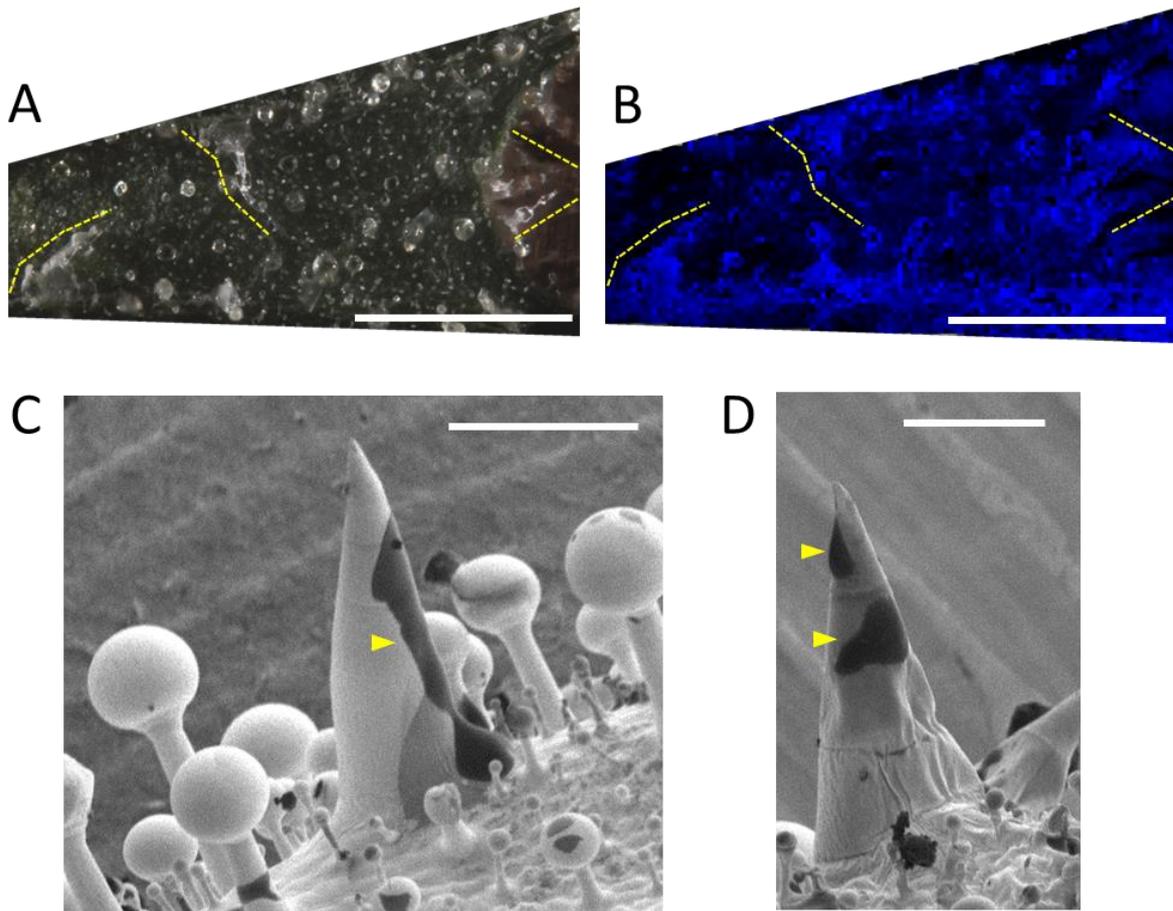


Figure 22: Potential secretory activity of calyx “non-glandular” trichomes.

A: Top view of a clary sage calyx sample observed with a zoom stereomicroscope. Scale bar: 1 mm.

B: Spatial distribution of $[\text{sclareol-K}]^+$ ions at the surface of the calyx sample shown in A, detected through laser desorption-ionization followed by FT-ICR mass spectrometry (LDI-FT-ICR). Scale bar: 1 mm. **Yellow dotted lines** indicate the position of “non-glandular” trichomes.

C-D: “Non-glandular” trichomes of a clary sage calyx observed by scanning electron microscopy. **Yellow arrows** indicate spills of electron-dense material similar to the glandular head content of capitate trichomes. Scale bars: C, 200 μm ; D, 100 μm .

3.3. A clary sage glabrous mutant produces less sclareol and linalyl acetate

A clary sage “glabrous” line was identified in a collection of clary sage mutant lines obtained by EMS-generated random mutagenesis (Drevensek et al., unpublished data) (**Figure 23a**). According to careful numbering of different trichome types (Drevensek et al., unpublished data), the “glabrous” aspect of calyces from these plants is essentially due to a drastic reduction of the number of large capitate glandular trichomes, while the number of peltate and small capitate glandular trichomes remains unchanged (**Figure 23b**). Metabolite quantification by GC-MS indicates that sclareol and linalyl acetate content is strongly reduced in glabrous calyces compared to wild-type calyces (**Figure 23c-d**). This lower metabolite content is likely due to the nearly absence of large capitate glandular trichomes, an idea reinforcing the hypothesis that capitate glandular trichomes are primarily responsible for calyx sclareol production. However, distinct EMS-induced point mutations could be the cause of the decrease in metabolite content and the reduced density of capitate glandular trichomes, respectively. The analysis of the co-segregation of these two phenotypes and the cloning of their causal mutations should shed light on this question and lead to insights in the understanding of glandular trichome development and sclareol production in clary sage.

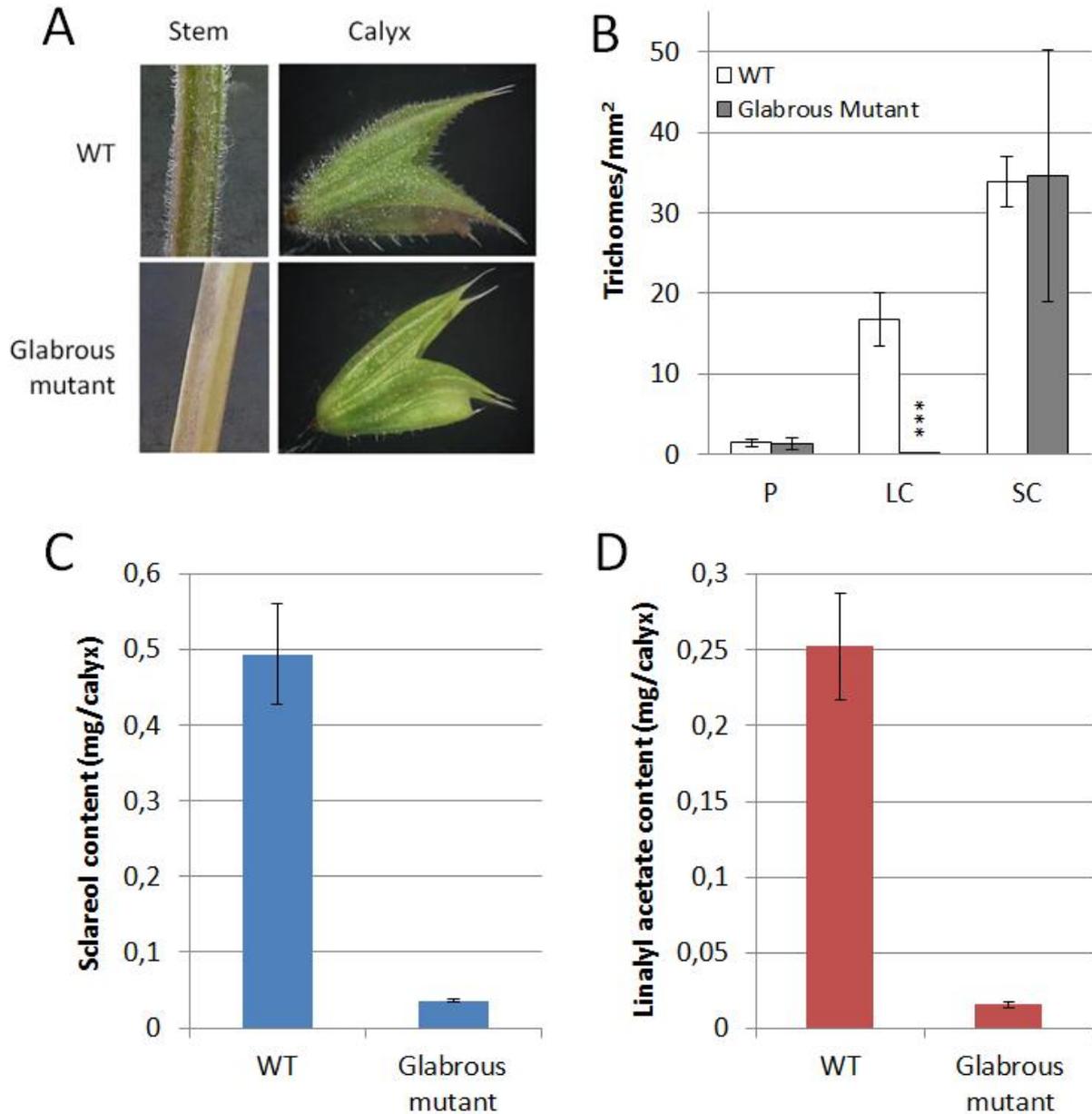


Figure 23: A glabrous mutant of clary sage displays reduced sclareol and linalyl acetate content.

A: Observation of stem and calyces from wild-type and glabrous plants with a binocular.

B: Density of different glandular trichome types at the surface of calyces from wild-type and glabrous plants. **P:** peltate glandular trichomes; **LC:** long capitate glandular trichomes; **SC:** small capitate glandular trichomes. Glandular trichomes were classified according to Schmiderer et al., 2008. Bars represent the average of 9 (WT) or 10 (glabrous mutant) measurements and error bars correspond to standard deviation. Stars indicate the results of Student's *t*-test (***: $p < 0.001$). (Drevensek et al., unpublished data)

C and D: Sclareol (**C**) and linalyl acetate (**D**) were quantified by GC-MS after hexane extraction. Bars represent the average of 3 replicates and error bars correspond to standard deviation.

4. Discussion

4.1. Sclareol is mainly produced by glandular trichomes

Glandular trichomes of the Lamiaceae are known to be specialized in the production of essential oil terpenes, mainly mono- and sesquiterpenes (Werker, 1993; Lange and Turner, 2013). In *Nicotiana* species, glandular trichomes produce diterpenes, for example sclareol in *Nicotiana glutinosa* (Guo and Wagner, 1995). These elements led the scientific community to assume that sclareol accumulated in clary sage calyces is produced by glandular trichomes. We provide here several lines of evidence reinforcing this hypothesis. First, spots of [linalyl acetate-K]⁺ and [(sclareol)₂-K]⁺ ions were detected on calyx samples by mass spectrometry imaging and coincide with capitate glandular trichomes (**Figure 19 and 20**). [Sclareol-K]⁺ ions were also detected on the rest of the epidermis (**Figure 19 and 20**). However, [(sclareol)₂-K]⁺ ions probably form only where sclareol is most abundant and may thus indicate the main sites of sclareol production. Secondly, a clary sage line with fewer capitate glandular trichomes produces less linalyl acetate and sclareol (**Figure 23**). Due to the presence of thousands of EMS-induced point mutations in this line, the two phenotypes may be caused by distinct mutations, but we can suppose that they have the same causal mutation. Since peltate and small capitate trichomes are unchanged compared to the wild-type (**Figure 23b**), this mutation may be located in a gene encoding a transcription factor specifically inducing large capitate trichome development. Taken together, these results support the idea that glandular trichomes are the main site of sclareol production. This theory is further underpinned by gene expression analyses performed on isolated glandular trichomes, showing that sclareol-synthase and LPP-synthase genes are more expressed in isolated glandular heads compared to the rest of the calyx (Drevensek et al., unpublished data) (**Figure 24**). Sclareol present on the major part of calyx epidermis, detected in the form of [sclareol-K]⁺ ions, could initially be produced in glandular heads and spread over pavement cells along with the rest of glandular head content. This process of accumulation was shown to occur in the case of the diterpene divatrienediol in *Nicotiana tabacum* (Chang and Grunwald, 1980; Keene and Wagner, 1985). Tobacco capitate glandular trichomes are known to regularly exude their content (Tissier, 2012) and our scanning electron microscopy images highlight the presence of glandular head content dripping down clary sage trichome stalks (**Figure 21**).

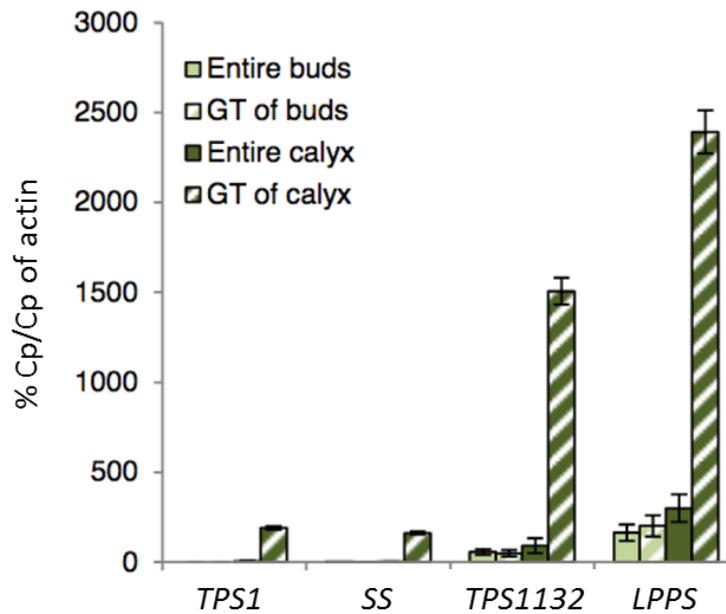


Figure 24: Expression level of sclareol biosynthesis genes in isolated glandular trichomes. (Drevensek et al., unpublished data)

Glandular trichomes (GT) were isolated from flower buds or calyces by abrasion with liquid nitrogen. Gene expression was quantified by reverse transcription-quantitative PCR (RT-qPCR) on RNAs purified from intact buds, intact calyces and purified glandular trichomes. **TPS1**: sclareol-synthase gene (Günnewich et al., 2013). **SS**: sclareol-synthase gene (Caniard et al., 2012). **TPS1132**: sclareol-synthase gene (Schalk et al., 2012). **LPPS**: LPP-synthase gene (Caniard et al., 2012). Bars represent the average of measurements performed on 2 replicates and error bars correspond to standard deviation.

Our scanning electron microscopy observations (**Figure 21**) do not confirm the presence of sclareol crystals described by Caissard et al. at the surface of calyx epidermis (Caissard et al., 2012) (**Figure 17**). We employed conventional scanning electron microscopy, while Caissard et al. performed environmental scanning electron microscopy observations. However, we did not apply any fixation or preparation protocol to calyces before observations. Samples were submitted to higher vacuum compared to environmental scanning electron microscopy, but it is unlikely that the totality of sclareol crystals were ripped off calyx surface during vacuum pumping. It is important to note that samples used in the study conducted by Caissard et al. were collected in the field, while our samples were collected on clary sage plants cultivated in the greenhouse. Less than one day after harvest, clary sage calyces collected in the field by Caissard et al. were covered with crystals, indicating that dehydration facilitates sclareol crystallization (Caissard et al., 2012). Consistently, we could

see sclareol crystals rapidly forming when calyces were heated by a lamp during observations with a binocular. The absence of crystals on our samples could be due to milder culture conditions. Indeed, culture conditions were controlled in the greenhouse (24 °C from 6 a.m. to 10 p.m., 20 °C from 10 p.m. to 6 a.m.; 60% hygrometry); it is thus likely that hygrometry was higher and temperature was lower compared to the field. Sclareol crystals observed by Caissard et al. were not evenly distributed at the surface of the epidermis and were clustered in patches (Caissard et al., 2012). Similarly, splashes of gland content that we observed by scanning electron microscopy were scattered over calyx epidermis (**Figure 21**). In the field, sclareol crystals may form progressively from splashes of gland content upon dehydration and evaporation of more volatile compounds like linalyl acetate. This hypothesis was already suggested by Caissard et al. (Caissard et al., 2012).

4.2. Some sclareol could also be produced by other epidermal cells

As previously underlined, splashes of gland content observed by scanning electron microscopy did not cover all epidermis (**Figure 21**). By contrast, [sclareol-K]⁺ signal is detected almost everywhere on calyx surface. Splashes of glandular head content may not account for all the sclareol present on the epidermis. In addition to capitate glandular trichomes, other epidermal cells may contribute to sclareol production, albeit at a lower extent compared to capitate glandular trichomes. Mass spectrometry imaging is a powerful technique, but false-negatives are possible if chemicals cannot be correctly ionized by the laser. In particular, the cuticle present at the surface of epidermal cells could prevent laser access to metabolites. Indeed, it is surprising that no sclareol and linalyl acetate signals could be associated with peltate glandular trichomes (**Figure 19 and 20**), given the fact that these trichomes are known to produce volatile monoterpenes (Lange and Turner, 2013) and were previously shown to contain large amounts of sclareol (Schmiederer et al., 2008). The analysis of the surface of glabrous calyces by LDI-FT-ICR could provide valuable information about the respective contributions of peltate trichomes, large capitate trichomes and pavement cells to sclareol production. Since large capitate trichomes are almost absent on glabrous calyces, signal potentially associated with peltate trichomes or pavement cells would be easier to distinguish.

Our results suggest that “non-glandular” trichomes may have a secretory activity. Indeed, [sclareol-K]⁺ signal seems to be associated with “non-glandular” trichomes, and electron-dense material similar to the glandular head content of capitate trichomes can sometimes be seen dripping down the sides of these trichomes (**Figure 22**). Similarly, recent publications suggest that T-shaped trichomes of *Artemisia annua*, generally considered non-glandular, may produce specialized metabolites (Soetaert et al., 2013; Judd et al., 2019). T-shaped trichomes were shown to express genes involved in artemisinin biosynthesis and mass spectrometry imaging confirmed the presence of artemisinin in these cells (Judd et al., 2019). Filamentous trichomes like T-shaped trichomes of *Artemisia annua* or “non-glandular” trichomes of clary sage are generally considered to be involved in plant mechanical defense against insects via hindering their locomotion (Soetaert et al., 2013). Even if their structure is not fully adapted to metabolite secretion, they could secrete metabolites, albeit in smaller amounts compared to glandular trichomes dedicated to this function. In clary sage, metabolites like sclareol could be secreted along with cuticular lipids via the same secretion pathway, as previously suggested (Caissard et al., 2012). Further investigation is needed to assess “non-glandular” trichome secretory activity and, if confirmed, to understand the process of metabolite secretion by these trichomes. Notably, the expression of genes involved in specialized metabolite biosynthesis could be compared in “non-glandular” and glandular trichomes sorted by laser microdissection, as performed previously in *Artemisia annua* (Soetaert et al., 2013).

4.3. Engineering higher glandular trichome density to increase sclareol yield

Since sclareol is mainly produced in glandular trichomes, increasing glandular trichome density appears to be a promising approach to enhance sclareol production in clary sage. Higher glandular trichome density can be engineered through the manipulation of transcription factors regulating glandular trichome initiation. This strategy is currently being developed to enhance the production of the anti-malarial sesquiterpene artemisinin in *Artemisia annua* (Xiao et al., 2016; Yan et al., 2017; Shi et al., 2018; Yan et al., 2018). Genes involved in artemisinin biosynthesis were previously shown to be preferentially expressed in glandular trichomes, highlighting the major role played by glandular trichomes in artemisinin

production (Xiao et al., 2016). Transcription factors positively regulating glandular trichome initiation were later identified: the MIXTA-like transcription factor AaMIXTA1 (Shi et al., 2018) and the homeodomain leucine zipper (HD-ZIP) transcription factors AaHD1 and AaHD8 (Yan et al., 2017, 2018). The overexpression of each of these positive regulators significantly enhanced glandular trichome density and also artemisinin yield, without any adverse effect on plant growth and fitness (Yan et al., 2017; Shi et al., 2018; Yan et al., 2018). The highest increase was observed after the overexpression of AaMIXTA1, which doubled artemisinin content (Shi et al., 2018).

Some regulators of glandular trichome initiation and density appear to be conserved in different plant lineages. For example, the tomato homolog of AaHD8, SICD2, is also a positive regulator of glandular trichome density in tomato (Nadakuduti et al., 2012). Homologs of conserved regulators could be looked for in the available transcriptome of clary sage calyces (Legrand et al., 2010). If further gene expression analysis shows that identified transcription factors are differentially expressed in glandular trichomes, these transcription factors would be good candidates for an involvement in glandular trichome initiation in clary sage. Enhancing or repressing the expression of positive or negative regulators of glandular trichome initiation, respectively, could increase glandular trichome density and thus sclareol yield in clary sage.

Chapter 2

A retro-biosynthetic approach to decipher sclareol metabolic origin

1. Introduction

1.1. Metabolic crosstalk between MVA and MEP pathways

In plants, MVA and MEP pathways are localized in two different subcellular compartments: the cytosol and the plast, respectively. Consequently, their activity supplies two physically separated pools of IPP and DMAPP. In accordance with the subcellular localization of biosynthetic enzymes, biosynthesis of downstream products generally depends on only one of these two pools (Lipko and Swiezewska, 2016). Most sesquiterpenoids originate from the MVA pathway, while the majority of mono- and diterpenoids are of MEP pathway origin (Reed and Osbourn, 2018).

However, a growing number of studies describe the production of terpenoids of mixed origin, i.e. made of five-carbon units originating from both pathways (Lipko and Swiezewska, 2016). Indeed, prenyl intermediates can be exchanged between the two compartments, opening up the possibility of cooperation between MVA and MEP pathways (Liao et al., 2016). Metabolite transfer from plastids to cytosol, and conversely, has been previously highlighted in BY-2 tobacco cells (Hemmerlin et al., 2003), but can also occur in whole plants, like for example in young cotton seedlings (Opitz et al., 2014). In tobacco cells, the percentage of precursors exchanged from plastids to cytosol can reach 100% and the percentage of precursors exchanged from cytosol to plastids can attain 80% (Hemmerlin et al., 2003; Liao et al., 2016). A unidirectional proton symport system has been suggested to transport prenyl diphosphates from plastids to cytosol (Bick and Lange, 2003). However, the identity of exchanged prenyl intermediates and the exact molecular mechanism responsible for their transport remain unknown (Lipko and Swiezewska, 2016).

Crosstalk between MVA and MEP pathways happens when prenyl intermediates start to accumulate in a given compartment (Liao et al., 2016). In that case, the excess of intermediates produced by one of the two pathways can be transferred to another compartment to supply the other pathway. This situation is encountered upon treatment

with an MVA- or MEP pathway-specific inhibitor, but also under particular physiological conditions in specific tissues or at given developmental stages. Notably, transport of prenyl intermediates from plastids to cytosol frequently occurs in photosynthetic organs (Tholl, 2015), whereas transport from cytosol to plastids can be substantial in the absence of light: in young cotton seedlings, the MVA pathway provides as much as 50% of the substrate used for phytol side chain and carotenoid biosynthesis (Opitz et al., 2014). It has been previously suggested that in developing seedlings, prenyl diphosphate precursors produced in the cytosol by the MVA pathway are transported to plastids to sustain carotenoid and gibberellin production (Rodríguez-Concepción et al., 2004).

1.2. Towards clary sage metabolic engineering

Metabolic engineering approaches aim at increasing the production of a compound of interest in an organism by genetically optimizing metabolic fluxes and regulatory processes. The compartmentalization of plant terpenoid biosynthesis pathways and the possibility of crosstalk between MVA and MEP pathways have already been taken in account to engineer terpenoid production in plant systems (Tholl, 2015; Lu et al., 2016; Liao et al., 2016; Mafu and Zerbe, 2018; Reed and Osbourn, 2018). A good understanding of the metabolic network including the pathway of interest is required to design a successful metabolic engineering strategy. In particular, it is important to know which metabolic pathways provide substrate to the pathway of interest, and to identify potential competing pathways (Reed and Osbourn, 2018).

Respective contributions of MVA and MEP pathways to sclareol production in clary sage have not been assessed so far. Terpene synthases responsible for sclareol biosynthesis have been fused to GFP and expressed in tobacco plants (*Nicotiana benthamiana*) in order to analyze their subcellular localization. Fusion proteins were specifically localized in plastids, suggesting that sclareol biosynthesis occurs in plastids and that sclareol precursors may be derived from the plastidial MEP pathway (Caniard et al., 2012). However, the MVA pathway may also sustain sclareol biosynthesis by producing prenyl intermediates which would then be transferred to plastids. The possibility of prenyl intermediate transport from cytosol to plastids has already been evidenced in spike lavender, another plant belonging to the Lamiaceae family (Mendoza-Poudereux et al., 2015).

The MVA contribution to the biosynthesis of mainly MEP-derived terpenoids can be high in non-photosynthetic organs like roots (Opitz et al., 2014). Clary sage calyces are photosynthetic, but it is unclear whether glandular trichomes, the multicellular epidermal structures producing the majority of calyx sclareol (see **Chapter 1**), are photosynthetic or not. Contrasted situations are encountered in other plant species. Tomato type VI trichomes are photosynthetic but still rely on sucrose imported from other leaf cells to achieve their high metabolic productivity (Balcke et al., 2017). *Artemisia annua* glandular trichomes have both photosynthetic and non-photosynthetic cells (Duke and Paul, 1993), while peltate glandular trichomes of *Mentha x piperita*, a Lamiaceae species closely related to clary sage, are not photosynthetic (Lange and Turner, 2013). If the photosynthetic activity of clary sage glandular trichomes is low, MVA-derived prenyl diphosphate precursors may be transported from the cytosol to the plast in glandular trichomes secretory cells, but this remains to be demonstrated.

To elucidate the metabolic origin of sclareol produced by clary sage calyces, we analyzed the impact of MVA and MEP pathway-specific inhibitors and carried out isotope labeling experiments based on the use of 1-¹³C-glucose and ¹³C-NMR spectroscopy. Eventually, knowing which metabolic pathways provide substrate for sclareol biosynthesis or compete with sclareol biosynthesis will help to design ways to redirect carbon fluxes towards sclareol biosynthesis.

2. Materials and methods

2.1. Plant material

Clary sage seeds from the “Vatican White” population were purchased from Jelitto (Germany). Clary sage plants were grown in pots for 1 month in growth chamber (16h of day; 27 °C during the day, 21 °C at night; 60% hygrometry) and then 3 months in the greenhouse (24 °C from 6 a.m. to 10 p.m., 20 °C from 10 p.m. to 6 a.m.; 60% hygrometry). Pots were then moved outdoor in December 2017 for natural vernalization and left outdoor until flowering in June-August 2018.

2.2. Treatment with MVA or MEP pathway-specific inhibitors

Mevinolin and fosmidomycin were purchased from Sigma-Aldrich. Stock solutions (50 mM) were prepared in ethanol and water, respectively. Nutritive solution containing 0, 1, 10 or 100 μ M of inhibitor was prepared in 5 mL vials (Murashige and Skoog basal salt mixture (Sigma-Aldrich); glucose 5% m/v (Euromedex); PPMTM 0.1% v/v (Plant Cell Technology); ethanol 0.2% v/v). Clary sage inflorescence axillary stems were cut between the 3rd and the 4th node (about 10 cm from the apex). Each stem was immediately placed in a vial in a way that its extremity was touching the bottom of the vial and dipped in nutritive solution (**Figure 25**). Treatment was performed by leaving inflorescence stems on the bench for 6 days. One calyx was sampled from each inflorescence stem, immediately frozen in liquid nitrogen and stored at -80 °C before sclareol and linalyl acetate quantification by GC-MS. Sampled calyces were at stage 2 at the beginning of the treatment and had reached stage 4 or 5 after 6 days.



Figure 25: Feeding of a clary sage inflorescence stem.

2.3. Sclareol and linalyl acetate quantification by GC-MS

Clary sage calyces were immersed in hexane or ethanol without grinding to extract metabolites present at the surface. 1.5 mL of solvent was used to extract metabolites from 1 mature calyx. 10-undecen-1-ol was employed as internal standard and was added in the solvent before extraction at a concentration of 0.75 mM. After 2 h of agitation, 1 mL of extract was centrifuged at 13,000 rpm for 10 min to eliminate potential dust or debris. Samples were directly analyzed by gas chromatography coupled to mass spectrometry (GC-MS) using a 7890B/5977A GC-MS system from Agilent and according to a protocol adapted from Laville et al. (Laville et al., 2012). The system was equipped with a 10 m guard column and an Rxi-5Sil MS column (length, 30 m; inner diameter, 0.25 mm; film thickness, 0.25 μ m) (Restek, Bellefonte, PA, USA). 1 μ l of sample was injected with a split ratio of 30:1. Oven temperature was set to 110 $^{\circ}$ C for 2 min, then increased to 270 $^{\circ}$ C at a rate of 10 $^{\circ}$ C/min, and finally set to 270 $^{\circ}$ C for 2 min. Other temperatures were set as follows: injector, 270 $^{\circ}$ C; transfer line, 270 $^{\circ}$ C; source, 230 $^{\circ}$ C; quadrupole, 150 $^{\circ}$ C. The carrier gas was helium at a constant flow of 1 mL/min. The quadrupole mass spectrometer was switched on after a solvent delay of 5 min and was programmed to scan from 35 u to 350 u. Data analysis was carried out using the MassHunter Quantitative Analysis software from Agilent. Sclareol and linalyl acetate were identified by comparison with analytical standards purchased from Sigma-Aldrich. Absolute quantification of sclareol and linalyl acetate was performed using a calibration curve prepared with analytical standards. Calyx dry weight was measured after lyophilization of dry material remaining after extraction.

2.4. Isotope labeling and analysis by 13 C-NMR or GC-MS

2.4.1. Isotope labeling and sample preparation

Nutritive solution containing $1\text{-}^{13}\text{C}$ -glucose was prepared in 5 mL vials (Murashige and Skoog basal salt mixture (Sigma-Aldrich); $1\text{-}^{13}\text{C}$ -glucose 5% m/v (Sigma-Aldrich); PPMTM 0.1% v/v (Plant Cell Technology); ethanol 0.2% v/v). Clary sage inflorescence axillary stems were cut between the 3rd and the 4th node (about 10 cm from the apex). Each stem was immediately placed in a vial in a way that its extremity was touching the bottom of the vial and dipped in nutritive solution. Feeding was performed by leaving inflorescence stems on the bench for 5

days. Young calyces (stage 4 or younger) were sampled from each inflorescence stem, immediately frozen in liquid nitrogen and stored at -80 °C before further analysis.

Calyces collected on one inflorescence stem were immersed in 8 mL of hexane without grinding to extract metabolites present at the surface. After 2 h of agitation, the extract was centrifuged at 13,000 rpm for 10 min to eliminate potential dust or debris. 300 µL of the extract were saved for GC-MS analysis of isotope cluster distortion and the rest of the extract was prepared for ¹³C-NMR analysis. Hexane was evaporated by leaving the samples under the fume hood for 2-3 days. Evaporation needed to be slow in order to avoid linalyl acetate volatilization. Dry extract was resuspended in 600 µL of CDCl₃ for ¹³C-NMR analysis.

2.4.2. Analysis of ¹³C-labeling patterns by ¹³C-NMR

¹³C Nuclear Magnetic Resonance (¹³C-NMR) experiments were performed using an Ascend 400 MHz NanoBay instrument equipped with an AVANCE III HD console and a Prodigy BBO cryoprobe (³¹P-¹⁵N) (Bruker). Analyses were performed at 277 K without tube spinning, using a proton-decoupled (decoupling sequence waltz16) carbon pulse program (zgig) with 90° pulses of 9.75 µs at 41 W for ¹³C. The relaxation delay (D1) was set to 15 s in order to enable complete acquisition of the signal corresponding to ¹³C nuclei relaxation. To get good signal-to-noise ratio, 5096 scans were done, representing 23 h 34 min of analysis per sample. Data were analyzed using the TopSpin software (Bruker). Free induction decay signal was processed by performing an exponential line broadening, calculating the Fourier transform, adjusting the phase and applying a baseline correction. The attribution of each peak of the spectrum to a carbon of sclareol or linalyl acetate (**Table 1**) was performed using ¹H, COSY, HSQC and HMBC experiments and inspired by previously published data (Fowler et al., 1999; Chain et al., 2015). Calibration curves with known concentrations of sclareol and linalyl acetate analytical standards (purchased from Sigma-Aldrich) confirmed that the acquisition was quantitative, i.e. that peak area was proportional to the amount of ¹³C present in the sample ($R^2 > 0.98$ for sclareol and $R^2 > 0.96$ for linalyl acetate). Potential incorporation of ¹³C at a given position was evaluated using the labeling ratio. This ratio was calculated by dividing peak area measured for labeled samples by peak area measured for unlabeled samples.

	Sclareol	Linalyl acetate	δ [ppm]
Peak 1		Carbon 11	170,89
Peak 2	Carbon 14		146,50
Peak 3		Carbon 6	142,51
Peak 4		Carbon 7	132,66
Peak 5		Carbon 2	124,44
Peak 6		Carbon 1	113,86
Peak 7	Carbon 15		112,11
Peak 8		Carbon 3	83,58
Peak 9	Carbon 8		75,62
Peak 10	Carbon 13		74,60
Peak 11	Carbon 9		62,20
Peak 12	Carbon 5		56,69
Peak 13	Carbon 7		45,51
Peak 14	Carbon 12		44,83
Peak 15	Carbon 3		42,63
Peak 16		Carbon 4	40,55
Peak 17	Carbon 1		40,24
Peak 18	Carbon 10		39,86
Peak 19	Carbon 18		34,20
Peak 20	Carbon 4		33,83
Peak 21	Carbon 16		27,90
Peak 22		Carbon 9	26,51
Peak 23	Carbon 17		25,00
Peak 24		Carbon 10	24,16
Peak 25		Carbons 5+12	23,01
Peak 26	Carbon 19		22,21
Peak 27	Carbon 11		21,19
Peak 28	Carbon 6		19,75
Peak 29	Carbon 2		19,11
Peak 30		Carbon 8	18,24
Peak 31	Carbon 20		16,12

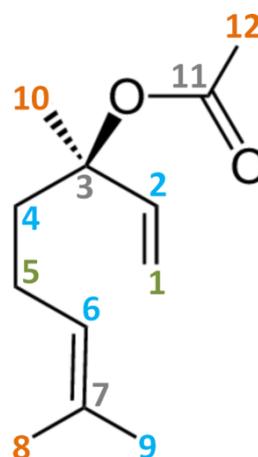
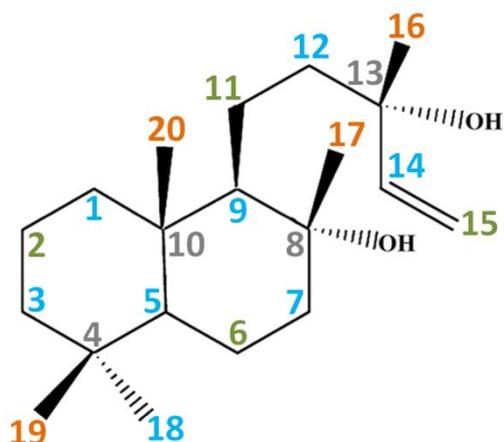


Table 1: Interpretation of the ^{13}C -NMR spectrum of clary sage calyx extracts: attribution of peaks to carbon positions in sclareol and linalyl acetate.

δ : ^{13}C -NMR chemical shift. In **blue**: carbons predicted to be labeled only if the MVA pathway is involved in IPP biosynthesis; in **green**, carbons predicted to be labeled only if the MEP pathway is involved in IPP biosynthesis; in **orange**, carbons predicted to be labeled whatever the pathway involved in IPP biosynthesis; in **grey**, carbons predicted to be unlabeled whatever the pathway involved in IPP biosynthesis.

2.4.3. Analysis of isotope cluster distortion by GC-MS

Extracts saved for GC-MS analysis were processed using the protocol described above (see page 67). During GC-MS analysis, compounds contained in the sample are fragmented and each fragment produces a signal in the form of a group of peaks known as isotope cluster. The mass spectrum of a given compound displays several isotope clusters, which correspond to the different fragments generated in the mass spectrometer. The different peaks of an isotope cluster represent ions with the same elemental composition, but different isotopic compositions: heavier peaks correspond to ions which contain one or more heavy isotopes like ^{18}O or ^{13}C . The isotope cluster is present even without artificial isotope labeling because ^{18}O and ^{13}C are present in nature, although in low abundance compared to light isotopes ^{16}O and ^{12}C (0.2 % and 1.1 % versus 99.8 % and 98.9 %). The incorporation of additional ^{13}C atoms leads to a distortion of the isotope cluster, with ions containing heavy isotopes becoming more abundant. For each compound, the distortion of several isotope clusters of the mass spectrum was evaluated by calculating their distortion ratio, which corresponds to the sum of heavy ions peak areas divided by light ion peak area. For β -caryophyllene, $134+135+136/133$ and $148+149/147$ ratios were calculated, and for linalyl acetate, $81+82+83/80$, $94+95+96/93$ and $122+123/121$ ratios were calculated. A higher distortion ratio reveals the incorporation of ^{13}C atoms in the compound.

3. Results

3.1. Sclareol and linalyl acetate are highly abundant terpenes of clary sage calyx surface

The solid-liquid extraction of sclareol from plant material is usually performed with an organic solvent like hexane (Caissard et al., 2012; Chain et al., 2015; Dufour et al., 2017). Sclareol is then quantified by GC-MS. This protocol is suitable for terpenes with a small hydrocarbon backbone harboring few decorations, displaying relatively low polarity and relatively high volatility (Jiang et al., 2016). Therefore, it was used to quantify sclareol and at the same time identify other metabolites present at the surface of clary sage calyxes.

According to GC-MS analyses, two metabolites are particularly abundant in clary sage calyx hexane extracts: sclareol and linalyl acetate (**Figure 26a**). ¹³C-NMR analysis confirmed that calyx hexane extracts are essentially composed of sclareol and linalyl acetate (**Figure 26c**). Hexane is a strongly nonpolar solvent, which preferentially extracts molecules displaying low polarity. To enable the extraction of more polar metabolites, ethanol was used instead of hexane. Again, sclareol and linalyl acetate were the main metabolites visible on the GC-MS chromatogram, respectively accounting for 54% and 43% of total peak area (**Figure 26b**). Linalool could not be detected in hexane or ethanol extracts either by GC-MS or ¹³C-NMR, although it is often reported as a major monoterpene of clary sage essential oil (Lattoo et al., 2006; Schmiderer et al., 2008; Wagner et al., 2012; Sharopov and Setzer, 2012; Nasermodeli and Rowshan, 2013). However, in a study comparing dichloromethane extracts and essential oil composition, sclareol, linalyl acetate and linalool respectively accounted for 74%, 22% and 1.5% of peak area in dichloromethane extracts, versus 6%, 63% and 18% in essential oil (Schmiderer et al., 2008). The absence or very low abundance of linalool in hexane and ethanol extracts is consistent with these previously published data.

Given the abundance of linalyl acetate in calyx extracts, we considered the hypothesis that linalyl acetate biosynthesis may significantly compete with sclareol biosynthesis for prenyl diphosphate substrates. For this reason, we focused our study on both sclareol and linalyl acetate biosynthesis and tried to decipher the metabolic origin of both metabolites.

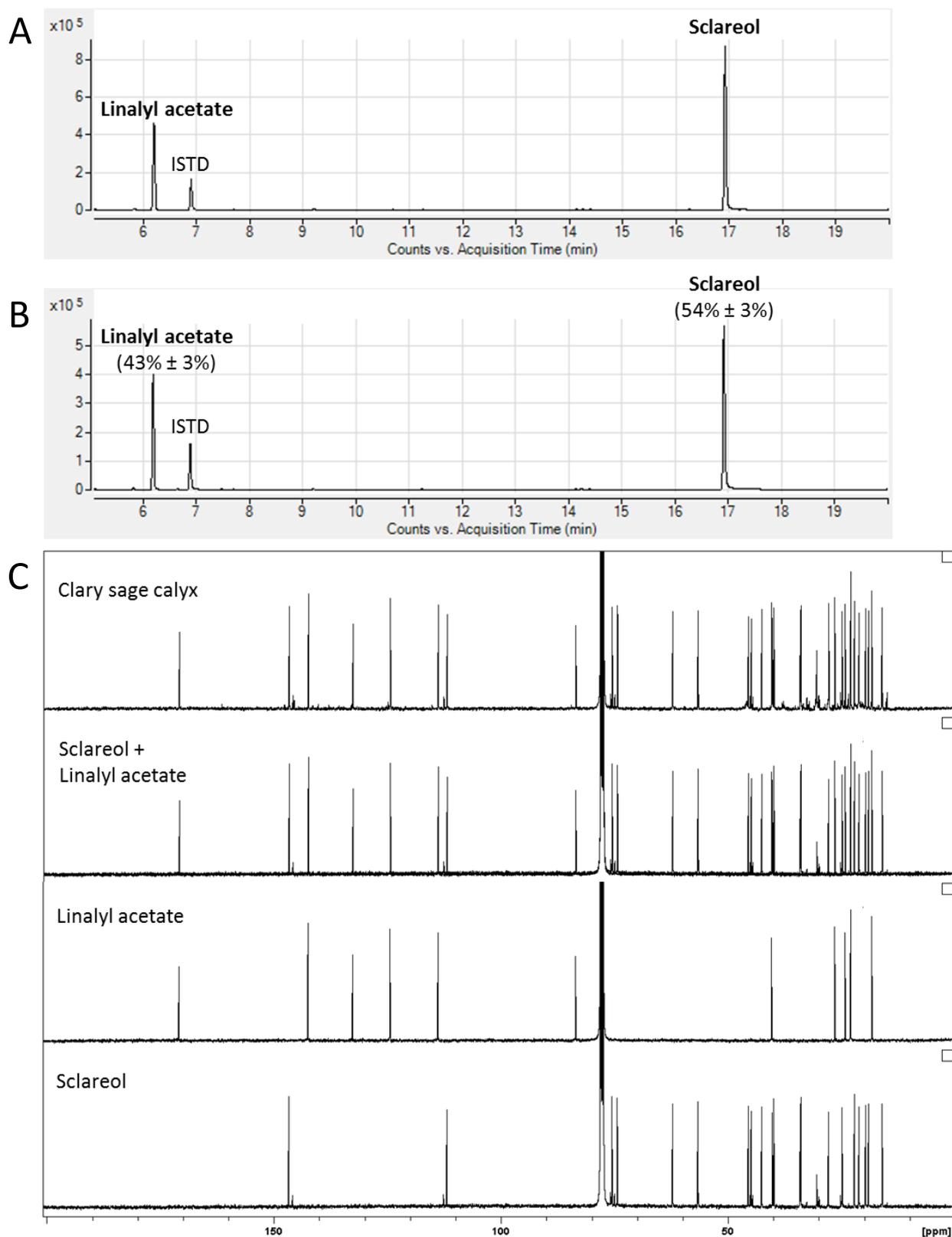


Figure 26: Analysis of the composition of clary sage calyx organic extracts by GC-MS and ^{13}C -NMR. **A, B:** GC-MS chromatograms of clary sage calyx extracts obtained after extraction in hexane (**A**) or ethanol (**B**). Values indicated in B correspond to average peak area expressed as a percentage of total peak area \pm standard deviation (8 biological replicates). **ISTD:** internal standard (10-undecen-1-ol). **C:** ^{13}C -NMR spectrum of sclareol and linalyl acetate analytical standards and of a clary sage calyx hexane extract.

3.2. Sclareol and linalyl acetate are MEP-derived terpenes

3.2.1. Dose-dependent inhibition of sclareol and linalyl acetate biosynthesis by fosmidomycin but not mevinolin

To investigate the contribution of MVA and MEP pathways to the production of IPP and DMAPP necessary for sclareol and linalyl acetate biosynthesis, we used specific inhibitors of each pathway, respectively mevinolin and fosmidomycin. Mevinolin is a competitive inhibitor of HMGR, a rate-limiting enzyme of the MVA pathway, while fosmidomycin competitively inhibits DXR, a key enzyme of the MEP pathway (Lipko and Swiezewska, 2016). Cut inflorescence stems were fed with different concentrations of inhibitor and sclareol and linalyl acetate content was quantified by GC-MS after 6 days. Mevinolin treatment does not lead to any significant modification of sclareol and linalyl acetate content (**Figure 27a**). By contrast, fosmidomycin treatment inhibits sclareol and linalyl acetate production in a dose-dependent way (**Figure 27b**). A significant drop in sclareol content can already be observed with a concentration of 1 μM of fosmidomycin in the feeding solution and the strongest inhibition is observed with 100 μM of fosmidomycin, with a decrease of 66% in sclareol content. A significant drop in linalyl acetate content (by 93%) can be observed with 100 μM of fosmidomycin. These results indicate that sclareol and linalyl acetate biosynthesis mainly depends on the MEP pathway.

However, 34% of sclareol content remains after 6 days of treatment with 100 μM of fosmidomycin (**Figure 27b**). The inhibition of the MEP pathway may be incomplete, or this amount of sclareol may have been produced before the start of the treatment. Alternatively, the MVA pathway may be activated and produce a part of the IPP and DMAPP needed for sclareol biosynthesis when the MEP pathway is inhibited by fosmidomycin. To test this hypothesis, we employed another approach to study the contribution of MVA and MEP pathways: a retro-biosynthetic approach based on the use of 1- ^{13}C -glucose.

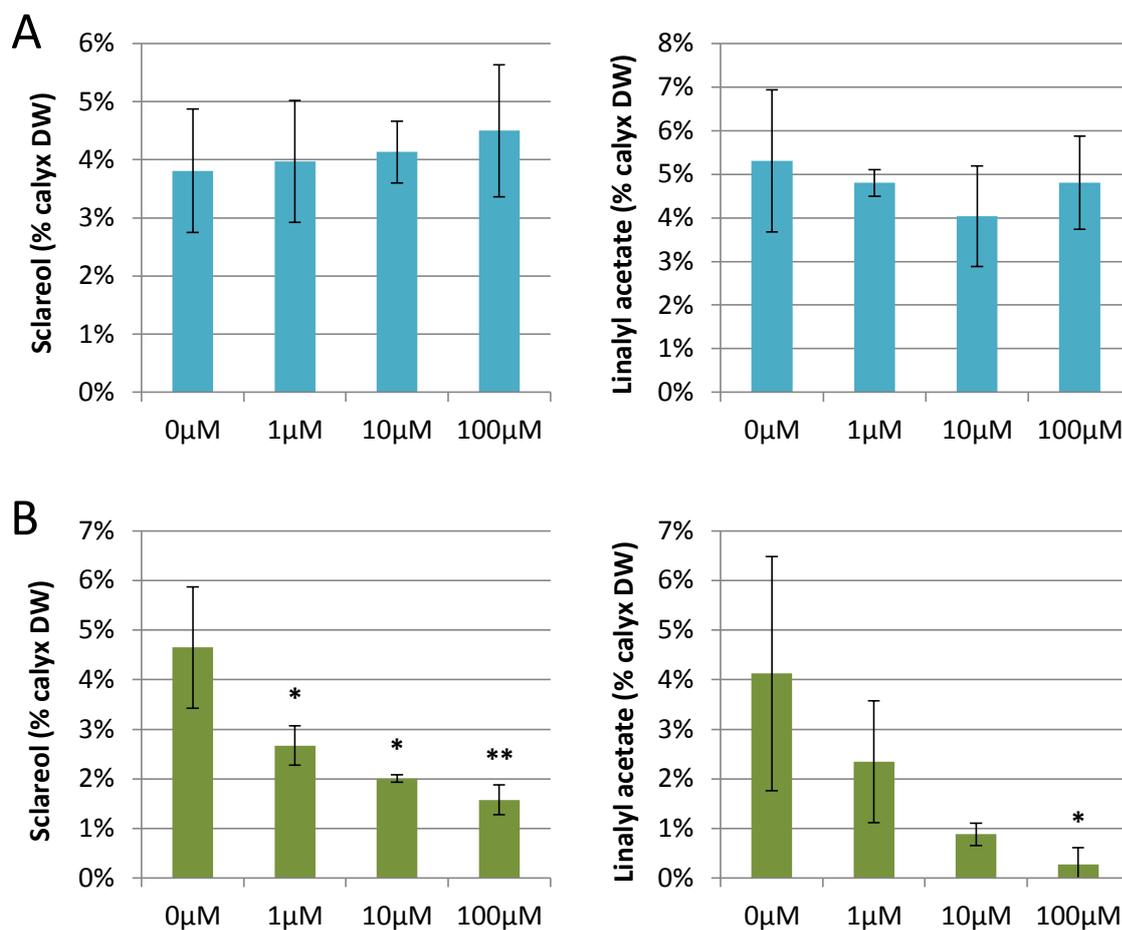


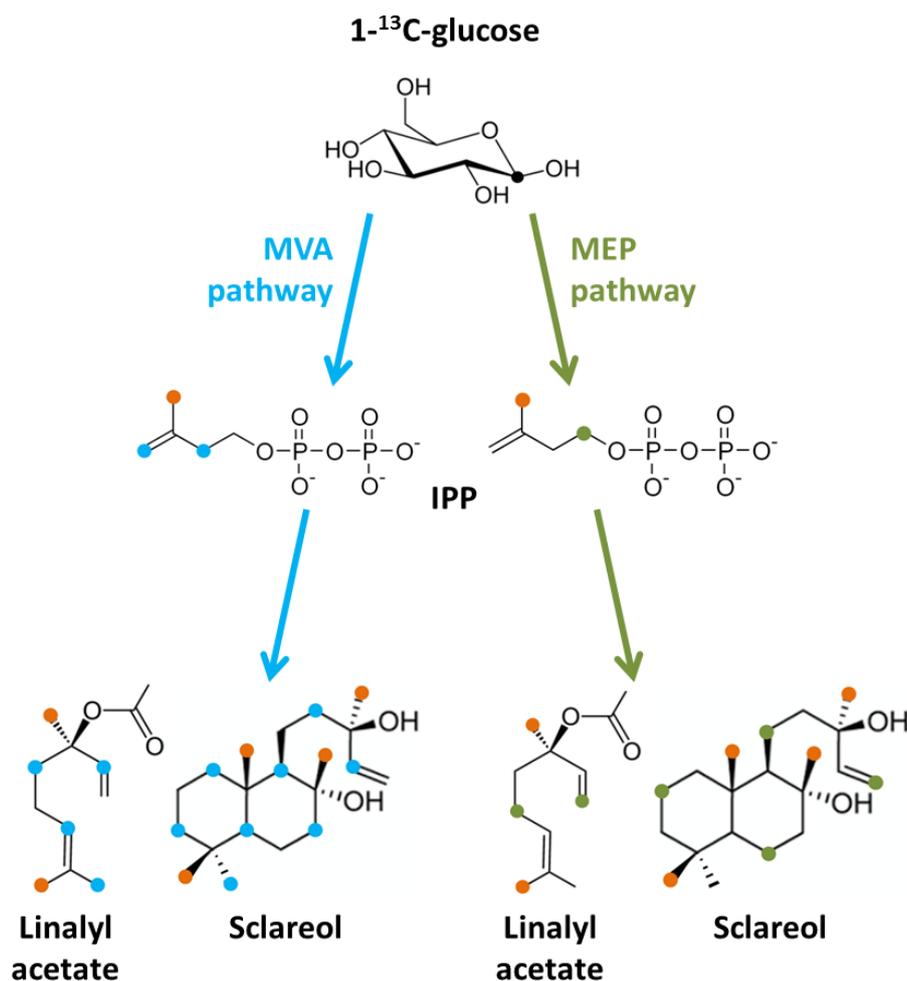
Figure 27: Impact of mevinolin or fosmidomycin treatment on sclareol and linalyl acetate content. Clary sage cut inflorescence stems were fed with a solution containing mevinolin (A) or fosmidomycin (B) for 6 days. Sclareol and linalyl acetate content was quantified by GC-MS. Bars correspond to the average of measurements performed on 4 biological replicates; error bars correspond to the standard deviation and stars indicate the results of Student's *t*-test (*: $p < 0.05$, **: $p < 0.01$).

3.2.2. Labeling with 1-¹³C-glucose suggests that sclareol and linalyl acetate are MEP-derived

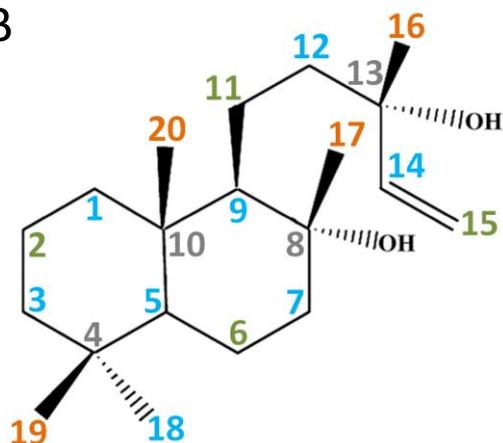
1-¹³C-glucose is a molecule of glucose labeled with ¹³C at position 1. In other words, it is a molecule of glucose where the proportion of ¹³C found at position 1 is 99% instead of the naturally occurring proportion of 1%, which is retained at all other positions. 1-¹³C-glucose can theoretically be used as a precursor by both MVA and MEP pathways to produce IPP and DMAPP. However, the enzymatic steps of each pathway are distinct and carbon atoms initially found at position 1 in glucose are not found at the same position in IPP and DMAPP after processing by each pathway (**Figure 28a**). Consequently, the theoretical ¹³C-labeling pattern obtained in IPP and DMAPP is different for each pathway: some carbons are labeled only if IPP and DMAPP have been produced by the MEP pathway, and other carbons are labeled only if the MVA pathway was involved. Theoretical ¹³C-labeling patterns expected on sclareol and linalyl acetate in each case were inferred based on enzymatic mechanisms that have been suggested in literature for enzymatic steps leading from IPP and DMAPP to sclareol (Caniard et al., 2012; Schalk et al., 2012; Günnewich et al., 2013) and linalyl acetate (Fowler et al., 1999; Zaks et al., 2008).

Cut inflorescence stems were fed with 1-¹³C-glucose and ¹³C-labeling patterns obtained on sclareol and linalyl acetate were analyzed by ¹³C-NMR. According to theoretical sclareol ¹³C-labeling patterns, carbons 2, 6, 11 and 15 should be labeled only if the MEP pathway was involved in the production of IPP and DMAPP, while carbons 1, 3, 5, 7, 9, 12, 14 and 18 should be labeled only if the MVA pathway was involved (**Figure 28b**). After 1-¹³C-glucose feeding, carbons 2, 6, 11 and 15 are strongly labeled while carbons 1, 3, 5, 7, 9, 12, 14 and 18 are weakly labeled (**Figure 29**). Interestingly, carbons 4, 8, 10 and 13, which should not be labeled whatever the pathway involved in IPP and DMAPP biosynthesis, show a weak labeling comparable to carbons specific of the MVA pathway. Similar results were obtained for linalyl acetate: carbons specific of the MEP pathway are strongly labeled whereas carbons specific of the MVA pathway are weakly labeled, and carbons which should not be labeled whatever the pathway involved also show a weak labeling (**Figure 30**). These results suggest that both sclareol biosynthesis and linalyl acetate biosynthesis mainly depend on the MEP pathway.

A



B



C

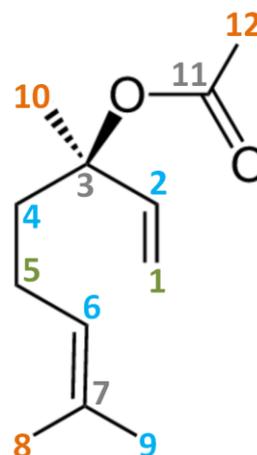


Figure 28: Theoretical ¹³C-labeling patterns of sclareol and linalyl acetate after 1-¹³C-glucose processing through MVA and MEP pathways.

In **grey**, carbons predicted to be unlabeled; in **blue**, carbons predicted to be labeled only if the MVA pathway is involved in IPP biosynthesis; in **green**, carbons predicted to be labeled only if the MEP pathway is involved in IPP biosynthesis; in **orange**, carbons predicted to be labeled whatever the pathway involved in IPP biosynthesis.

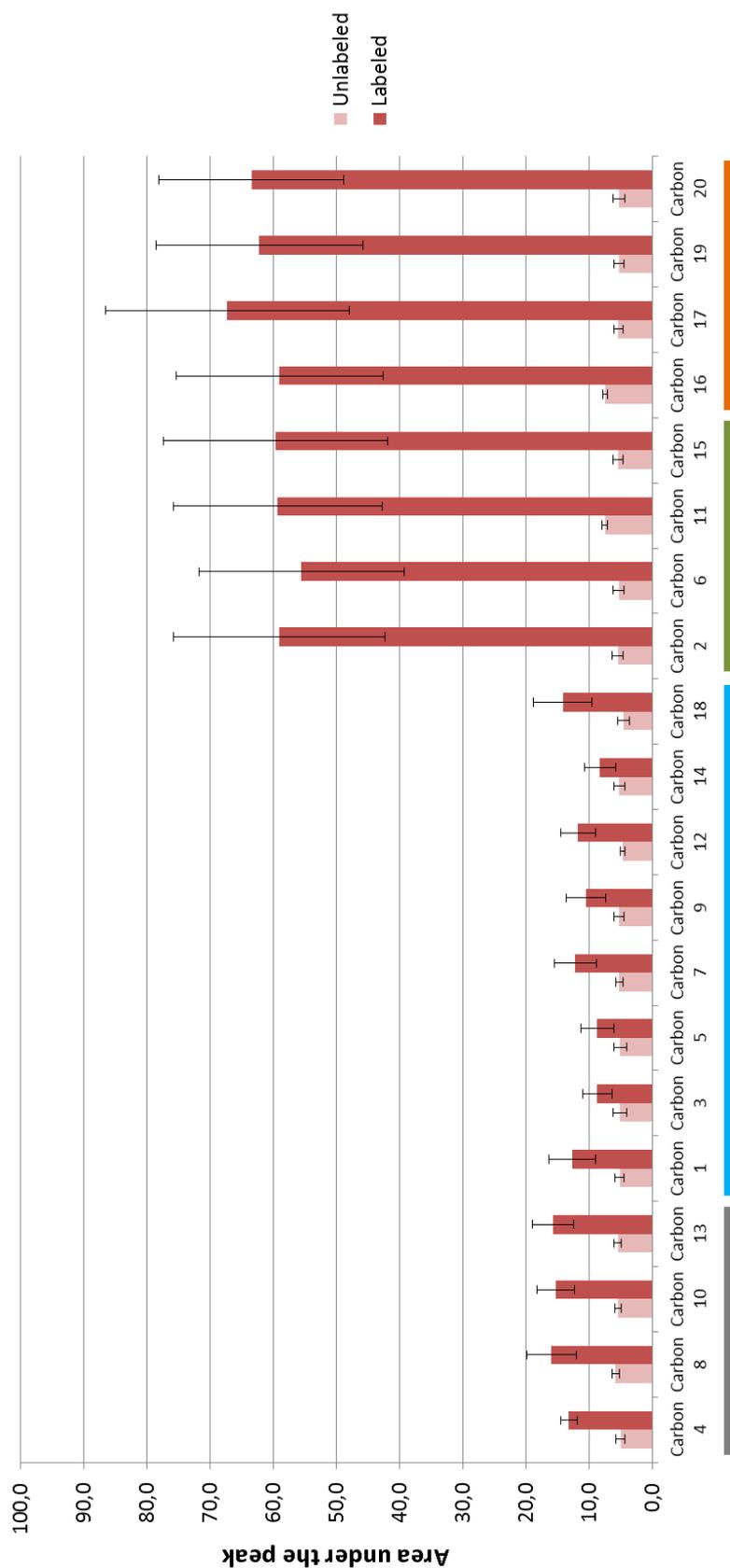


Figure 29: Sclareol ¹³C-labeling after feeding with 1-¹³C-glucose.

Clary sage cut inflorescence stems were fed with a solution containing 1-¹³C-glucose (**Labeled**) or standard glucose (**Unlabeled**) for 5 days. Sclareol ¹³C-labeling was analyzed by quantitative ¹³C-NMR after hexane extraction. Bars correspond to the average of peak area measurements performed on 3 biological replicates and error bars correspond to the standard deviation. In **grey**, carbons predicted to be unlabeled; in **blue**, carbons predicted to be labeled only if the MVA pathway is involved in IPP biosynthesis; in **green**, carbons predicted to be labeled only if the MEP pathway is involved in IPP biosynthesis; in **orange**, carbons predicted to be labeled whatever the pathway involved in IPP biosynthesis.

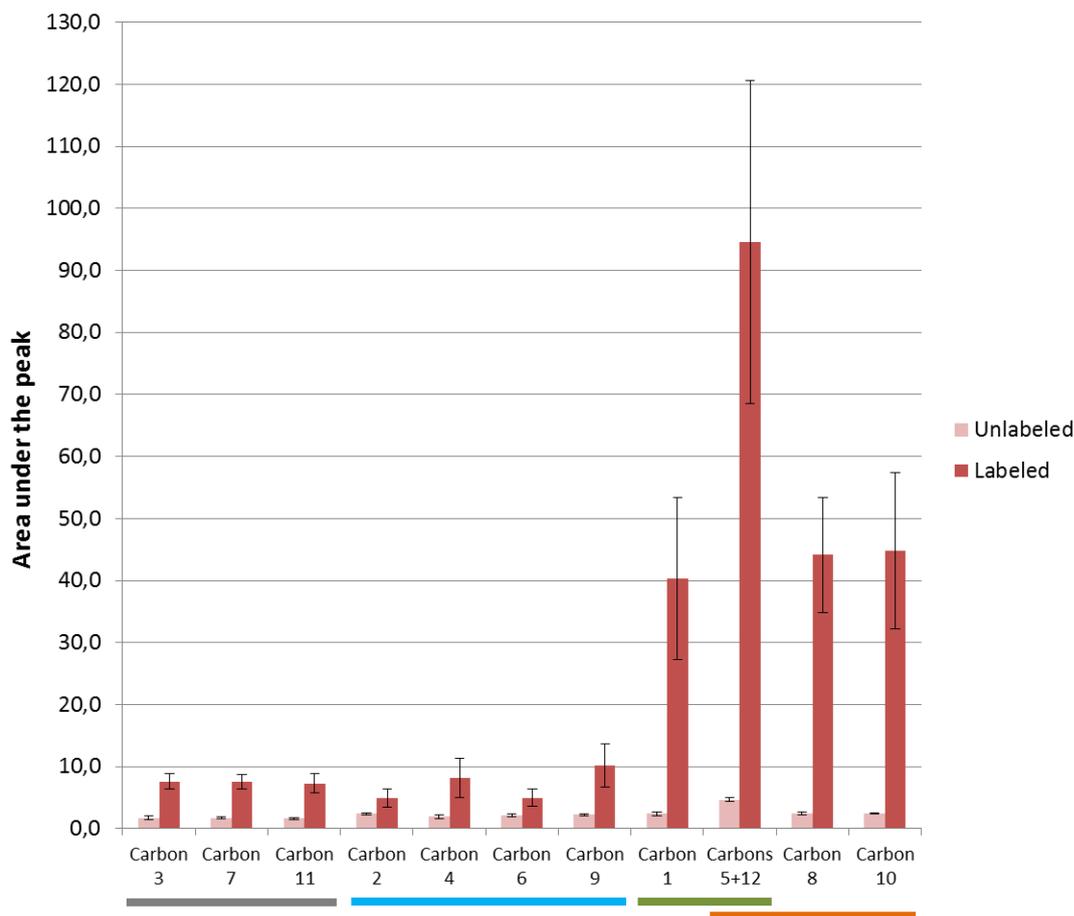


Figure 30: Linalyl acetate ^{13}C -labeling after feeding with $1\text{-}^{13}\text{C}$ -glucose.

Clary sage cut inflorescence stems were fed with a solution containing $1\text{-}^{13}\text{C}$ -glucose (**Labeled**) or standard glucose (**Unlabeled**) for 5 days. Linalyl acetate ^{13}C -labeling was analyzed by quantitative ^{13}C -NMR after hexane extraction. Bars correspond to the average of peak area measurements performed on 3 biological replicates and error bars correspond to the standard deviation. In **grey**, carbons predicted to be unlabeled; in **blue**, carbons predicted to be labeled only if the MVA pathway is involved in IPP biosynthesis; in **green**, carbons predicted to be labeled only if the MEP pathway is involved in IPP biosynthesis; in **orange**, carbons predicted to be labeled whatever the pathway involved in IPP biosynthesis.

3.2.3. Feeding with 1-¹³C-glucose together with MVA and MEP pathways inhibitors confirms sclareol and linalyl acetate metabolic origin

To further investigate the contribution of each pathway, mevinolin or fosmidomycin were added to the feeding solution together with 1-¹³C-glucose. In presence of fosmidomycin, all carbons of sclareol show a comparable very weak labeling with a labeling ratio ranging between 1.0 and 1.9 (**Figure 31a**). This result indicates that the MEP pathway is mainly responsible for the labeling observed at all carbon positions of sclareol, whether weak or strong. By contrast, labeling ratios observed in presence of mevinolin are comparable to the control (**Figure 31a**). Interestingly, the weak labeling of carbons specific of the MVA pathway is retained in presence of mevinolin, suggesting that this labeling is not due to MVA pathway activity. Similar results are obtained for linalyl acetate (**Figure 31b**). Taken together, these results exclude a possible minor contribution of the MVA pathway and show that sclareol and linalyl acetate biosynthesis fully depends on IPP and DMAPP produced by the MEP pathway. Given the fact that glucose is a central intermediate in primary metabolism, the very weak labeling observed for all sclareol carbons in presence of fosmidomycin is probably due to an aspecific labeling of all cell metabolites. Consistently, in the context of first clary sage feeding tests using 1-¹³C-glucose, an incorporation of ¹³C was detected in fructose, serine and palmitic acid (Drevensek et al., unpublished data).

3.3. Comparison of results obtained for sclareol and linalyl acetate

Higher labeling ratios (ranging between 16.7 and 20.2) are observed for labeled carbons of linalyl acetate compared to labeled carbons of sclareol (labeling ratios ranging between 7.3 and 13.0) (**Figure 31**). This observation suggests that during the feeding, more IPP and DMAPP units were targeted towards linalyl acetate biosynthesis compared to sclareol biosynthesis. Consistently, in presence of fosmidomycin, linalyl acetate carbons specific of the MEP pathway are more labeled than sclareol carbons specific of the MEP pathway (labeling ratios ranging between 2.0 and 3.0, versus labeling ratios ranging between 1.0 and 1.9). This result indicates that the MEP pathway is not completely inhibited by fosmidomycin, and that remaining MEP pathway activity produces IPP and DMAPP that is preferentially used for linalyl acetate biosynthesis compared to sclareol biosynthesis.

Although linalyl acetate biosynthesis seems to be preferred in presence of fosmidomycin according to ^{13}C -labeling data, the drop in linalyl acetate content obtained in presence of 100 μM of fosmidomycin is more important than the drop in sclareol content (93% versus 66%) (**Figure 27b**). This apparent discrepancy between the two experiments could be explained by a higher turn-over of the linalyl acetate pool. Indeed, without inhibitor treatment, linalyl acetate content of calyx organic extracts is comparable to sclareol content. However, linalyl acetate is more volatile than sclareol and probably evaporates more rapidly at the surface of the calyx. A higher carbon flux through linalyl acetate biosynthesis is probably necessary to maintain a linalyl acetate content equivalent to sclareol content in spite of higher linalyl acetate evaporation. In presence of fosmidomycin, it is likely that linalyl acetate produced before the start of the treatment evaporates whereas sclareol remains at the surface of the calyx. This would explain why only 7% of linalyl acetate content remains after 6 days of treatment, versus 34% in the case of sclareol.

3.4. The case of the acetate group of linalyl acetate

Carbons 11 and 12 of linalyl acetate correspond to the acetate moiety. In lemon mint (*Mentha aquatica* var. *citrate*, Lamiaceae), a linalool acetyltransferase activity has been detected and uses acetyl-CoA as the acetyl group donor to produce linalyl acetate from linalool (Zaks et al., 2008). It is likely that a similar activity exists in clary sage and that acetyl-CoA is used as the acetyl group donor for the production of linalyl acetate. After 1- ^{13}C -glucose processing through glycolysis and pyruvate oxidation, ^{13}C -labeling is expected to be found on carbon 2 of acetyl-CoA, which corresponds to carbon 12 of linalyl acetate (Carvalho et al., 2013; Jones, 2014). Peak 25 of the calyx extract ^{13}C -NMR spectrum integrates for 2 carbons of linalyl acetate: carbon 12 and carbon 5 (**Table 1**). Carbon 5 is expected to be labeled only if the MEP pathway contributes to the biosynthesis of IPP and DMAPP required for linalyl acetate biosynthesis. On the spectrum obtained for samples fed with 1- ^{13}C -glucose without inhibitor, the area measured under peak 25 is about twice the area measured under peaks corresponding to single labeled carbons (carbons 1, 8 and 10), indicating that both carbon 12 and carbon 5 are labeled (**Figure 30**). This result is in agreement with acetyl-CoA being the acetyl group donor and with the involvement of the MEP pathway in linalyl acetate biosynthesis.

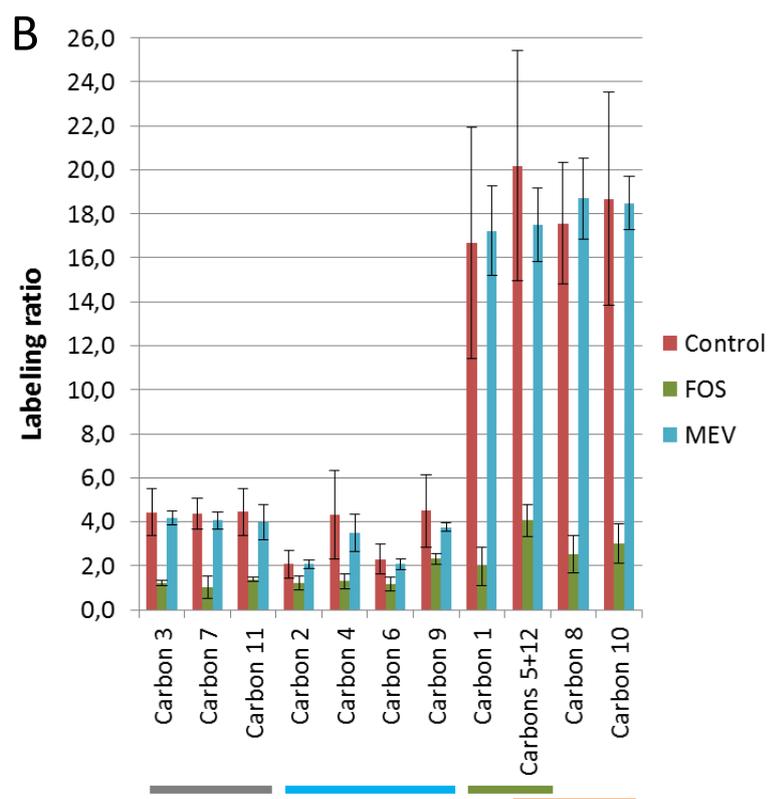
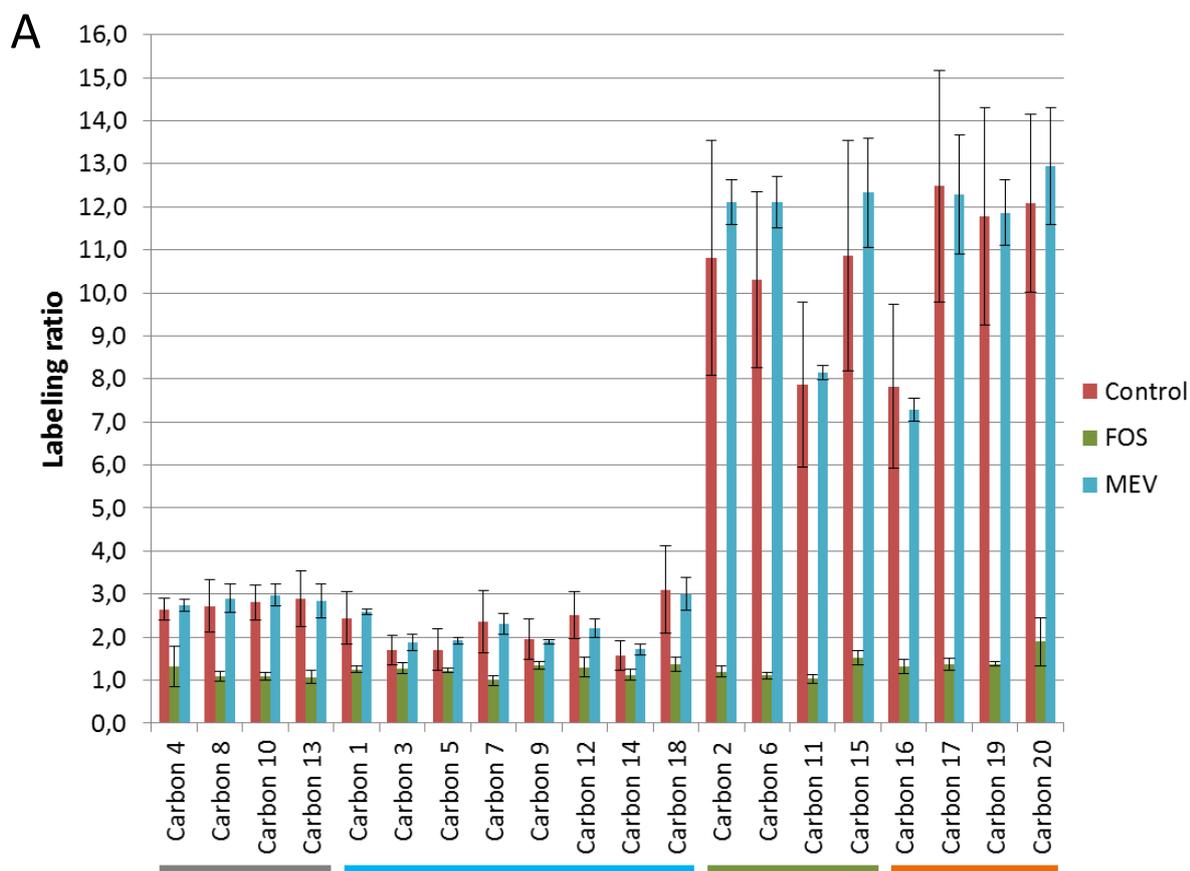


Figure 31: Sclareol and linalyl acetate ^{13}C -labeling after feeding with $1\text{-}^{13}\text{C}$ -glucose and terpene biosynthesis inhibitors.

Clary sage cut inflorescence stems were fed with a solution containing $1\text{-}^{13}\text{C}$ -glucose or standard glucose with $100\mu\text{M}$ of mevinolin (**MEV**), $100\mu\text{M}$ of fosmidomycin (**FOS**) or no inhibitor (**Control**) for 5 days. Sclareol (**A**) and linalyl acetate (**B**) ^{13}C -labeling was analyzed by quantitative ^{13}C -NMR after hexane extraction. The **labeling ratio** is calculated by dividing peak area measured for labeled samples by peak area measured for unlabeled samples. Bars correspond to the average of labeling ratio measurements performed on 3 biological replicates and error bars correspond to the standard deviation. In **grey**, carbons predicted to be unlabeled; in **blue**, carbons predicted to be labeled only if the MVA pathway is involved in IPP biosynthesis; in **green**, carbons predicted to be labeled only if the MEP pathway is involved in IPP biosynthesis; in **orange**, carbons predicted to be labeled whatever the pathway involved in IPP biosynthesis.

3.5. The mixed origin of the sesquiterpene β -caryophyllene

Sclareol and linalyl acetate are major compounds produced by clary sage calyces (**Figure 26**), but calyx organic extracts also contain minor compounds such as the sesquiterpenes germacrene D and β -caryophyllene (Schmiderer et al., 2008). The abundance of these minor compounds is too low for them to generate a signal in ^{13}C -NMR, but β -caryophyllene ^{13}C -labeling was successfully analyzed by GC-MS. ^{13}C -labeling was estimated through the analysis of isotope cluster distortion induced by the incorporation of ^{13}C atoms. The shapes of two isotope clusters of β -caryophyllene mass spectrum and three isotope clusters of linalyl acetate mass spectrum were evaluated by calculating the ratio of heavy ions peak areas over light ion peak area. As expected, feeding inflorescence stems with 1- ^{13}C -glucose leads to a distortion of β -caryophyllene and linalyl acetate isotope clusters in favor of heavy ions: the distortion ratio is multiplied by a factor 2 to 5 depending on considered cluster (**Figure 32**). This result indicates that 1- ^{13}C -glucose has been metabolized by plant cells and progressively incorporated into β -caryophyllene and linalyl acetate, although this observation alone does not provide information about which metabolic pathways have been involved in the production of each metabolite. Interestingly, terpene biosynthesis inhibitors have contrasted effects on the distortion of β -caryophyllene and linalyl acetate isotope clusters. Unlike mevinolin, fosmidomycin almost completely abolishes the distortion of linalyl acetate isotope clusters induced by 1- ^{13}C -glucose feeding (**Figure 32a**). This result is in agreement with ^{13}C -NMR data and demonstrates that linalyl acetate biosynthesis essentially depends on IPP and DMAPP production by the MEP pathway. By contrast, fosmidomycin and mevinolin both reduce the distortion of β -caryophyllene isotope clusters, suggesting that both MVA and MEP pathways contribute to the production of IPP and DMAPP required for β -caryophyllene biosynthesis (**Figure 32b**).

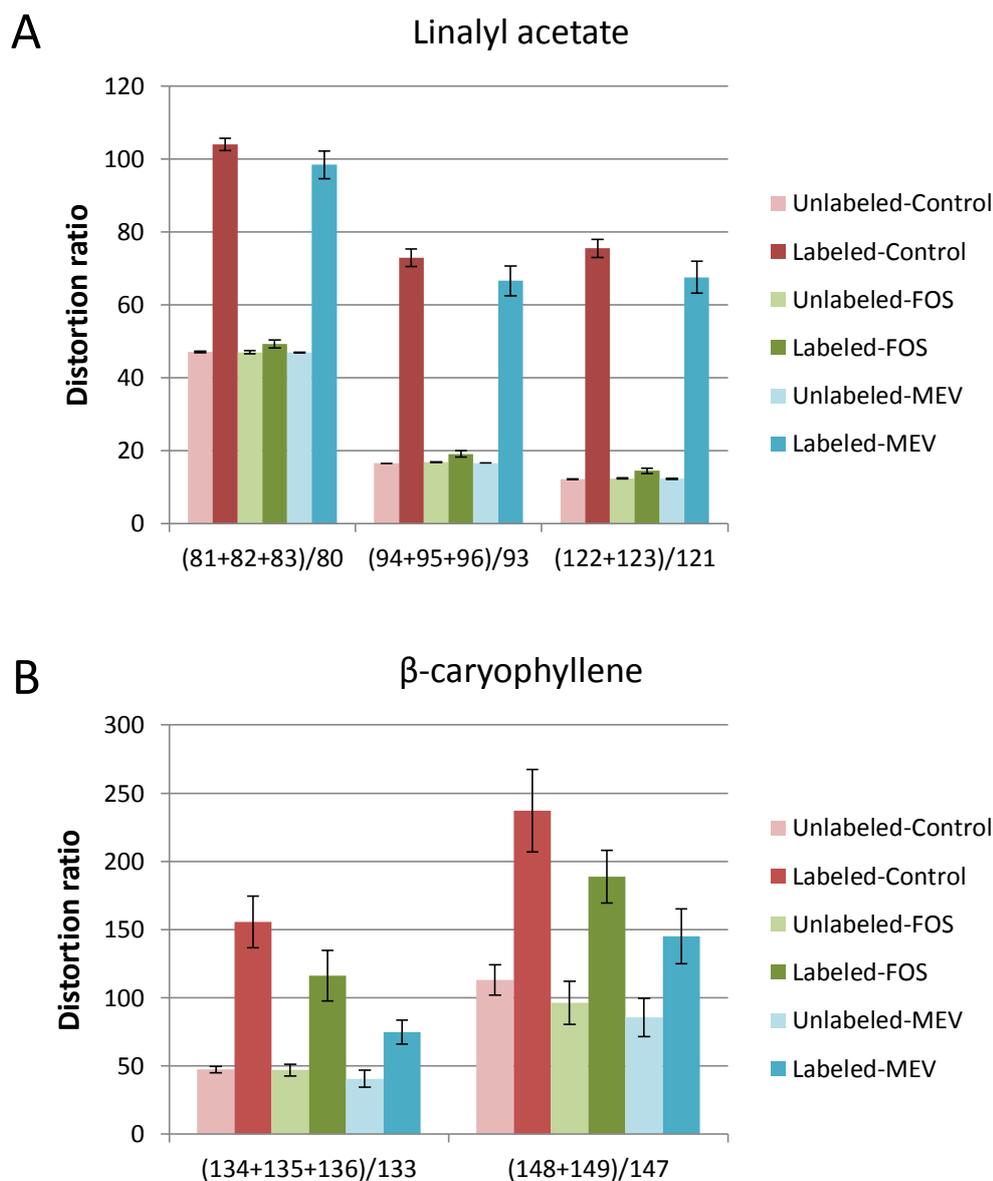


Figure 32: Estimation of linalyl acetate and β -caryophyllene ^{13}C -labeling after feeding with $1\text{-}^{13}\text{C}$ -glucose and terpenoid biosynthesis inhibitors.

Clary sage cut inflorescence stems were fed with a solution containing $1\text{-}^{13}\text{C}$ -glucose (**Labeled**) or standard glucose (**Unlabeled**) with $100\mu\text{M}$ of mevinolin (**MEV**), $100\mu\text{M}$ of fosmidomycin (**FOS**) or no inhibitor (**Control**) for 5 days. Linalyl acetate and β -caryophyllene ^{13}C -labeling was estimated by GC-MS analysis of isotope cluster distortion induced by the incorporation of ^{13}C atoms. Distortion ratio was calculated by dividing the sum of heavy ions peak areas by light ion peak area. Bars correspond to the average of distortion ratio measurements performed on 3 biological replicates and error bars correspond to standard deviation.

4. Discussion

4.1. Biosynthetic origin of mono-, sesqui- and diterpenes in clary sage and other plants

Isotope labeling strategies are powerful tools to study metabolic fluxes and have been successfully employed to characterize the metabolism of many plant species (Lipko and Swiezewska, 2016). Notably, the specific labeling patterns obtained on terpenoids after labeling with $1\text{-}^{13}\text{C}$ -glucose are useful to analyze the contribution of MVA and MEP pathways to the production of terpenoids of interest (**Figure 28**). Here, we employed this strategy to assess the biosynthetic origin of 3 terpenoids produced by clary sage calyces: the diterpene sclareol, the sesquiterpene β -caryophyllene and the monoterpene linalyl acetate.

The strong impact of a treatment with the MEP pathway inhibitor fosmidomycin already suggested a strong involvement of the MEP pathway in sclareol and linalyl acetate biosynthesis (**Figure 27**). However, one could argue that the drop in sclareol and linalyl acetate content is due to a decrease in photosynthetic activity caused by the inhibition of carotenoid synthesis by fosmidomycin (Lipko and Swiezewska, 2016). The feeding of clary sage inflorescence stems with $1\text{-}^{13}\text{C}$ -glucose together with fosmidomycin rules out this hypothesis and clearly demonstrates that sclareol and linalyl acetate both originate from the plastidial MEP pathway (**Figure 31**), like the majority of plant monoterpenes and diterpenes (Reed and Osbourn, 2018). These results are consistent with the plastidial localization of clary sage terpene synthases involved in sclareol biosynthesis (Caniard et al., 2012). Isotope labeling experiments had previously shown that linalyl acetate is MEP-derived in lemon mint (*Mentha citrata*) (Fowler et al., 1999). Comparable results were obtained for other monoterpenes and diterpenes in other species of the Lamiaceae family. In *Salvia divinorum*, the same labeling strategy based on the use of $1\text{-}^{13}\text{C}$ -glucose demonstrated that the diterpene salvinatorin A is MEP-derived (Kutrzeba et al., 2007). In spike lavender (*Lavandula latifolia*), the main monoterpenes contained in the essential oil, 1,8-cineole and camphor, both originate from the MEP pathway, as illustrated by pulse-chase experiments based on $^{13}\text{CO}_2$ labeling (Mendoza-Poudereux et al., 2017).

Sesquiterpenes are usually MVA-derived. However, the GC-MS analysis of isotope cluster distortion after labeling with $1\text{-}^{13}\text{C}$ -glucose suggests that in clary sage calyces, the sesquiterpene β -caryophyllene is of mixed origin (**Figure 32**). A similar GC-MS analysis was

previously employed to investigate the contribution of MVA and MEP pathways to the biosynthesis of mono- and sesquiterpenes in snapdragon flowers (Dudareva et al., 2005). Other plants produce sesquiterpenes with IPP stemming from both MVA and MEP pathways, like for example the artemisinin-producing plant *Artemisia annua* (Towler and Weathers, 2007). The contribution of the plastidial MEP pathway to the biosynthesis of a sesquiterpene in the cytosol suggests that the MEP pathway is highly active in glandular trichomes of clary sage calyces, although these epidermal structures have been reported to be non-photosynthetic in other Lamiaceae species like mint (Lange and Turner, 2013). The analysis of clary sage glandular trichome ultrastructure by transmission electron microscopy would provide useful information about the structure-function adaptation of secretory cells. In particular, it would help to determine whether glandular trichome plastids are functional chloroplasts like in tomato type VI capitate trichomes (Balcke et al., 2017), or non-photosynthetic leucoplasts similar to those of mint peltate trichomes (Lange and Turner, 2013). If clary sage glandular trichomes are not photosynthetic, carbon assimilates must be transported from neighboring cells to glandular trichomes to supply MEP pathway activity. A recent study demonstrated that although being photosynthetic, tomato type VI glandular trichomes rely on leaf sucrose to maintain their high metabolic activity (Balcke et al., 2017). This conclusion emerged from the comparison of ^{13}C incorporation in primary metabolites of leaves and isolated type VI glandular trichomes during $^{13}\text{CO}_2$ labeling. The same approach could be employed on clary sage to analyze the contribution of non-glandular cells sucrose to glandular trichome metabolism.

4.2. Metabolic engineering of terpene production in plants

Different strategies can be conceived in order to increase the production of a compound of interest in plant cells. A first category of approaches attempts to increase the amount of substrate available for the biosynthesis of the molecule of interest by removing bottlenecks in upstream pathways (Reed and Osbourn, 2018). Suppressing competition by other metabolic pathways using the same substrate is another way to reach the same goal (Lu et al., 2016; Reed and Osbourn, 2018). Such strategies have been developed in heterologous terpene production platforms like *Nicotiana benthamiana* (Reed and Osbourn, 2018), but significant results have also been obtained directly in species naturally producing the

compound of interest, like *Artemisia annua* for the anti-malarial compound artemisinin (Lu et al., 2016).

4.2.1. Increasing the production of prenyl intermediates

When the metabolite of interest is a diterpene, a promising metabolic engineering approach is to attempt to enlarge the pool of GGPP, the universal precursor of diterpene biosynthesis. Strategies aiming at increasing GGPP biosynthesis have been tested in different plant systems with contrasted results. In some cases, the overexpression of a GGPP-synthase alone was enough to enhance diterpene biosynthesis, for example tanshinone biosynthesis in *Salvia miltiorrhiza* hairy roots (Kai et al., 2011). By contrast, in *Nicotiana benthamiana*, the overexpression of a GGPP-synthase in addition to a cembatrien-ol synthase did not increase cembatrien-ol content compared to plants overexpressing the cembatrien-ol synthase alone, suggesting that GGPP-synthase activity is not limiting (Brückner and Tissier, 2013). In such cases, enlarging the IPP and DMAPP pool may be necessary.

The majority of plant diterpenoids being mainly MEP-derived, efforts in enlarging the IPP and DMAPP pool available for diterpenoid production have historically focused on increasing the activity of the MEP pathway rather than the MVA pathway. Enzymes catalyzing the first steps of the MEP pathway, DXS and DXR, are often reported as rate-limiting (Tholl, 2015). For this reason, enhancing their activity appears to be a good strategy to increase IPP and DMAPP production by the MEP pathway. In *Salvia miltiorrhiza* hairy roots, overexpression of *DXS* or *DXR* could enhance tanshinone content by up to 5-fold (Kai et al., 2011; Shi et al., 2014). In *Nicotiana benthamiana*, the overexpression of *DXS2* in addition to a cembatrien-ol synthase did not increase cembatrien-ol content compared to plants overexpressing the cembatrien-ol synthase alone. However, the overexpression of *DXS2* and a GGPP-synthase together with the cembatrien-ol synthase led to a 3.5-fold increase of cembatrien-ol content (Brückner and Tissier, 2013). In *Vitis vinifera*, a specific *DXS* allele with better catalytic activity is responsible for the distinct aroma of Muscat cultivars, due to the higher monoterpene content of berries (Battilana et al., 2011). The overexpression of the muscat allele of *DXS1* increased monoterpene production in microvine (Dalla Costa et al., 2018).

Enhanced MVA pathway activity can also positively impact MEP-derived terpene biosynthesis. In *Lavandula latifolia*, the overexpression of *HMGR* increased the biosynthesis

of 1,8-cineole and camphor, the major monoterpenes contained in the essential oil: camphor and 1,8-cineole contents were enhanced by 30% and 18%, respectively (Mendoza-Poudereux et al., 2017). Surprisingly, both metabolites remained produced with IPP stemming only from the MEP pathway, suggesting that the MVA pathway affects monoterpene production otherwise than by exporting prenyl intermediates to plastids (Mendoza-Poudereux et al., 2017). This crosstalk may be due to protein prenylation with MVA-derived prenyl groups, which could positively regulate the expression of monoterpene biosynthesis genes as previously evidenced in *Catharanthus roseus* cell cultures (Courdavault et al., 2005).

4.2.2. Suppressing competing pathways

Another way to increase the amount of substrate available for the biosynthesis of the compound of interest is to eliminate competition by pathways using the same substrate. In *Artemisia annua*, several metabolic pathways compete with the biosynthesis of artemisinin. For example, the biosynthesis of three other sesquiterpenes, β -caryophyllene, β -farnesene and germacrene A use the same MVA-derived substrate FPP. Downregulation of a β -caryophyllene synthase, a β -farnesene synthase or a germacrene A synthase by RNA interference increased artemisinin content by 77%, 77% and 103%, respectively (Lv et al., 2016). Reducing the flux through a competing pathway has also proved to be an efficient strategy to enhance the production of MEP-derived terpenoids. For example, in sweet potato, lutein biosynthesis competes with the biosynthesis of β -carotene for the use of the same MEP-derived substrate, lycopene. To increase β -carotene production, the enzyme catalyzing the first step of lutein biosynthesis from lycopene, the lycopene ϵ -cyclase, was downregulated through RNA interference. Lutein content was reduced to undetectable levels in sweet potato transgenic calli, while β -carotene content was approximately 21-fold higher (Kim et al., 2013). These results highlight the potential of blocking competing pathways to enhance the yield of a compound of interest.

4.3. Perspectives for clary sage metabolic engineering

4.3.1. Increasing MEP pathway activity

Since sclareol is clearly MEP-derived in clary sage calyces (**Figure 31**), enhancing MEP pathway activity may be a straightforward approach to increase sclareol content. This strategy has already been employed successfully to enhance the production of abietane diterpenes by clary sage roots (Vaccaro et al., 2014). The ectopic expression of *AtDXS* and *AtDXR* genes in hairy roots led to a 2.2-fold and 3.3-fold increase in abietane diterpene content, respectively. Some hairy root lines overexpressing *AtDXS* also grew significantly slower. However, this negative effect on growth is not due to increased abietane diterpene accumulation, because *AtDXR* overexpression led to even higher abietane diterpene levels without any adverse effect on root growth. Some lines overexpressing *AtDXR* even grew faster than the control, thereby combining higher abietane diterpene content and improved growth (Vaccaro et al., 2014).

4.3.2. Increasing MVA pathway activity

At first sight, blocking the MVA pathway appears to be a better strategy than enhancing it, because it would block sesquiterpene biosynthesis, which theoretically competes with sclareol biosynthesis for carbon assimilates. However, this strategy would probably be inefficient for two reasons. First, in clary sage calyces, sesquiterpenes are produced in very low amounts compared to sclareol (**Figure 26**), and thus divert very few carbon precursors away from the MEP pathway and sclareol biosynthesis. Secondly, MVA and MEP pathways rather cooperate than compete, due to the existence of metabolic crosstalks (Hemmerlin et al., 2012; Liao et al., 2016). Therefore, boosting MVA pathway activity may also elevate sclareol content. In *Salvia miltiorrhiza* hairy roots, *HMGR* overexpression increased the biosynthesis of the presumably MEP-derived diterpene tanshinone by up to 3-fold (Kai et al., 2011; Shi et al., 2014). Moreover, an even higher elevation of tanshinone content (by up to 5-fold) was obtained by co-overexpressing *HMGR* together with *DXR* (Shi et al., 2014), indicating that strategies based on MEP or MVA pathway enhancement can be successfully combined.

4.3.3. Blocking monoterpene biosynthesis

In addition to being both MEP-derived (**Figure 31**), the monoterpene linalyl acetate and the diterpene sclareol are mainly produced in the same cells of clary sage calyces: glandular trichome secretory cells (see **Chapter 1**). Hence, their respective biosynthesis pathways are in competition for the use of the same IPP and DMAPP pool produced by the MEP pathway in glandular trichomes plastids. Reducing or blocking the biosynthesis of linalyl acetate may therefore increase sclareol production, because IPP and DMAPP not converted to linalyl acetate would be available for sclareol biosynthesis.

In lemon mint (*Mentha aquatica* var. *citrata*, Lamiaceae), linalyl acetate is obtained after acetylation of the monoterpene linalool, which is produced in one step from GPP through the action of a linalool-synthase (Zaks et al., 2008). To our knowledge, no linalool acetyltransferase gene was identified yet, whereas several linalool-synthase genes were characterized in Lamiaceae species including lavender (Landmann et al., 2007) and sweet basil (Iijima et al., 2004). Clary sage genes homologous to these linalool-synthase genes probably encode enzymes responsible for linalool biosynthesis in calyces. The suppression of linalool-synthase activity would lead to GPP accumulation. The accumulation of its product may in turn inhibit GPP-synthase activity, thereby increasing the amount of IPP and DMAPP available for GGPP biosynthesis. However, multiple linalool-synthase isoforms might be functional in calyx glandular trichomes. In that case, the reduction of linalool-synthase activity would require the inactivation of several, if not all, linalool-synthase isoforms, otherwise remaining isoforms would probably complement the loss of function. Moreover, in absence of linalool-synthase activity, other terpene synthases may convert GPP into other monoterpenes, again diverting prenyl intermediates away from sclareol biosynthesis.

A way to block the biosynthesis of all monoterpenes would be to suppress GPP-synthase activity. In *Mentha piperita* and *Antirrhinum majus*, two species relatively closely related to *Salvia sclarea*, GPP-synthases are heteromeric, comprising a large catalytic subunit and a small regulatory subunit (Burke et al., 1999; Tholl et al., 2004). In absence of the regulatory subunit, the catalytic subunit produces GGPP instead of GPP (Burke and Croteau, 2002). If clary sage GPP-synthases are also heteromeric, inactivating regulatory subunits would at the same time repress GPP biosynthesis and increase GGPP-synthase activity, because catalytic subunits would display GGPP-synthase activity and work together with authentic GGPP-

synthases. Alternatively, inactivating catalytic subunits would also suppress GPP biosynthesis, thereby increasing the amount of IPP and DMAPP available for GGPP production by GGPP-synthases. As mentioned previously in the case of linalool-synthases, these strategies may be difficult to implement if multiple GPP-synthase isoforms are functional in calyx glandular trichomes, because the inactivation of several, if not all, GPP-synthase isoforms would be necessary. In any case, this approach requires the characterization of genes encoding clary sage GPP-synthase catalytic subunit(s) and putative regulatory subunit(s), and the identification of GPP-synthase isoform(s) active in calyx glandular trichomes. Transcriptomic data obtained from clary sage calyces are a precious resource for such analyses. Interestingly, only one GPP-synthase isoform was identified in the available clary sage calyx transcriptome, suggesting that only one GPP-synthase isoform is expressed in the calyx (Legrand et al., 2010). If further gene expression analysis on isolated glandular trichomes confirms that only one GPP-synthase isoform is expressed in calyx glandular trichomes, the inactivation of this isoform would be a promising and straightforward strategy to redirect carbon fluxes towards sclareol biosynthesis.

4.3.4. Engineering increased sclareol secretion

If sclareol production is strongly induced, existing transport mechanisms responsible for sclareol secretion may not be sufficient to export all sclareol out of glandular trichome cells. Sclareol may accumulate inside cells and impact glandular trichome productivity negatively. Indeed, sclareol has been shown to display cytotoxic activity to animal and fungal cells, but also to plant cells, as evidenced in *Arabidopsis thaliana* (Van Den Brûle et al., 2002). An efficient induction of sclareol biosynthesis may require the engineering of adapted transport systems. Today, the mechanism of sclareol secretion in clary sage remains elusive and may involve passive diffusion and/or active transport systems (Caissard et al., 2012). A transporter belonging to the Pleiotropic Drug Resistance-like (PDR-like) subfamily of ATP-Binding Cassette (ABC) transporters was shown to export exogenous sclareol out of *Arabidopsis thaliana* cells (Van Den Brûle et al., 2002). In *Nicotiana plumbaginifolia*, an ABC transporter of the same PDR-like subfamily is involved in plant defense and is more expressed in leaf glandular trichomes and flower petals (Stukkens et al., 2005). If sclareol

transport proves to be limiting, increasing the expression of PDR-like ABC transporters in clary sage may further enhance sclareol secretion by clary sage.

Chapter 3

Is sclareol biosynthesis regulated by jasmonates?

1. Introduction

1.1. Regulation of plant specialized metabolism by jasmonates

Jasmonates are phytohormones involved in the regulation of plant stress responses and other various developmental processes. For example, they induce the production of specialized metabolites upon herbivore feeding or attack by a necrotrophic pathogen (Goossens et al., 2016), but also play a role in primary root growth, flower development and leaf senescence (Huang et al., 2017). The jasmonate family of phytohormones includes jasmonic acid and its inactive and active forms, among which a volatile derivative, methyljasmonate (MeJA).

Endogenous jasmonic acid production by plant cells is induced by diverse biotic and abiotic stresses and developmental signals (Wasternack and Strnad, 2019). Jasmonic acid can undergo various chemical modifications including conjugation with amino acids and methylation. Jasmonic acid conjugation with isoleucine leads to the formation of the active form JA-Ile. JA-Ile is transported to the nucleus and specifically interacts with the SCF^{CO1}-JAZ co-receptor complex (Goossens et al., 2016). JASMONATE ZIM DOMAIN (JAZ) proteins interact with positively-acting transcription factors to repress their action. SCF^{CO1} is an E3-ubiquitin ligase complex which is part of the ubiquitin-proteasome degradation machinery and specifically targets JAZ proteins. Upon binding of JA-Ile, SCF^{CO1} mediates the degradation of JAZ proteins by the proteasome, thus removing the repression exerted by JAZ proteins. Positively-acting transcription factors are released and bind to JA responsive elements located in JA response gene promoters, thus triggering the expression of these genes (Goossens et al., 2016).

JA response genes include transcription factors which control the JA-induced accumulation of specialized metabolites. These regulators often belong to the AP2/ERF, bHLH, MYB and WRKY transcription factor families (Zhou and Memelink, 2016). For example, in *Catharanthus roseus*, monoterpene indole alkaloid production is controlled by AP2/ERF

transcription factors of the ORCA family and by the bHLH transcription factors BIS1 and BIS2, which are all JA-inducible (Wasternack and Strnad, 2019). In *Artemisia annua*, the biosynthesis of the sesquiterpene artemisinin is activated in response to JA through the action of the bHLH transcription factor MYC2 and also the WRKY transcription factor GSW1. *GSW1* is expressed specifically in glandular trichomes, which are the site of artemisinin production in *Artemisia annua* (Shen et al., 2016; Chen et al., 2017b).

JA-triggered accumulation of specialized metabolites is primarily achieved through the induction of biosynthetic gene expression. In the case of JA-inducible terpenoids, the expression of downstream pathway genes like terpene synthase genes is enhanced, as well as MVA or MEP pathway genes. For example, in *Artemisia annua*, JA increases the expression of a DXS gene, an FPP-synthase gene and a cytochrome P450 gene, which respectively encode the first enzyme of the MEP pathway, the enzyme producing FPP (the linear precursor of artemisinin) and an enzyme involved in the conversion of FPP to artemisinin (Maes et al., 2011). When specialized metabolites are produced in glandular trichomes, JA can also impact the yield of a compound of interest through the induction of glandular trichome development, as demonstrated in *Artemisia annua* (Maes et al., 2011) and tomato (Chen et al., 2018).

1.2. How to use jasmonate signaling to enhance the production of compounds of interest

Many terpenoids exploited in industry are specialized metabolites produced by plants in response to biotic and abiotic stresses. Strategies based on the induction of plant stress responses have therefore been designed to increase the biosynthesis of terpenoids of interest in plant systems (Lu et al., 2016). Genes involved in the same biosynthetic pathway are often co-regulated. Hence, manipulating biosynthesis pathway regulation offers the possibility to upregulate all biosynthetic genes together in appropriate proportions, a goal which is more difficult to reach when targeting biosynthetic enzymes one by one (Wang et al., 2016).

Methyljasmonate (MeJA) is a volatile derivative of jasmonate which activates JA signaling after conversion to JA-Ile in plant cells. Commercial MeJA is often used as an elicitor to enhance the biotechnological production of terpenes of interest in plant systems (Lu et al.,

2016). For example, MeJA has been successfully employed at industrial scale to increase the production of the anti-cancer molecule paclitaxel (Taxol®) in *Taxus cuspidata* cell cultures. The combination of MeJA elicitation and cell culture optimization methods led to high paclitaxel titers of up to 900 mg/L, in spite of the negative impact of MeJA on cell growth (Patil et al., 2014).

Other strategies aiming at increasing terpenoid production employ genetic modifications to enhance JA signaling. In tomato, constitutive activation of JA signaling was achieved through the overexpression of the precursor protein of systemin, a peptide hormone known to activate JA signaling in tomato, and led to the accumulation of the phenylpropanoid-polyamine conjugate caffeoyl putrescine (Chen et al., 2006). Different components of JA signaling have also been targeted to enhance the production of JA-inducible terpenoids. In *Salvia miltiorrhiza* hairy roots, the production of the diterpene tanshinone was increased by up to 2.5-fold through downregulation of JAZ repressors (Shi et al., 2016). In *Catharanthus roseus* cells, overexpression of the JA-inducible transcription factors ORCA4, BIS1 or BIS2 led to an increase in monoterpene indole alkaloid content (Wasternack and Strnad, 2019). Artemisinin content of *Artemisia annua* plants was enhanced by up to 50% through the overexpression of either of the AP2/ERF transcription factors AaERF1, AaERF2 or AaORA1 (Yu et al., 2011; Lu et al., 2013), or through the overexpression the bHLH transcription factor AaMYC2 (Shen et al., 2016). Moreover, the overexpression of the WRKY transcription factor AaWRKY1 could increase artemisinin content by 80% (Han et al., 2014).

1.3. Regulation of clary sage specialized metabolism by jasmonates: state of the art

Several studies demonstrate a positive impact of JA signaling on specialized metabolite content in different clary sage aerial and underground organs. Increased levels of metabolites belonging to the phenylpropanoid family (caffeic acid, rosmarinic acid, luteolin and apigenin) were detected in clary sage leaves 3 days after spraying with MeJA at a concentration of 7.5 mM (Hao et al., 2015). A limited, but significant increase in rosmarinic acid content was also visible 6 days after spraying with MeJA at a concentration of 0.75 mM. Besides, clary sage hairy roots treated with MeJA produced higher levels of abietane diterpenes (aethiopinone, 1-oxoaethiopinone, salvipisone and ferruginol) (Kuzma et al.,

2009a; Vaccaro et al., 2017). The strongest induction was observed for aethiopinone content, which could be enhanced by up to 25-fold (Vaccaro et al., 2017). The increase in abietane diterpene content was strongly correlated to enhanced transcription of biosynthetic genes, including MEP pathway genes (especially *DXS2* and *DXR*) and GGPP-synthase genes (Vaccaro et al., 2017). Interestingly, MeJA treatment could elevate total abietane diterpene content by more than 20-fold, i.e. to much higher levels than *AtDXS* and *AtDXR* ectopic expression in hairy roots, which only led to a 2 to 3-fold increase in abietane diterpene content (Vaccaro et al., 2014, 2017).

Abietane diterpene content in clary sage hairy roots was recently increased by ectopic expression of the JA-responsive transcription factors *AtMYC2*, *AtWRKY18* or *AtWRKY40* (Alfieri et al., 2018). Ectopic expression of each of these transcription factors enhanced the transcription of MEP pathway genes, suggesting that JA stimulates abietane diterpene production partially through increasing substrate supply (Alfieri et al., 2018). Best results were obtained through the ectopic expression of *AtMYC2*, which increased total abietane diterpene content by more than 5-fold. This lower increase compared to results obtained by direct MeJA treatment could be explained by the fact that MeJA enhances abietane diterpene production through the upregulation of more than one transcription factor (Alfieri et al., 2018).

Taken together, results obtained on clary sage hairy roots highlight the potential of JA signaling manipulation to engineer higher terpenoid production in clary sage. However, the impact of JA signaling on terpenoid production in clary sage aerial organs, including calyces, has not been assessed so far. Here, we treated clary sage inflorescences with MeJA in order to determine whether sclareol biosynthesis in calyces is regulated by JA.

2. Materials and methods

2.1. Calyx development stages

Early developmental stages were defined according to the size of the bud: stage 1 corresponds to a small bud (less than 3 mm long) and stage 2 corresponds to a large bud (more than 4 mm long). Later developmental stages were defined according to the appearance of the corolla and the calyx. At stage 3, calyx has opened but the corolla is still closed; at stage 4, calyx and corolla have both opened; at stage 5, corolla is wilted or has fallen; and at stage 6, the calyx has completely dried (**Figure 33a**).

2.2. MeJA treatment by feeding

Methyljasmonate (MeJA) was purchased from Sigma-Aldrich and stock solutions were prepared in ethanol. Nutritive solution containing 0 μ M, 20 μ M, 100 μ M, 500 μ M, 1 mM, 3 mM, 5 mM or 7.5 mM of MeJA was prepared in 5 mL vials (Murashige and Skoog basal salt mixture (Sigma-Aldrich); sucrose 2% m/v (Euromedex); PPMTM 0.1% v/v (Plant Cell Technology); ethanol 0.4% v/v). Clary sage inflorescence axillary stems were cut between the 3rd and the 4th node (about 10 cm from the apex). Each stem was immediately placed in a vial in a way that its extremity was touching the bottom of the vial and dipped in nutritive solution. Treatment was performed by leaving inflorescence stems on the bench for 2 or 6 days. At the end of the treatment, one mature calyx (stage 4 or 5) was sampled from each inflorescence stem (at the 2nd node, starting from the base), immediately frozen in liquid nitrogen and stored at -80 °C before sclareol and linalyl acetate quantification by GC-MS.

2.3. MeJA treatment by spraying

Spraying solution containing 3 mM of MeJA was prepared in water. Ethanol was added to the control solution to reach the same concentration as in the solution containing MeJA (0.4% v/v). All aerial parts of clary sage plants were sprayed with control or MeJA solution, including rosette leaves, calyces and inflorescence stems. 75 mL of spraying solution were sprayed on one plant in order to saturate the surface of all aerial organs with solution. Plants were placed in the shade during the evaporation of the spraying solution. After 2 days,

mature calyces (stage 4 or 5) were sampled from each plant, immediately frozen in liquid nitrogen and stored at -80 °C before sclareol and linalyl acetate quantification by GC-MS.

2.4. Sclareol and linalyl acetate quantification by GC-MS

Clary sage calyces were immersed in hexane without grinding to extract metabolites present at the surface. 1.5 mL of solvent was used to extract metabolites from 5-6 calyces at stage 1, 2 calyces at stage 2 or 1 calyx at stage 3, 4, 5 or 6. The extraction was performed without removing the developing corolla from the bud at stages 1 and 2 and without removing achenes at stages 5 and 6. 10-undecen-1-ol was employed as internal standard and was added in the solvent before extraction at a concentration of 0.75 mM. After 2 h of agitation, 1 mL of extract was centrifuged at 13,000 rpm for 10 min to eliminate potential dust or debris. Samples were directly analyzed by gas chromatography coupled to mass spectrometry (GC-MS) using a 7890B/5977A GC-MS system from Agilent and according to a protocol adapted from Laville et al. (Laville et al., 2012). The system was equipped with a 10 m guard column and an Rxi-5Sil MS column (length, 30 m; inner diameter, 0.25 mm; film thickness, 0.25 µm) (Restek, Bellefonte, PA, USA). 1 µl of sample was injected with a split ratio of 30:1. Oven temperature was set to 110 °C for 2 min, then increased to 270 °C at a rate of 10 °C/min, and finally set to 270 °C for 2 min. Other temperatures were set as follows: injector, 270 °C; transfer line, 270 °C; source, 230 °C; quadrupole, 150 °C. The carrier gas was helium at a constant flow of 1 mL/min. The quadrupole mass spectrometer was switched on after a solvent delay of 5 min and was programmed to scan from 35 u to 350 u. Data analysis was carried out using the MassHunter Quantitative Analysis software from Agilent. Sclareol and linalyl acetate were identified by comparison with analytical standards purchased from Sigma-Aldrich. Absolute quantification of sclareol and linalyl acetate was performed using a calibration curve prepared with analytical standards. Calyx dry weight was measured after lyophilization of dry material remaining after extraction. For dry weight measurements, achenes were removed from stage 6 calyces, but not from stage 5 calyces (because of their very small size at this stage, their mass was considered negligible).

3. Results

3.1. Regulation of sclareol biosynthesis during calyx development

Calyces have been shown to be the organs of clary sage where sclareol accumulates at highest levels (Caissard et al., 2012), but the dynamic of sclareol production during calyx development has not been investigated yet. To explore the temporal dimension of sclareol accumulation in the calyx, sclareol content was quantified by GC-MS at different stages of calyx development. Linalyl acetate, another major metabolite produced by the calyx, was quantified at the same time (**Figure 33**).

Sclareol and linalyl acetate production starts early during calyx development as these metabolites are already detectable at stage 1 (**Figure 33c**), where they already represent respectively 7.5% and 4.5% of calyx dry mass (**Figure 33d**). Metabolite content then gradually increases until reaching its maximum at stage 5, where sclareol and linalyl acetate are present in comparable amount (**Figure 33c**), each of them accounting for about 6% of calyx dry mass (**Figure 33d**). A decrease in metabolite content is observed at stage 6 (**Figure 33c**) whereas calyx dry weight at stage 6 is higher than at stage 5 (**Figure 33b**). The decrease in linalyl acetate content can be explained by an evaporation of this volatile molecule. Given its high boiling point (398 °C), sclareol is unlikely to have evaporated, but sclareol crystals formed during calyx drying may have fallen because of the wind. Consequently, stage 6 was not considered to be the best stage to measure metabolite content and stage 5 was used instead.

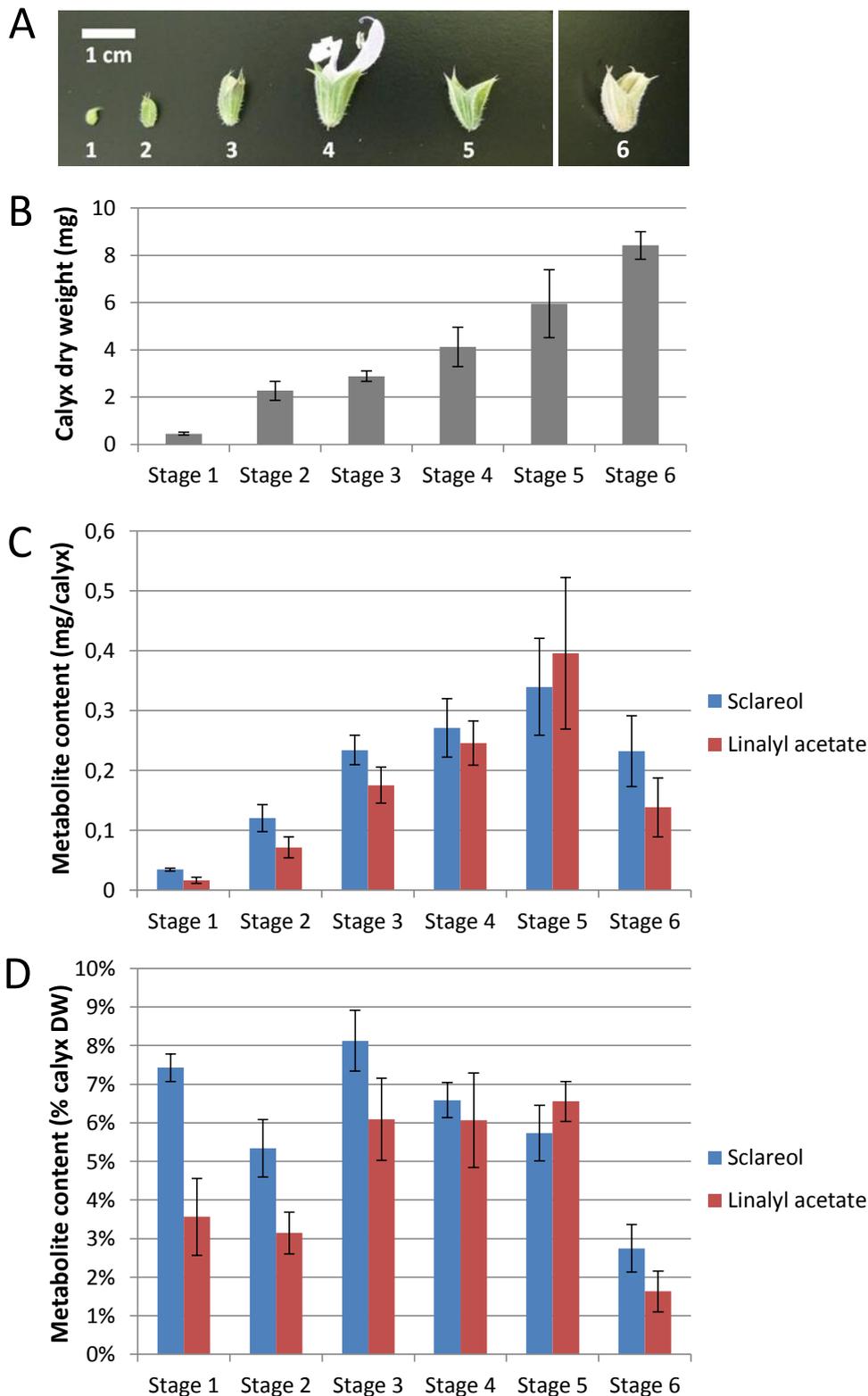


Figure 33: Sclareol and linalyl acetate accumulation during calyx development.

Sclareol and linalyl acetate content at each developmental stage of the calyx (A) were quantified by GC-MS after hexane extraction and are expressed in mg/calyx (C) or as a percentage of calyx dry weight (D). Dry weight was measured after hexane extraction (B). Hexane extraction was performed without removing the developing corolla from the bud at stages 1 and 2 and without removing achenes at stages 5 and 6. To measure dry weight, achenes were removed from stage 6 calyces, but not from stage 5 calyces (because of their very small size at this stage, their mass was considered negligible).

3.2. MeJA treatment on flowers by feeding

In order to analyze the impact of MeJA on sclareol and linalyl acetate production by calyces, clary sage cut inflorescence stems were treated with MeJA by feeding. A study performed on *Achyranthes bidentata* has shown that MeJA can be transported to distal leaves through stem vasculature. Moreover, an increased emission of volatile organic compounds was detected in distal leaves of *Achyranthes bidentata* cut stems fed with a solution containing MeJA (Tamogami et al., 2012). Thus, we could reasonably expect that in our experiment, MeJA provided in the feeding solution would be transported to clary sage calyces through the inflorescence stem vasculature, and that it may impact calyx metabolism. MeJA concentration in the feeding solution was 7.5 mM, the same concentration as that used for leaf spraying by Hao et al. (Hao et al., 2015). Inflorescence stems were fed with MeJA during 6 days in order to let enough time for the youngest calyces to fully develop from bud stage to mature stage.

After 6 days, cut inflorescence stems fed with 7.5 mM of MeJA displayed a brown color and were wilted, whereas control stems were still green (**Figure 34**). During other feeding experiments with solutions containing glucose or terpene inhibitors, cut inflorescence stem length visibly increased, indicating that internode elongation is to some extent maintained during the feeding. By contrast, after 6 days of feeding, stems fed with MeJA were clearly shorter than control stems, suggesting that the presence of MeJA in the feeding solution led to an interruption of stem growth during the feeding (**Figure 34**). The browning of organs (stem, bracts and calyces) located at the apex supports our initial assumption that MeJA is able to reach calyces and to induce a response in calyx cells after transport through the vasculature (**Figure 34**, white arrows). However, some calyces that were already mature at the beginning of the feeding remained green, and some calyces that were still at bud stage at the beginning of the feeding did not get brown (**Figure 34**, black arrows), underlying the possibility that all calyces may not receive the same amount of MeJA. In general, all calyces located at the second node (starting from the base) browned, except very young ones. Consequently, in further feeding experiments, sampling was performed among mature calyces located at the second node.

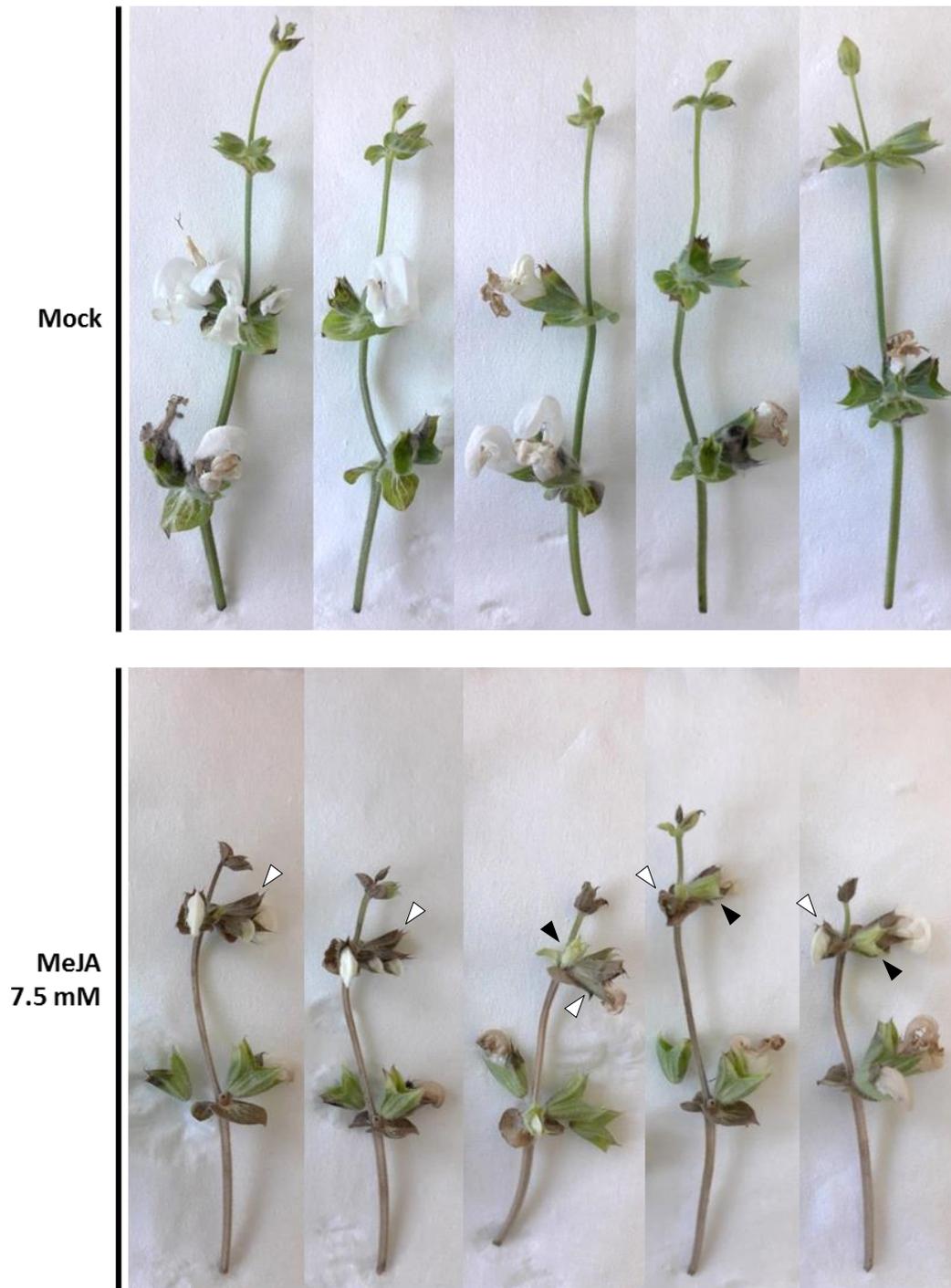


Figure 34: Impact of MeJA treatment on cut inflorescence stem growth and fitness. Clary sage cut inflorescence stems were fed with a solution containing 7.5mM of MeJA for 6 days. **White arrows** indicate brown calyces and **black arrows** indicate green calyces.

Protocol optimization tests were carried out in order to find milder treatment conditions that could potentially impact calyx metabolism without impairing inflorescence stem growth and fitness. Browning and wilting of inflorescence stems is still observed after 6 days of feeding with 3 or 5 mM of MeJA, whereas inflorescence stems fed with only 1 mM of MeJA for 6 days remain green and turgid like control stems. Therefore, 2 independent 6-day-feeding experiments with MeJA concentrations lower than 1 mM were conducted. Five different MeJA concentrations were tested in order to detect a potential dose-dependent effect of MeJA. Sclareol and linalyl acetate content was quantified by GC-MS and the 2 experiments yielded the same results (results of the second experiment are shown in **Figure 35a**): no impact of MeJA was detected on sclareol and linalyl acetate content. Concentrations tested may be too low to trigger any modification of calyx metabolism.

Further optimization tests were conducted, with higher MeJA concentrations while reducing the length of the treatment. Inflorescence stems fed with 5 mM of MeJA were completely brown and wilted in less than 48h, whereas browning and wilting started only after 3 days with 3 mM of MeJA. Therefore, a 2-day-feeding experiment with 3 mM of MeJA was performed. Again, no impact on sclareol and linalyl acetate content was detected by GC-MS analysis (**Figure 35b**).

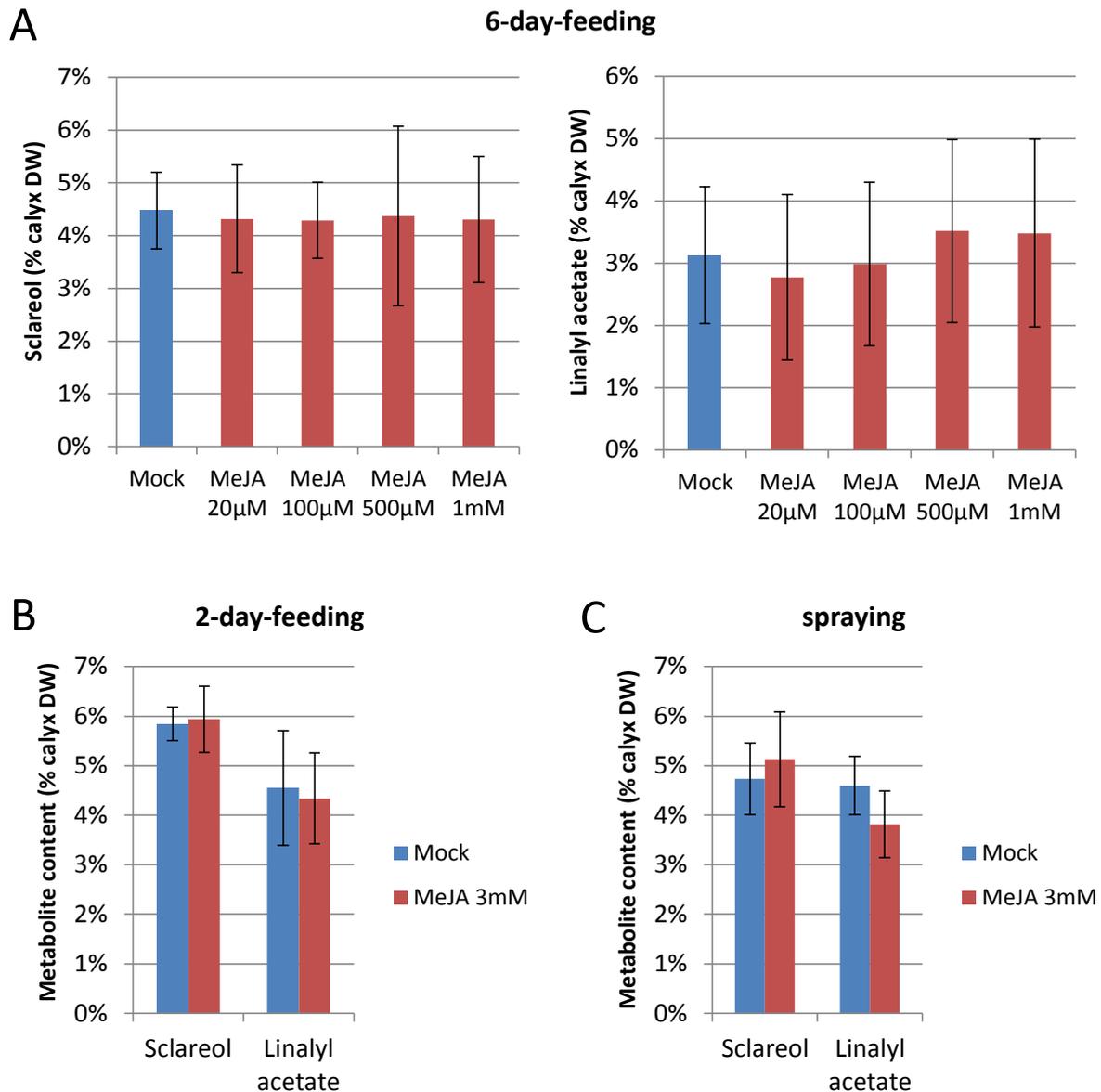


Figure 35: Impact of MeJA treatment on sclareol and linalyl acetate content.

A: Clary sage cut inflorescence stems were fed with a solution containing MeJA for 6 days. Bars represent the average of measurements performed on 4 biological replicates and error bars correspond to standard deviation.

B: Clary sage cut inflorescence stems were fed with a solution containing 3 mM of MeJA for 2 days. Bars represent the average of measurements performed on 10 calyces (5 stems, 2 calyces per stem) and error bars correspond to standard deviation.

C: Clary sage whole plants were sprayed once with a solution containing 3 mM of MeJA and calyces were sampled 2 days after treatment. Bars represent the average of measurements performed on 4 calyces (MeJA) or 8 calyces (Mock) and error bars correspond to standard deviation. Sclareol and linalyl acetate were quantified by GC-MS after hexane extraction.

3.3. MeJA treatment on flowers by spraying

A treatment protocol closer to that used by Hao et al. on leaves (Hao et al., 2015) was tested on the inflorescence. Inflorescence stems were sprayed with a solution containing MeJA without separating the stems from the rest of the plant. In order to avoid any impact on calyx growth and fitness, a concentration of 3 mM was used (instead of 7.5 mM in the work of Hao et al.) and calyces were sampled 48h after spraying (instead of 3 days in the work of Hao et al.) for GC-MS analysis. As expected, no browning and wilting was observed on floral organs. Again, no modification of sclareol and linalyl acetate content was detected with this protocol (**Figure 35c**).

To interpret the results of this last experiment and of the 2-day-feeding, it is important to note that all GC-MS analyses were performed on calyces at mature stage (stage 4 or 5). At the beginning of 2-day experiments (spraying or feeding), calyces intended to be sampled were already almost at mature stage (late stage 3 or stage 4). At this stage, in normal conditions, sclareol content has almost reached its maximum (**Figure 33c**). Since sclareol accumulates at the surface of the calyx, a modification of sclareol production rate between late stage 3 and stage 5 may have a limited impact on final sclareol content.

4. Discussion

4.1. MeJA apparently does not impact sclareol production, but further protocol optimization is needed

The phytohormone MeJA has already been shown to induce abietane diterpene biosynthesis in clary sage hairy roots (Kuźma et al., 2009a; Vaccaro et al., 2017; Alfieri et al., 2018) and phenylpropanoid production in clary sage leaves (Hao et al., 2015). For these reasons, we hypothesized that MeJA may increase terpenoid production in clary sage calyces, including sclareol biosynthesis. The treatment protocols that we employed to test this hypothesis do not highlight any clear impact of MeJA on sclareol and linalyl acetate content in clary sage calyces (**Figure 35**). Sclareol and linalyl acetate biosynthesis may be insensitive to jasmonate signaling. Alternatively, sclareol and linalyl acetate biosynthesis might be controlled by jasmonates, but the experimental design we used may not be appropriate to demonstrate the role of these phytohormones. Further optimization of treatment and sampling protocols may be necessary to highlight an impact of MeJA on sclareol and linalyl acetate production.

We tested two types of treatment protocol: cut inflorescence stem feeding with nutritive solution containing MeJA, and whole plant spraying with MeJA dissolved in water. Whole plant spraying is closer to physiological conditions compared to the stress caused by stem cutting. However, it is difficult to spray a reproducible amount of MeJA solution on each part of the plant, while inflorescence stem feeding requires less plant material and is easily reproduced. Unfortunately, feeding with MeJA had a negative impact on cut inflorescence stem growth and fitness (**Figure 34**). This result is not surprising, given the well reported negative effect of jasmonate signaling on plant growth and its ability to induce leaf senescence (Huang et al., 2017). Notably, MeJA inhibits plant cell division by inducing cell cycle arrest at G2 phase in *Arabidopsis thaliana* cell cultures (Pauwels et al., 2008). The impact of MeJA on cut inflorescence stem growth and fitness was more or less marked depending on the concentration of MeJA in the nutritive solution and the duration of the treatment. This observation is consistent with studies conducted on clary sage hairy roots: treatment with MeJA for 7 days had no visible impact on hairy root fitness, while a longer treatment of 28 days severely impaired growth (Vaccaro et al., 2017).

Due to the strong effect of MeJA feeding on clary sage cut inflorescence stem growth and fitness, a milder treatment protocol based on whole plant spraying may be preferable to test the impact of MeJA on calyx specialized metabolism. This type of protocol was successfully employed on tomato to demonstrate the positive effect of jasmonate signaling on type VI glandular trichome density on leaves (Boughton et al., 2005; Chen et al., 2018). Glandular trichome density was significantly increased 7 days after spraying with a solution of MeJA at a concentration of 7.5 mM. Here, we treated clary sage plants with a solution containing 3 mM of MeJA and collected calyces two days after spraying. The concentration we used may be too low and two days between the start of the treatment and sampling may be too short for glandular trichome development to be induced. Repeated applications of MeJA at higher concentration and a longer period between the first treatment and sampling may be necessary to impact calyx specialized metabolism and glandular trichome initiation.

A recent study on tomato highlighted a differential impact of JA on defense chemical production in leaves, depending on the developmental stage of the leaf: the increase in defense chemical content was higher in developing leaves compared to fully developed leaves (Chen et al., 2018). A more detailed analysis showed that in developing leaves, the increase in defense chemical content was mainly explained by a higher type VI glandular

trichome density, whereas it was essentially due to a higher metabolic activity of type VI glandular trichomes in fully developed leaves. Taken together, these results indicate that the impact of MeJA may be different depending on the developmental stage of the considered organ. In our study, we analyzed the impact of MeJA treatment on mature calyces which were already almost at mature stage at the beginning of the treatment. In further experiments, it may be interesting to focus the analysis on young calyces (stage 2 to 3) instead, because the possible effect of MeJA may be more marked than in mature calyces.

At the start of the MeJA treatment, a certain amount of sclareol is already present at the surface of younger calyces as well as of mature ones (**Figure 33**). Variations in this initial quantity of sclareol may mask a minor, but significant impact of MeJA on the amount of sclareol produced after the start of the treatment. To distinguish newly produced sclareol from sclareol initially present on the calyx, MeJA treatment may be combined with stable isotope labeling. For example, clary sage plants could be sprayed with MeJA and then exposed to $^{13}\text{CO}_2$ in a similar way as the protocol previously used on lavender plants (Mendoza-Poudereux et al., 2017). MeJA could also be applied in gaseous form, as previously performed on *Arabidopsis thaliana* (Thomma et al., 2000).

4.2. Other potential ways to manipulate sclareol biosynthesis regulation

Besides jasmonates, other phytohormones may play a role in sclareol biosynthesis regulation. The phytohormone abscisic acid (ABA), which plays diverse roles in plant development and adaptation to environment, has been shown to induce terpenoid production in several plants. For example, it stimulates lycopene and β -carotene accumulation in tomato fruits and increases artemisinin production in *Artemisia annua* (Lu et al., 2016). In the latter, ABA signaling has already been successfully engineered in order to enhance artemisinin production. The expression of the transcription factor AabZIP1 was shown to be induced upon ABA treatment, and its overexpression led to a 2.5-fold increase in artemisinin content (Zhang et al., 2015b). The impact of phytohormones like ABA on sclareol production by clary sage calyces could be tested by feeding or spraying. If sclareol content is increased by ABA, homologs of *AabZIP1* identified in clary sage available

transcriptomes (Legrand et al., 2010; Hao et al., 2015) would be good candidates for an involvement in sclareol biosynthesis regulation.

Many regulators of plant specialized metabolism have been identified among transcription factors involved in plant stress response, but plant development regulators also participate in the spatiotemporal control of specialized metabolites biosynthesis (Colinas and Goossens, 2018). In *Arabidopsis thaliana* and rice, transcription factors of the YABBY family play diverse roles in plant development, but do not appear to regulate specialized metabolism (Wang et al., 2016). However, in spearmint (*Mentha spicata*), downregulation or upregulation of *YABBY5* respectively increased or decreased the production of essential oil monoterpenes, indicating that *YABBY5* is a negative regulator of terpenoid biosynthesis in this plant. Interestingly, the ectopic expression of *MsYABBY5* in another Lamiaceae (*Ocimum basilicum*), and in a Solanaceae (*Nicotiana sylvestris*) also decreased terpenoid biosynthesis, suggesting that the function of *YABBY5* may be conserved in Lamiaceae and Solanaceae (Wang et al., 2016). Initially, the authors focused their study on *YABBY5* because it is preferentially expressed in spearmint peltate glandular trichomes, which are responsible for the production of essential oil monoterpenes. Data available on genes expressed in glandular trichomes appear to be a precious resource for the identification of novel regulators of the biosynthesis of specialized metabolites produced by glandular trichomes. If a clary sage homolog of *YABBY5* is more expressed in isolated glandular trichomes than in other calyx cells, it would be promising to decrease its expression and analyze the effect of this genetic modification on sclareol content.

Chapter 4

Clary sage natural diversity

1. Introduction

1.1. The use of natural genetic diversity for breeding

During the past century, agricultural practices caused a dramatic erosion of biodiversity in major crop plants like wheat or rice (Kilian and Graner, 2012). The concomitant need for genetic improvement to enhance crop performance prompted the search for new genetic resources. Natural populations of wild species closely related to domesticated crop plants carry distinct alleles which can be introgressed into elite varieties. Huge seed collections of wild relatives of crop plants were established and rapid progress in molecular biology provided effective tools to tap into these genetic resources (Kilian and Graner, 2012). Notably, it enabled the development of marker-assisted selection, whereby molecular markers are used for the indirect selection of traits of interest in crops (He et al., 2014). Once markers have been associated with a trait of interest, breeding can be performed through selecting plants possessing these markers, instead of phenotyping the plants and selecting the ones presenting the desired phenotype. In recent years, next-generation sequencing technologies further increased the potential of genetic resources to enhance plant breeding programs (He et al., 2014). For instance, Genome Wide Association Studies (GWAS) allow the identification of genetic loci involved in the quantitative variation of a trait of interest among a collection of populations. This approach is based on the use of a large set of genetic markers easily generated by high-throughput sequencing (He et al., 2014). Deeper analysis of identified genetic loci can eventually lead to the characterization of genes regulating the variation of the trait of interest, thereby improving our understanding of plant development and physiology.

The raising global demand for sclareol currently encourages breeding initiatives aiming at enhancing clary sage productivity. In addition to sclareol content, flowering time is another important trait to take into account. Indeed, since sclareol is mainly produced in floral organs (Caissard et al., 2012), the productivity of clary sage plants flowering late is potentially more impacted by summer drought. Early flowering is therefore a desirable trait.

The same idea has previously been suggested for the breeding of chamomile, a medicinal plant which also produces valuable terpenoids in flowers (Otto et al., 2017). Wild *Salvia sclarea* populations found around the Mediterranean Sea represent a pool of potentially new alleles which could be introgressed into cultivated varieties to improve sclareol content and flowering time. However, the characterization of such alleles requires the identification of genetic loci involved in the quantitative variation of these traits of interest. If natural populations displaying distinct sclareol levels were identified, genetic loci impacting sclareol yield could be characterized using state-of-the-art technologies and approaches like GWAS. Detailed study of these loci would be highly valuable to unravel regulatory mechanisms controlling sclareol production. As a first step towards this goal, we analyze here the sclareol content and flowering time of several wild populations of *Salvia sclarea* sampled in the region of Makarska in South Croatia (**Figure 36**).

1.2. Impact of environmental conditions on terpene production and genetic diversity

Specialized metabolite levels observed in plants result from the combination of physiological variations, environmental conditions and genetic factors (Figueiredo et al., 2008; Moore et al., 2014). Environmental conditions impacting metabolite content include abiotic factors, such as temperature and soil composition, and biotic factors, like the presence of specific pathogens. Altitude potentially influences plant specialized metabolism as a combination of abiotic and biotic factors. Indeed, temperature, hygrometry, wind speed and water availability change with altitude, promoting the development of distinct microbial and insect communities (Mahzooni-Kachapi et al., 2014; Foutami et al., 2018). Higher elevation (3200 m or 2400 m instead of 1600 m above sea level) was shown to impact negatively the yield and quality of essential oil extracted from *Stachys lavandulifolia*, a species belonging to the Lamiaceae family (Mahzooni-Kachapi et al., 2014).

Plants harboring the same genome often have distinct chemical profiles if they are submitted to different growth conditions, due to up- or down-regulation of biosynthesis pathways in response to these different environments (Moore et al., 2014). Over evolutionary time scales, the selective pressure exerted by environmental factors may lead to the selection of new genotypes determining a better chemical adaptation of the plant to

its environment (Foutami et al., 2018). Genetic mutations selected through this evolution process can determine the presence or absence of a metabolite, but in the large majority of cases, they are responsible for quantitative variations in metabolite content (Moore et al., 2014).

Here, we study clary sage populations sampled in three different types of macro-environment encountered around the Croatian city of Makarska: the coast, the mountain range alongside the coast and the plateau behind the mountain range (**Figure 36**). *In situ*, populations growing in these three macro-environments are submitted to different environmental conditions specific to coastal, mountain and plateau environments. Sampling sites are notably characterized by their distinct elevation (40-80 m for coastal and plateau sites and 360-550 m for mountain sites). The selective pressure associated with these environmental conditions may have led to the selection of genetic polymorphisms which could potentially have an impact on sclareol production. To assess the presence of genetic variation, genes involved in sclareol production were sequenced in individuals belonging to the different Croatian populations. Detected single nucleotide polymorphisms were used to estimate the genetic diversity found among Croatian populations and to compare it to genetic diversity found in reference populations. Finally, individuals from Croatian and reference populations were grown in the same environment in order to investigate whether genetic variation between the populations has an impact on sclareol content and flowering time.

2. Materials and methods

2.1. Plant material

Seeds from 6 wild populations of clary sage were sampled in summer 2016 around the city of Makarska in Croatia. Localization of sampling sites is indicated in **Figure 36**, along with their elevation above sea level. For each population, seeds from 10 individuals were collected and stored at 4 °C and 10% humidity until sowing. Vatican White seeds were purchased from Jelitto (Germany). Toscalia and Milly seeds were kindly donated by the French National Conservatory of Perfume, Medicinal, Aromatic and Industrial Plants (CNPMAI).

A



B



C

Site	Population	Environment	Altitude
1	Makarska	Coast	80 m
2	Blato	Coast	80 m
3	Zaostrog	Coast	40 m
4	Ravca	Mountain	550 m
5	Kljenak	Mountain	360 m
6	Banja	Plateau	80 m

Figure 36: Localization of Croatian populations used in the study.

A: Aerial view of Makarska, Croatia.

B: Localization of seed sampling for the 6 Croatian populations of wild clary sage.

C: Elevation above sea level for each sampling site.

2.2. Genetic diversity analysis

For each population, 3-4 individuals were grown in growth chamber (16h of day; 27 °C during the day, 21 °C at night; 60% hygrometry) for 6 weeks. DNA was extracted from leaves using the DNeasy 96 Plant Kit from Qiagen. Genomic fragments from *DXS2*, *CMK* and *ITS* loci were amplified by PCR using primers listed in **Table 2**. Sequences of PCR products were obtained through Sanger sequencing performed by Eurofins Genomics. For each individual, sequences from *DXS2*, *CMK* and *ITS* loci were concatenated in order to integrate data from the 3 loci in a single phylogenetic tree. Single nucleotide polymorphisms (SNPs) were located by aligning the sequences using the ClustalW algorithm in the MEGA7 software. Heterozygous sites were visualized in Sanger sequencing chromatograms provided by Eurofins Genomics and manually annotated in MEGA7 using the IUPAC code. Encoding heterozygous sites using the IUPAC code was necessary to reconstruct haploid sequences using the PHASE algorithm in the DnaSP software. Haploid sequences obtained (2 per individual) were used to build a phylogenetic tree in MEGA7 using the Neighbor Joining method with 500 bootstrap repetitions.

Target	Primer	Sequence
<i>DXS2</i>	<i>DXS2_F</i>	5'-GCAGTTTCTTGCCATTGCTCC-3'
	<i>DXS2_R</i>	5'-TAATACGTACCTGGTGACCC-3'
<i>CMK</i>	<i>CMK_F</i>	5'-CGAGAGGTACAGGTGGAGGA-3'
	<i>CMK_R</i>	5'-CAAGAGTGGTCGGGTGAGAT-3'
<i>ITS</i>	<i>ITS_F</i>	5'-GCATCGATGAAGAACGTAGC-3'
	<i>ITS_R</i>	5'-TCCTCCGCTTATTGATATGC-3'

Table 2: Primers used to amplify genomic fragments from *DXS2*, *CMK* and *ITS* loci.

2.3. Field trial

Seeds from Makarska, Ravca, Banja, Milly and Vatican White populations were sown on August 3rd, 2017. Plantlets were grown in growth chamber (16h of day; 27 °C during the day, 21 °C at night; 60% hygrometry) until being transplanted in the field in Chemillé (France) on November 7th, 2017. Plants were placed every 0.4 m x 0.7 m according to a randomized layout. Flowering time was defined as the date of opening of the first flowers and was recorded between May 31st and June 29th, 2018. For calyx sampling, inflorescences were cut

on July 19th, 2018 at the beginning of the afternoon. Mature calyces (harboring green achenes) were cut from all over the inflorescence, immediately placed in dry ice and stored at -80 °C at the end of the day until further analysis.

2.4. Sclareol and linalyl acetate quantification by GC-MS

Clary sage calyces were immersed in hexane without grinding to extract metabolites present at the surface. For each individual, 12 mL of solvent were used to extract metabolites from 8 mature calyces. 10-undecen-1-ol was employed as internal standard and was added in the solvent before extraction at a concentration of 0.75 mM. After 2 h of agitation, 1 mL of extract was centrifuged at 13,000 rpm for 10 min to eliminate potential dust or debris. Samples were directly analyzed by gas chromatography coupled to mass spectrometry (GC-MS) using a 7890B/5977A GC-MS system from Agilent and according to a protocol adapted from Laville et al. (Laville et al., 2012). The system was equipped with a 10 m guard column and an Rxi-5Sil MS column (length, 30 m; inner diameter, 0.25 mm; film thickness, 0.25 µm) (Restek, Bellefonte, PA, USA). 1 µl of sample was injected with a split ratio of 30:1. Oven temperature was set to 110 °C for 2 min, then increased to 270 °C at a rate of 10 °C/min, and finally set to 270 °C for 2 min. Other temperatures were set as follows: injector, 270 °C; transfer line, 270 °C; source, 230 °C; quadrupole, 150 °C. The carrier gas was helium at a constant flow of 1 mL/min. The quadrupole mass spectrometer was switched on after a solvent delay of 5 min and was programmed to scan from 35 u to 350 u. Data analysis was carried out using the MassHunter Quantitative Analysis software from Agilent. Sclareol and linalyl acetate were identified by comparison with analytical standards purchased from Sigma-Aldrich. Absolute quantification of sclareol and linalyl acetate was performed using a calibration curve prepared with analytical standards.

3. Results

3.1. Croatian populations show genetic polymorphism in MEP pathway genes

As mentioned above, we hypothesized that distinct environmental conditions associated with coastal, mountain and plateau environments may have led to the selection of genetic polymorphisms potentially impacting sclareol production. To support this hypothesis, we checked whether genetic polymorphisms can be found among Croatian populations in genes involved in sclareol biosynthesis. We looked for genetic polymorphisms in the sequences of 2 genes of the MEP pathway: *DXS2* and *CMK*. These two genes were selected because substantial parts of their genomic DNA could be readily amplified by PCR and sequenced. In addition to sequences from the 6 Croatian populations, sequences from 3 reference populations were analyzed: Vatican White and Milly, which were both used as reference populations in the context of the PISTILL project, and Toscalia, which is a population of cultivated clary sage. Single nucleotide polymorphisms (SNPs) were detected in partial sequences obtained from *DXS2* and *CMK* (respectively 4 and 9 SNPs) (**Figure 37a**). Five of these SNPs are found in individuals from Croatian populations but not in reference populations, indicating that Croatian populations represent a source of new alleles (**Figure 37b**). Most of these SNPs probably have no impact on enzyme activity, for example if they are located in introns or if they are silent due to genetic code redundancy. However, analyzed sequences are small and other SNPs could probably be found in non-sequenced parts of *DXS2* and *CMK* genes and also in other genes submitted to an equivalent selection pressure. Since the existence of specific genetic polymorphisms in genes involved in sclareol biosynthesis has been assessed among Croatian populations, their sclareol content may be different compared to reference populations.

A

<i>DXS2</i> partial sequence (434 bp)	GAATTTATCTATGTTTTGTTCTCTATCAAGATTATTTGAGATTGATCGCATCTTGATGTTTCTCTTTCCAGT TCGCCGTGGTAGCARCTCTTCAACAGGATAACACCAACGACATGGTTGCAAGTGGYGGAGAGAGTCTGAC GACGACGAGGCACAAAACAAGAGCTCTGAATTTACGGGAGAGAAGCCTCCTACCCAATATTGGATACC ATCAACTATCCAATCCACATGAAAAACCTCTCTCAGGTAATATAGATWAWTTTTCGAAAATTCTTGAG AAATTTAGATGTCCATAAATTTGCTAAAAATTGTGATGTAACAAGATGAAAATTGATCGTGTAGGACGA CTCATGCATATAGAAATTGGAATGTGTTTTATTTCATAAAAACCTCTTGATCTTGATACACATTAAGGAA AGTCAA
<i>CMK</i> partial sequence (366 bp)	TATACAAAAGAAATTAGTAACAAGAGGAAASYTTGTTATWCTCAAGACTGCTTCAGCRTTSAGCATTATAT GTTGGTGAATCATGTGAGGTGAGAAATAACAACATTTCAGGTCTACCTTGATTAAGGATTAACCTTATATA GCCTCCAATAAGCAGARTCAGATGTATAAGCATATGATTAATCATGTAAAGCTTTGTGCTTATTCAAATCAA AATCGACACCTACAAAATGCTACAGAAAGTTGAAAGACATGAACAAGGATGAGACGGGCTGRATGCAC ACAACTTAAGACTTACCTCTCTACCCGTGCAATATGCAGCACCRGTGWGAGAAAAAGAAGGGGATAT CGGAGCCAATCT

B

	Population	Individual	SNPs found in <i>DXS2</i>				SNPs found in <i>CMK</i>								
			R	Y	W	A	C	Y	A	R	C	R	A	R	T
Croatian populations	Makarska	1	R	Y	W	A	C	Y	A	R	C	R	A	R	T
		2	G	C	A	A	C	C	A	A	C	A	A	A	T
		3	G	C	A	A	C	Y	A	R	C	R	A	R	T
		4	R	Y	W	A	C	Y	A	R	C	R	A	R	T
	Ravca	1	G	Y	W	W	C	T	A	G	C	G	A	G	T
		2	G	T	T	T	C	Y	A	G	S	R	R	R	T
		3	G	T	T	T	C	T	A	G	C	G	A	G	T
		4	G	T	T	T	C	Y	A	G	S	R	R	R	T
	Kljenak	1	G	Y	W	W	C	Y	A	R	C	R	A	R	T
		2	G	C	A	A	C	C	A	A	C	A	A	A	T
		3	G	C	A	A	C	Y	A	R	C	R	A	R	T
	Banja	1	A	T	T	A	C	T	A	G	C	G	A	G	T
		2	A	T	T	A	C	T	A	G	C	G	A	G	T
		3	A	T	T	A	C	T	A	G	C	G	A	G	T
		4	R	Y	W	A	C	Y	A	R	C	R	A	R	T
	Zaostrog	1	G	Y	W	W	C	Y	A	R	C	R	A	R	T
2		G	Y	W	W	C	Y	A	R	C	R	A	R	T	
3		G	Y	W	W	C	Y	A	G	S	R	R	R	T	
4		R	T	T	W	C	Y	A	G	S	R	R	R	T	
Blato	1	G	Y	W	W	C	C	A	R	S	A	R	A	T	
	2	G	Y	W	W	C	C	A	G	G	A	G	A	T	
	3	G	C	A	A	C	C	A	R	S	A	R	A	T	
Reference populations	Milly	1	G	T	T	T	C	C	A	A	C	A	A	A	T
		2	G	T	T	T	C	C	A	A	C	A	A	A	T
		3	G	T	T	T	C	C	A	A	C	A	A	A	T
		4	G	T	T	T	C	C	A	R	C	A	R	A	T
	Toscalia	1	A	T	T	A	C	Y	A	R	C	A	R	A	W
		2	A	T	T	A	S	C	W	R	C	A	A	A	T
		3	R	T	T	W	C	Y	A	G	C	A	A	A	W
		4	A	T	T	A	S	C	W	R	C	A	A	A	T
	Vatican White	1	G	T	T	T	C	C	A	R	C	A	R	A	T
		2	G	T	T	T	C	C	A	R	C	A	R	A	T
		3	G	T	T	T	C	C	A	R	C	A	R	A	T
		4	G	T	T	T	C	C	A	R	C	A	R	A	T

Figure 37: Single nucleotide polymorphisms (SNPs) found in *DXS2* and *CMK* sequences from Croatian and reference clary sage populations.

A: SNPs found in *DXS2* and *CMK* sequences are highlighted in yellow and encoded using the IUPAC code. **B:** Genotype of individuals used to estimate genetic diversity. Heterozygous genotypes are encoded using the IUPAC code (**R** = A or G; **S** = C or G; **W** = A or T; **Y** = C or T).

3.2. Choice of the populations to be analyzed in the field

To test whether genetic differences between Croatian and reference populations have an impact on sclareol production, a field trial was carried out in order to compare their sclareol content. The first step was the choice of populations which are the most relevant for the study. Indeed, Croatian populations being geographically close from each other (**Figure 36**), some of these populations may be almost genetically identical. Since the field trial is aimed at investigating the impact of genetic variation on plant phenotype, populations selected for the field trial should be genetically distinct. Moreover, the study of the link between genotype and phenotype is easier if each population used in the study is as genetically homogenous as possible.

Genetic diversity among Croatian and reference populations and genetic homogeneity of each population were evaluated using SNPs found in *DXS2* and *CMK* sequences, along with 3 additional SNPs detected in Internal Transcribed Spacer (*ITS*) sequences. To integrate polymorphism data detected at these 3 loci, the 3 sequences obtained for each individual were concatenated. Raw concatenated sequences could not be used directly to build a phylogenetic tree because heterozygous sites are not properly taken in account by tree building algorithms. The phasing algorithm PHASE was used to reconstruct the 2 haploid sequences carried by each individual. In total, 6 or 8 haploid sequences per population were obtained, depending on the number of individuals sequenced (3 or 4, respectively). All haploid sequences were considered to be independent and used to build a phylogenetic tree (**Figure 38**). For example, the 4 individuals belonging to the Vatican White population are heterozygous and carry the same 2 haploid sequences (**Figure 37b**). As a consequence, the 8 haploid sequences of the Vatican White population form 2 groups of 4 identical sequences which cluster together in the tree (**Figure 38**).

Chapter 4 - Clary sage natural diversity

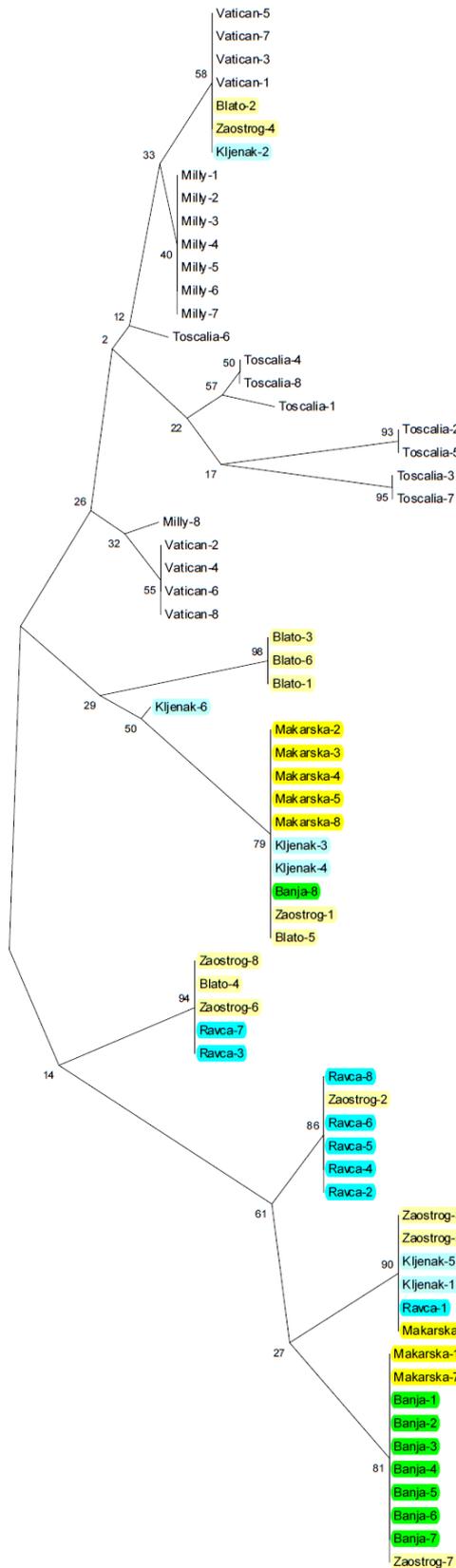


Figure 38: Estimation of genetic diversity found in Croatian and reference clary sage populations.

The tree was built with the Neighbor Joining method using concatenated haploid sequences from *ITS*, *DXS2* and *CMK* loci. The 2 haploid sequences carried by each individual were considered to be independent. The percentage of replicate trees in which haploid sequences clustered together in the bootstrap test (500 replicates) is shown next to the branches.

Some haploid sequences from Vatican White and Croatian populations cluster together. However, in general, reference populations on the one hand, and Croatian populations on the other, are found in distinct clades (**Figure 38**). This result illustrates that Croatian populations carry alleles which are not found in reference populations and represent a distinct genetic resource. All populations have their haploid sequences scattered in different clades of the tree and the scattering pattern is different for each population. Notably, haploid sequences from Banja and Ravca populations are not found in the same clades. Haploid sequences from Banja and Makarska populations cluster together, but the proportions of haploid sequences found in each clade are different (**Figure 38**). These observations suggest that Croatian and reference populations used in the study are not genetically equivalent between themselves. Besides, haploid sequences from Kljenak, Zaostrog, Blato and Toscalia populations are dispersed in at least 4 different clades, whereas haploid sequences from Makarska, Ravca, Banja, Milly and Vatican White populations are clustered in 2 or 3 different clades (**Figure 38**). For this reason, the 5 latter populations were considered more homogenous and selected for the field trial.

3.3. Comparison of flowering date and metabolite content of the different populations

In autumn 2017, 93 plants (between 15 and 23 per population) were transplanted in the field using a random layout. In total, 73 plants (between 10 and 22 per population) flowered between May 31st and June 29th, 2018, that is, between 301 and 330 days after sowing. These flowering dates are consistent with data available in the literature; indeed, clary sage is usually reported to flower in June-July (Wagner et al., 2012). The date of flowering does not significantly vary between the three Croatian populations (**Figure 39**). Interestingly, Vatican White plants flowered significantly earlier than Milly and Makarska plants (10 days earlier in average) (**Figure 39**).

Calyces from each individual were sampled in the field in summer 2018 and sclareol and linalyl acetate were quantified by GC-MS (**Figure 40**). No significant difference in sclareol and linalyl acetate content of calyces could be detected between the three Croatian populations, which produce the same amount of sclareol and linalyl acetate as Milly plants. Interestingly,

a significant difference was detected between the 2 reference populations: Vatican White plants produce 20% less sclareol and 30% less linalyl acetate than Milly plants (**Figure 40**).

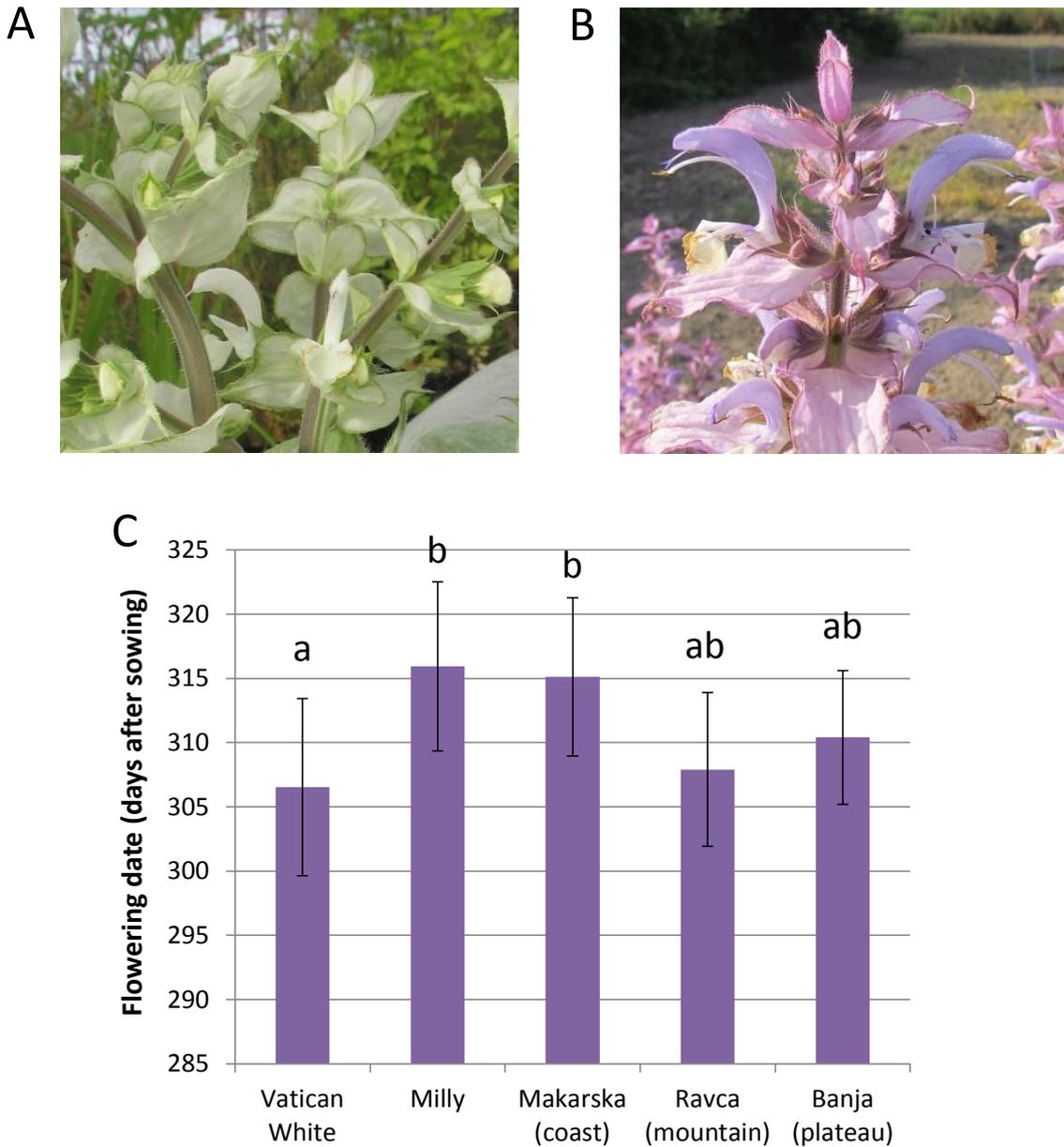


Figure 39: Flowering of Croatian and reference clary sage populations.

A, B: Flowers of representative Vatican White (**A**) and Milly (**B**) clary sage plants.

C: Date of flowering, expressed in days after sowing. The day of opening of the first flowers was considered as the flowering date. Letters indicate the result of a Kruskal-Wallis rank test: values harboring the same letter are not significantly different.

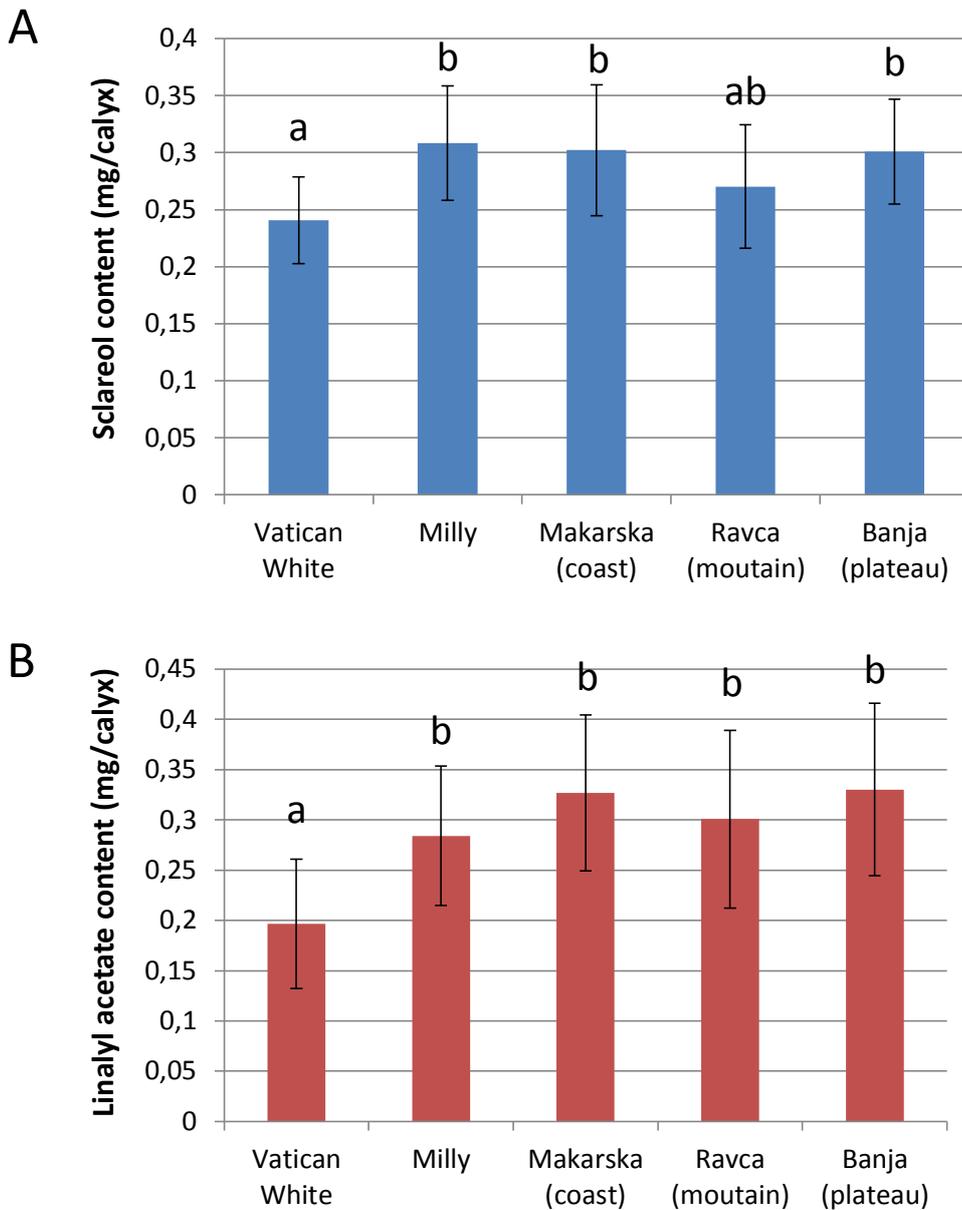


Figure 40: Calyx metabolite content of Croatian and reference clary sage populations.

Sclareol (A) and linalyl acetate (B) were quantified by GC-MS after hexane extraction. Letters indicate the result of a Kruskal-Wallis rank test: values harboring the same letter are not significantly different.

4. Discussion

4.1. Complex relationships between genetic diversity, chemical diversity and environmental conditions

The production of specialized metabolites by plants is known to be impacted by environmental conditions. In *Salvia desoleana*, a sage species closely related to *Salvia sclarea*, sclareol content was shown to be influenced by soil chemical composition: plants grown on soil rich in silt produced less sclareol compared to plants grown on soil rich in sand (Rapposelli et al., 2015). Other abiotic factors such as temperature or hygrometry may influence sclareol production by clary sage. Here, we study Croatian populations of clary sage naturally growing in three different types of environment: coast, mountain and plateau. These three types of environment are probably submitted to distinct climatic conditions, notably because they are located at different altitudes above sea level (40-80 m for coast and plateau, 360-550 m for mountain) (**Figure 36**).

The impact of altitude on clary sage terpene production was recently investigated using herbarium samples. Sampling sites were located in several European and Middle Eastern countries at different altitudes, ranging from 34 m to 1300 m (Foutami et al., 2018). According to this study, altitude had no significant effect on clary sage sample chemical composition. Another study focused on the impact of environmental conditions on terpene production in a related species, *Salvia officinalis* (Russo et al., 2013). Plants were cultivated at 18 sites located around Campobasso in Italy and grouped in three categories according to global environmental conditions: lowland, low hill and high hill. Terpene production was evaluated through the analysis of essential oil chemical composition. The main essential oil components were always α -thujone, camphor, borneol, γ -muurolene and sclareol, but their relative abundances in essential oil varied according to cultivation site. Notably, the percentage of sclareol in essential oil ranged between 5.9 and 23.1%, although these variations could not be attributed to a particular environmental factor such as altitude or water availability (Russo et al., 2013).

In addition to the influence of environmental conditions, the production of specialized metabolites by plants also depends on genetic factors. Relationships between genetic diversity, chemical diversity and environmental conditions were investigated in 50 Sardinian populations of the terpene-producing plant *Helichrysum italicum* (Melito et al., 2013).

Genetic diversity was estimated using the AFLP (Amplified Fragment Length Polymorphism) DNA fingerprinting method, and chemical diversity was evaluated through the characterization of leaf essential oil composition. AFLP-based clades were in general congruent with the geographical localization of *Helichrysum italicum* populations, indicating limited gene exchanges between populations. The genetic structure was best explained by environmental conditions at sampling site, suggesting an involvement of adaptive selection in shaping genetic diversity. The amount of oxygenated monoterpenes present in the essential oil was significantly correlated with AFLP-based clades. Therefore, genetic factors may be partly responsible for differences in metabolite content. However, the strong association of the genetic structure with environmental conditions at sampling site prevented the analysis of the respective contributions of genetic factors and environmental conditions to chemical diversity (Melito et al., 2013).

In our study, we cultivated three Croatian populations in the same environment (field trial) to focus on differences in metabolite content attributable to genetic variation. Beforehand, alignment of *ITS*, *DXS2* and *CMK* sequences underlined the presence of SNPs in the genome of individual plants belonging to Croatian populations (**Figure 37**). Haploid sequences (harboring various combinations of SNPs) were used to estimate genetic diversity among Croatian populations (**Figure 38**). Each haploid sequence was present in at least two populations, highlighting the existence of gene exchanges between Croatian populations, or pointing to an incomplete genetic differentiation from the same initial pool of alleles. Frequencies of the different haploid sequences varied among Croatian populations, suggesting that Croatian populations were not genetically equivalent. However, flowering time, sclareol content and linalyl acetate content were not significantly different between the three populations analyzed in the field trial (**Figures 39 and 40**). Climatic conditions encountered in coastal, mountain and plateau environments may not have led to significant genetic differentiation of clary sage populations. If the impact of selective pressure is limited, it may be masked by potential gene flows due to pollen transport between populations. The extent of gene exchanges may have been underestimated due to the limited number of genetic markers used to evaluate genetic diversity among Croatian populations.

4.2. Perspectives for the study of clary sage natural diversity

Different types of genetic markers can be developed and employed to study the genetic diversity of plant populations. For example, SSR (Simple Sequence Repeats, also known as microsatellites) and ISSR (Inter Simple Sequence Repeats) markers have been developed for the study of patchouli populations (Sandes et al., 2016), and the AFLP method (Amplified Fragment Length Polymorphism) was used to analyze 50 Sardinian populations of *Helichrysum italicum* (Melito et al., 2013). The AFLP method is cheap and does not require any genomic data about the species of interest, but provides limited information due to the relatively low number of markers generated by this approach. For instance, only 125 AFLP markers were generated to analyze the genetic structure of the 50 Sardinian populations of *Helichrysum italicum* (Melito et al., 2013).

Generation of SNP markers through genotyping-by-sequencing provides much higher marker density compared to other methods (He et al., 2014). This approach was recently employed to evaluate the genetic diversity and reveal the genetic structure of cultivated chamomile populations (Otto et al., 2017). 16,972 SNPs were simultaneously identified and genotyped by sequencing the genome of 2-4 individuals from each of the 33 populations used in the study (91 individuals in total). In absence of a reference genome, the chamomile genome was assembled *de novo* with sequences obtained from the 91 individuals. Unlike other methods like AFLP or SSR marker identification, genotyping-by-sequencing requires significant genome-wide sequencing efforts, but technological advances continuously reduce the cost of high-throughput sequencing (He et al., 2014). Due to their abundance in genomes and their increasingly affordable genotyping, SNPs tend to gradually replace other markers (Kilian and Graner, 2012).

The recent study of chamomile diversity mentioned above illustrates how genotyping-by-sequencing and GWAS approaches can contribute to increase knowledge on non-model plants and facilitate their breeding (Otto et al., 2017). The flowers of chamomile contain numerous medicinally active specialized metabolites, notably the sesquiterpene α -bisabolol. Therefore, flowering time and α -bisabolol content are important traits for chamomile breeding. SNP markers generated through genotyping-by-sequencing were used for GWAS, leading to the identification of genetic loci controlling flowering time and α -bisabolol content (Otto et al., 2017). A similar approach could be employed to identify genetic loci

controlling flowering time and sclareol content in clary sage. However, to perform successful GWAS, the collection of populations must be large enough and display sufficient genetic diversity (Kilian and Graner, 2012). To meet these criteria, clary sage seed sampling should be extended to a larger geographic area than the region of Makarska in Croatia. Indeed, more natural genetic diversity is probably present in the native area of clary sage and would complement new alleles already provided by Croatian populations used in this study (**Figures 37 and 38**). It could also be relevant to sample cultivated populations of clary sage grown around the world.

4.3. Vatican White plants flower earlier and produce less specialized metabolites

Phenotyping of flowering date and metabolite content during the field trial highlighted significant differences between the two clary sage populations used as reference populations, Vatican White and Milly. Vatican White plants flowered 10 days earlier and produced 20% less sclareol and 30% less linalyl acetate than Milly plants (**Figures 39 and 40**). For this reason, if a GWAS approach is planned in the context of clary sage breeding, it would be judicious to include Vatican White and Milly populations in the collection of clary sage populations.

The Milly population is a French population of clary sage maintained by the French National Conservatory of Perfume, Medicinal, Aromatic and Industrial Plants (CNPMAI), whereas the Vatican White population is propagated and commercialized for ornamental purposes, notably owing to the distinct color of its flowers. Indeed, Vatican White flowers are generally white or light purple, whereas flowers of other populations, including Milly and Croatian populations, display darker shades of pink and purple (**Figure 39**). In other plant species, the pink or purple color of flowers is often due to the presence of specialized metabolites belonging to the flavonoid family (Tanaka et al., 2008). The lighter color of Vatican White flowers may be the consequence of reduced flavonoid biosynthesis.

A simultaneous decrease in flavonoid biosynthesis and terpenoid biosynthesis may be explained by a global downregulation of specialized metabolism in Vatican White flowers. If this downregulation is operated at transcriptome level, it would be relevant to conduct a comparative transcriptomic study of Milly and Vatican White flower organs. A differential

expression of biosynthetic genes would validate the initial hypothesis, and differentially expressed transcription factors would be good candidates for an involvement in specialized metabolism regulation or flowering time control. Since sclareol and linalyl acetate are mainly produced by glandular trichomes (**Chapter 1**), the lower metabolite content observed in Vatican White plants may also be due to a lower glandular trichome density. In *Artemisia annua*, the amount of artemisinin found in leaves is tightly correlated to glandular trichome density (Xiao et al., 2016). Careful measurements of glandular trichome density are needed to test this hypothesis.

Conclusion

1. Context and objectives of the project

Clary sage produces sclareol, a terpenoid widely used in perfume industry for the hemisynthesis of ambroxide, a valued perfume component with remarkable fixative properties. In response to the development of ambroxide use in functional perfumery, the global demand for sclareol currently raises, prompting efforts aiming at increasing sclareol supply. The PISTILL project, driven by the Flower and Carpel Development Group at the Institute of Plant Sciences Paris-Saclay, aims at increasing the yield of sclareol production from clary sage through targeted genetic improvement. The purpose of the work presented in this manuscript was to enhance knowledge on sclareol biosynthesis and its regulation in clary sage, with a dual objective: orientating the PISTILL project towards promising strategies to boost sclareol yield, and improving our general understanding of terpenoid production in plants. We investigated the cellular and metabolic origin of sclareol in clary sage calyces, tried to identify regulation mechanisms controlling sclareol accumulation, and started to explore how clary sage natural diversity could be exploited. This study opens up perspectives for basic research aiming at understanding sclareol production in clary sage, and at the same time highlights interesting avenues for targeted genetic enhancement of clary sage performances.

2. Perspectives for research on sclareol production in clary sage

In the first chapter of this manuscript, we provide several lines of evidence supporting the hypothesis that sclareol is mainly secreted by calyx glandular trichomes. Mass spectrometry imaging suggests that sclareol is most abundant on large capitate glandular trichomes, and clary sage plants lacking these trichomes produce less sclareol. First, it is important to assess whether the absence of large capitate glandular trichomes and the lower sclareol content are caused by the same mutation, and if so, to identify this mutation by positional cloning. The characterization of this mutation would improve our understanding of biological processes underlying glandular trichome development and sclareol production in clary sage.

Conclusion

Approaches developed in the second chapter of this manuscript underline the respective contributions of MVA and MEP pathways to the biosynthesis of three terpenoids in clary sage calyces: sclareol, linalyl acetate and β -caryophyllene. Sclareol and linalyl acetate both originate from the plastidial MEP pathway, while the sesquiterpene β -caryophyllene seems to be of mixed origin. However, the origin of carbon assimilates needed for MVA and MEP pathways activity is still unknown. They may be produced by photosynthesis in clary sage glandular trichomes, or be imported from neighboring cells. The ultrastructural and metabolic specificities of glandular trichomes and their relationships with other calyx cells could be investigated through transmission electron microscopy and dynamic isotope labeling experiments.

Experiments described in the third chapter of this manuscript could not demonstrate any role of jasmonates in the regulation of sclareol production. Two possibilities remain plausible: either jasmonates are not involved in sclareol biosynthesis regulation, or the experimental protocol needs further optimization. In the latter case, employing a mild treatment protocol and focusing the study on young developing calyces may help to highlight an impact of jasmonates on sclareol content. The role of other phytohormones like ABA, known to control terpenoid biosynthesis in other plants, could also be investigated.

The fourth chapter of this manuscript explores genetic and phenotypic diversity of Croatian wild populations originating from three different types of environment: coast, mountain and plateau. Although genetic polymorphisms were detectable in sclareol biosynthesis genes, Croatian populations were not different in terms of sclareol content and flowering time when grown in the same environment. Croatian climatic conditions apparently did not result in genetic differentiation affecting sclareol production. However, the difference in sclareol content observed between “Vatican White” and “Milly” reference populations deserves further investigation. It could be the consequence of a global downregulation of specialized metabolism or of a decrease in glandular trichome density in “Vatican White” plants. To test these hypotheses, glandular trichome density and expression levels of biosynthetic genes should be compared between the two reference populations.

3. Perspectives for clary sage genetic improvement

Several strategies aiming at enhancing sclareol production in clary sage emerge from results described in this manuscript. Sclareol being MEP-derived and mainly produced in glandular trichomes, increasing glandular trichome density or boosting MEP pathway activity are promising ways to elevate sclareol content. Moreover, since linalyl acetate is also MEP-derived and produced in glandular trichomes, blocking linalyl acetate biosynthesis may increase sclareol production through reducing competition for substrate. The MEP pathway and linalyl acetate biosynthesis could be manipulated through genetic modifications targeting enzymes involved in these pathways. Similarly, glandular trichome density could be increased through the manipulation of glandular trichome initiation regulators. Candidate genes putatively encoding DXS, DXR and GPP-synthase enzymes have been identified in the transcriptome previously obtained from clary sage calyces (Legrand et al., 2010), but transcription factors involved in glandular trichome initiation remain to be characterized. Candidate genes could be identified in the same transcriptome by looking for genes displaying high sequence similarity with conserved regulators of glandular trichome initiation.

Clary sage breeding programs could be enhanced through an efficient exploitation of clary sage natural diversity. Croatian populations studied in the fourth chapter of this manuscript represent a valuable genetic resource because they have been shown to carry new alleles compared to reference populations. Extending clary sage seed sampling to the whole Mediterranean basin, where clary sage is naturally growing, would increase the amount of genetic resources available for clary sage breeding and for the study of sclareol biosynthesis regulation.

The feasibility of approaches described above will depend on the development of tools dedicated to clary sage genetic improvement. The sequencing of clary sage genome and robust genetic markers would enable the identification of genetic loci controlling traits of interest and the rapid introgression of new alleles into elite varieties. The development of a clary sage transformation protocol would also facilitate the functional characterization of candidate genes involved in the quantitative variation of traits of interest. Clary sage transgenic hairy roots can be obtained (Kuźma et al., 2009a; Vaccaro et al., 2014, 2017; Alfieri et al., 2018), but to our knowledge, no protocol allowing stable transformation of

Conclusion

whole clary sage plants has been described so far. Genomic data and biotechnological tools specific to clary sage would be highly valuable for basic and applied research projects aiming at understanding and improving this perfume plant.

Annex

Review: Genetic control of glandular trichome development

1. Introduction

1.1. Why study the genetic control of glandular trichome development?

Scents, pigments, medicines... Life would be bleaker and more difficult for humans if plants were not producing such a wide variety of compounds. These chemicals are critical for the capacity of plants to adapt to their environment and to overcome the various challenges they are facing every day, like pollinator attraction or defense against pathogen attacks. Due to their frequent self-toxicity, many of these compounds are produced by specialized secretory structures, for example glandular trichomes. Glandular trichomes are quite common as they can be found in approximately 30% of all vascular plants (Fahn, 2000; Glas et al., 2012). The development of these multicellular structures originating from the epidermis has been suggested as a model to study plant cell differentiation (Yang and Ye, 2013). Glandular trichomes are also remarkable for their very high metabolic productivity (Balcke et al., 2017; Tissier, 2018).

Compounds produced by glandular trichomes are harnessed by human industries that benefit from their various properties. Organic synthesis often made it possible to dispense with the initial plant source. However, total synthesis can be complex and expensive, and many compounds of interest are still directly extracted from plants. For example, the perfume ingredient ambrox is synthesized from sclareol extracted from *Salvia sclarea* (Yang et al., 2016), and world supply in the anti-malarial drug artemisinin still relies on artemisinin extraction from *Artemisia annua* (Xiao et al., 2016; Tissier, 2018). The amount of specialized metabolites produced by a plant is often tightly correlated to the density of glandular trichomes present at the surface of the epidermis (Tissier, 2012; Xiao et al., 2016). Increasing glandular trichome density has recently emerged as a new plant breeding strategy to enhance the yield in compounds of interest for the pharmaceutical sector (Xiao et al., 2016). This strategy could also be used to breed crops with improved resistance to herbivores (Glas

et al., 2012). In some cases, decreased trichome density is also desirable in order to reduce the amount of compounds toxic to humans, like gossypol in cotton (Ma et al., 2016).

Engineering glandular trichome density requires reliable data about the genetic network controlling glandular trichome initiation. Several lines of evidence indicate that glandular (and in general multicellular) trichome formation is probably controlled by a different network than the one controlling non-glandular trichome formation in *Arabidopsis thaliana* (Yang and Ye, 2013). Compared to non-glandular trichome formation, our knowledge about genes involved in glandular trichome formation is still limited. Nevertheless, recent studies have led to significant advances, which are summarized in this review.

1.2. Which species could serve as model(s)?

As no glandular trichomes are found in *Arabidopsis thaliana*, research on glandular trichome development has been carried out on various other plant species. From the careful analysis of recent literature, three species emerge as main working materials: tomato (*Solanum lycopersicum*, Solanaceae), sweet wormwood (*Artemisia annua*, Asteraceae) and cucumber (*Cucumis sativus*, Cucurbitaceae). In tomato, glandular trichomes are essential for plant defense against herbivores (Kang et al., 2010; Glas et al., 2012); in *Artemisia annua*, they produce the anti-malarial drug artemisinin (Xiao et al., 2016); and in cucumber, their size and number are important fruit quality traits (Liu et al., 2016; Yang et al., 2018). Tomato and *Artemisia annua* have both already been suggested as good potential models for the study of glandular trichome development (Tissier, 2012; Shi et al., 2018). Tomato has long been established as a model plant in other fields of plant research; sequenced genome, reliable genetic tools and extensive genetic resources are available for *Solanum lycopersicum* and other related species (Tissier, 2012). Research efforts focusing on *Artemisia annua* are more recent, but genetic transformation protocols are available (Shen et al., 2016) and a draft assembly of the genome has been recently published (Shen et al., 2018).

Eight trichome types have been described on tomato leaves, among which four are glandular: type I, type IV, type VI and type VII. Type VI glandular trichomes are the most abundant and secrete mainly terpenoids, whereas type I and type IV are involved in acyl sugar biosynthesis (Glas et al., 2012). The current classification of tomato glandular trichomes is still discussed; for instance, type I and type IV were suggested to be the same

type according to the molecules they secrete (McDowell et al., 2011), but are generally considered to be different according to their morphology and patterning (Glas et al., 2012; Vendemiatti et al., 2017). *Artemisia annua* leaves display two types of trichomes: T-shaped non-glandular ones and glandular ones, which are able to accumulate artemisinin along with various other compounds (Duke and Paul, 1993; Xiao et al., 2016; Yan et al., 2017). Eight trichome types have been recently characterized on cucumber fruits including two glandular types: type I and type VI, type I being the most frequent of the two (Xue et al., 2018). Type I glandular trichomes are also called bloom trichomes and are believed to be involved in fruit cuticle formation and in the secretion of mineral substances (Xue et al., 2018). Given the fact that the genetic control of multicellular trichome development in cucumber fruits has been recently reviewed (Liu et al., 2016), this review focuses on recent discoveries concerning genes involved in glandular trichome development in tomato and *Artemisia annua*.

2. Description of glandular trichome morphology and development

Glandular trichome morphology has been described in a large number of plant species and this abundant literature highlights their tremendous diversity of shape and size (Tissier, 2012). Nevertheless, a common organization scheme shared by most glandular trichomes arises from the description of their structure. Glandular trichomes are usually multicellular and composed of 3 parts: a base, a stalk and a gland. The gland is responsible for the secretion of specialized metabolites. Gland cells are highly differentiated, with an ultrastructure adapted to their very high metabolic activity. Ultrastructural adaptations depend on the nature of secreted compounds and have been previously reviewed by Lange and Turner (Lange and Turner, 2013). The stalk is the structure bearing the gland and the base connects the stalk to surrounding epidermal cells. Each of these 3 parts can be unicellular or multicellular and cells can be more or less elongated. This variability in cell number and shape accounts for a large part of the high morphological diversity found among glandular trichomes and is also responsible for their wide variety of sizes. For example, 10-celled glandular trichomes of *Artemisia annua* have a biseriate structure of only 40-50 μm long (Duke and Paul, 1993), whereas tomato type I glandular trichomes are 2-3 mm long with a long stalk (Glas et al., 2012).

Secreted metabolites often accumulate in a storage cavity. This storage cavity can be subcuticular: in that case, molecules secreted at the top of gland cells accumulate under the cuticle which is gradually pushed away from the cell wall, as seen in *Artemisia annua* glandular trichomes (Duke and Paul, 1993). The storage cavity can also be intercellular, as seen in type VI glandular trichomes of tomato (Bergau et al., 2015). The size of the storage cavity has an impact on glandular trichome shape. For example, in cultivated tomato, type VI glandular trichome glands have a four-leaf clover shape due to the small size of the storage cavity, whereas in the wild tomato species *Solanum habrochaites*, a larger storage cavity is responsible for their spherical shape (Bergau et al., 2015). An abscission zone between the stalk and the gland, allowing quick separation of the gland from the rest of the trichome, has been described in tomato type VI glandular trichomes (Bergau et al., 2015).

In addition to the wide inter-specific diversity, intra-specific morphological diversity of glandular trichomes can also be very high, as illustrated by the distinct morphologies of the four glandular trichome types described in tomato (Glas et al., 2012). The proportion of each trichome type can be different in closely related species of tomato (Glas et al., 2012) and is highly dependent on leaf developmental stage (Vendemiatti et al., 2017). For example, in cultivated tomato, type IV trichomes are present on juvenile leaves and absent in adults leaves (Vendemiatti et al., 2017). These developmental variations being rarely well described, the description of a modification of the proportion of each trichome type is difficult. For example, *woolly* mutant tomato lines have been known for decades as “hairy” mutants of tomato, but the precise phenotype of certain lines in terms of glandular trichome type proportions is controversial (Vendemiatti et al., 2017).

Given their common organization scheme, glandular trichomes must share common key developmental events. A number of studies have attempted to describe the different steps of glandular trichome development in different plant species (Duke and Paul, 1993; Bergau et al., 2015; Huchelmann et al., 2017). Trichome initiation occurs when an epidermal cell acquires a trichome identity according to signals received from surrounding cells. This cell then undergoes tightly controlled cell divisions; the number and the orientation of these divisions and the extent of cell elongation contribute to shape various trichome morphologies. These first developmental steps are common to glandular and non-glandular trichomes. Additionally, in glandular trichomes one or more cells differentiate into gland cells. The acquisition of the secretory activity implies a profound remodeling of cell

ultrastructure (Lange and Turner, 2013) and the activation of specialized metabolism pathways, for example terpenoid biosynthesis in *Artemisia annua* glandular trichomes (Soetaert et al., 2013). All these developmental events are critical for glandular trichome patterning, morphogenesis and differentiation, but their genetic control remains poorly understood (Huchelmann et al., 2017). However, a certain number of genes involved in glandular trichome initiation have been recently characterized in *Artemisia annua* and tomato, along with a few genes involved in glandular trichome morphogenesis. These genes and their respective functions are detailed below.

3. Genes controlling glandular trichome initiation

Several transcription factors involved in glandular trichome initiation have been identified both in tomato and *Artemisia annua*. They belong to various families of transcription factors, but the majority of them belong to two subfamilies: the R2R3-MYB subfamily and the HD-ZIP IV subfamily.

3.1. Transcription factors

3.1.1. R2R3-MYB transcription factors

Transcription factors of the R2R3-MYB subfamily have two DNA-binding MYELOBLASTOSIS-RELATED (MYB) domain repeats. Different members of this subfamily are known to be involved in the regulation of secondary metabolism (Chezem and Clay, 2016) or in the regulation of epidermal cell fate, for example *MIXTA* and *MIXTA*-like genes (Brockington et al., 2013). The first *MIXTA* gene was characterized in snapdragon (*Antirrhinum majus*) and controls the differentiation of conical epidermal cells from flat epidermal cells (Brockington et al., 2013).

In *Artemisia annua* and tomato, three members of the R2R3-MYB subfamily have been characterized as positive regulators of glandular trichome initiation: *AaMYB1* and *AaMIXTA1* in *Artemisia annua* (Matías-Hernández et al., 2017; Shi et al., 2018) and *SIMX1* in tomato (Ewas et al., 2016, 2017) (**Figure 41**). Indeed, *Artemisia annua* plants overexpressing *AaMYB1* show an increase in glandular trichome density (Matías-Hernández et al., 2017). Moreover, down- or upregulation of *AaMIXTA1* in *Artemisia annua* (Shi et al., 2018) and *SIMX1* in tomato (Ewas et al., 2016, 2017) respectively decreases and increases glandular

trichome density. *AaMIXTA1* and *SIMX1* are both *MIXTA*-like genes (Ewas et al., 2016; Shi et al., 2018) whereas *AaMYB1* belongs to another clade of the R2R3-MYB subfamily (Matías-Hernández et al., 2017).

3.1.2. HD-ZIP IV transcription factors

HD-ZIP transcription factors are characterized by a homeodomain (HD) DNA-binding domain and a leucine-zipper (ZIP) motif responsible for protein dimerization (Gao et al., 2015; Chen et al., 2017a). Members of the HD-ZIP IV subfamily of transcription factors also have a START domain and are known to be involved in epidermal cell differentiation in plants, including cuticle biosynthesis and patterning of trichomes and stomata (Yang et al., 2011; Nadakuduti et al., 2012).

In *Artemisia annua*, two HD-ZIP IV transcription factors, namely *AaHD1* and *AaHD8*, have recently been shown to positively regulate glandular trichome initiation (**Figure 41**). Overexpression of *AaHD1* (Yan et al., 2017) or *AaHD8* (Yan et al., 2018) increases glandular trichome density, whereas downregulation of any of the two genes has the opposite effect. *AaHD8* acts upstream of *AaHD1* by directly promoting its expression (Yan et al., 2018).

The closest homolog of *AaHD8* in tomato is *CUTIN DEFICIENT 2 (SICD2)* (Yan et al., 2018). A loss-of-function mutation in *SICD2* is responsible for the phenotype of the *sticky peel* mutant of tomato, which displays a lower number of glandular trichomes (especially type VI) (Nadakuduti et al., 2012). Therefore, the function of *AaHD8/SICD2* in positive regulation of glandular trichome initiation seems to be conserved between tomato and *Artemisia annua*. Another HD-ZIP IV transcription factor, WOOLLY (*Wo*), appears to be an important regulator of glandular trichome initiation in tomato (**Figure 41**). Dominant point mutations in the C-terminus part of *Wo* are responsible for the phenotype of *woolly* mutants, which show dramatically increased trichome density (Yang et al., 2011, 2015). According to the first characterization of *Wo*, type I glandular trichome density is increased in plants carrying dominant *woolly* mutations and reduced in *Wo*-RNAi plants, suggesting that *Wo* enhances type I glandular trichome initiation (Yang et al., 2011). However, a recent re-analysis of the phenotype of *woolly* mutants indicates instead that *woolly* mutants show a higher density of type III and type V non-glandular trichomes and a lower density of type IV glandular trichomes in adult leaves (Vendemiatti et al., 2017). According to this study, the effect of the

dominant *woolly* point mutation is different depending on leaf developmental stage: indeed, a higher density of type IV glandular trichomes was observed in juvenile leaves of *woolly* mutants, whereas it was lower in adult leaves compared to the wild-type (Vendemiatti et al., 2017).

3.1.3. Other transcription factors

Two other transcription factors involved in glandular trichome initiation and belonging neither to the R2R3-MYB subfamily, nor to the HD-ZIP IV subfamily, have also been recently characterized in tomato: the C2H2 zinc-finger protein HAIR (*SIH*) (Chang et al., 2018) and the bHLH protein MYELOCYTOMATOSIS-RELATED 1 (*SIMYC1*) (Xu et al., 2018) (**Figure 41**). Downregulation of *SIMYC1* by RNAi or missense mutations in *SIH* reduce type VI or type I glandular trichome density, respectively. Moreover, type VI glandular trichomes are absent in *myc1* knockout mutants and *SIH* knockout leads to a hair-absent phenotype (Xu et al., 2018; Chang et al., 2018). These results indicate that *SIMYC1* positively regulates the initiation of type VI glandular trichomes and that *SIH* is a key positive regulator of all glandular trichome types. Interestingly, type VI glandular trichomes have smaller glands and shorter stalks in *SIMYC1*-RNAi plants, suggesting that *SIMYC1* is also an important regulator of later steps of type VI glandular trichome morphogenesis in tomato (Xu et al., 2018).

3.2. Cyclins

The induction of cell divisions in early steps of glandular trichome development requires not only transcription factors, but also cell cycle regulators like cyclins. The tomato gene *SIcycB2* encodes a B-type cyclin, which is a type of cyclin promoting the G2/M transition (Gao et al., 2017). Type I glandular trichome density is reduced in *SIcycB2*-RNAi plants, highlighting the role of *SIcycB2* in glandular trichome initiation (Yang et al., 2011; Gao et al., 2017) (**Figure 41**). *SIcycB2* may promote a shift from endoreduplication to mitosis in epidermal cells, thereby inducing the first cell divisions of type I glandular trichome development (Yang et al., 2011). However, *SIcycB2* overexpression does not seem to be an efficient tool to increase glandular trichome density because it inhibits the initiation of type I and type VI glandular trichomes (Gao et al., 2017). *SIcycB2* expression is upregulated in *Wo*-overexpressing plants,

downregulated in *Wo*-RNAi plants and upregulated in *SIMX1*-overexpressing plants, suggesting that *S/CycB2* expression may be positively regulated by *Wo* and *SIMX1* (Ewas et al., 2016; Yang et al., 2011) (Figure 41).

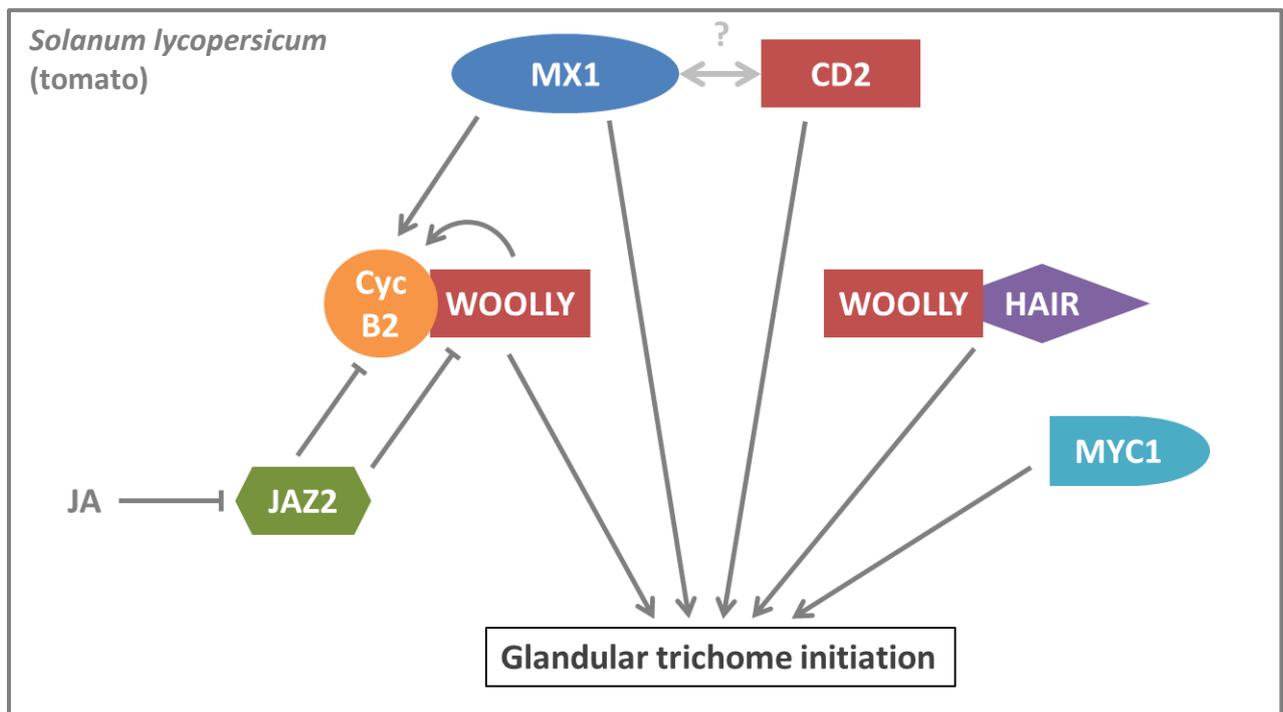
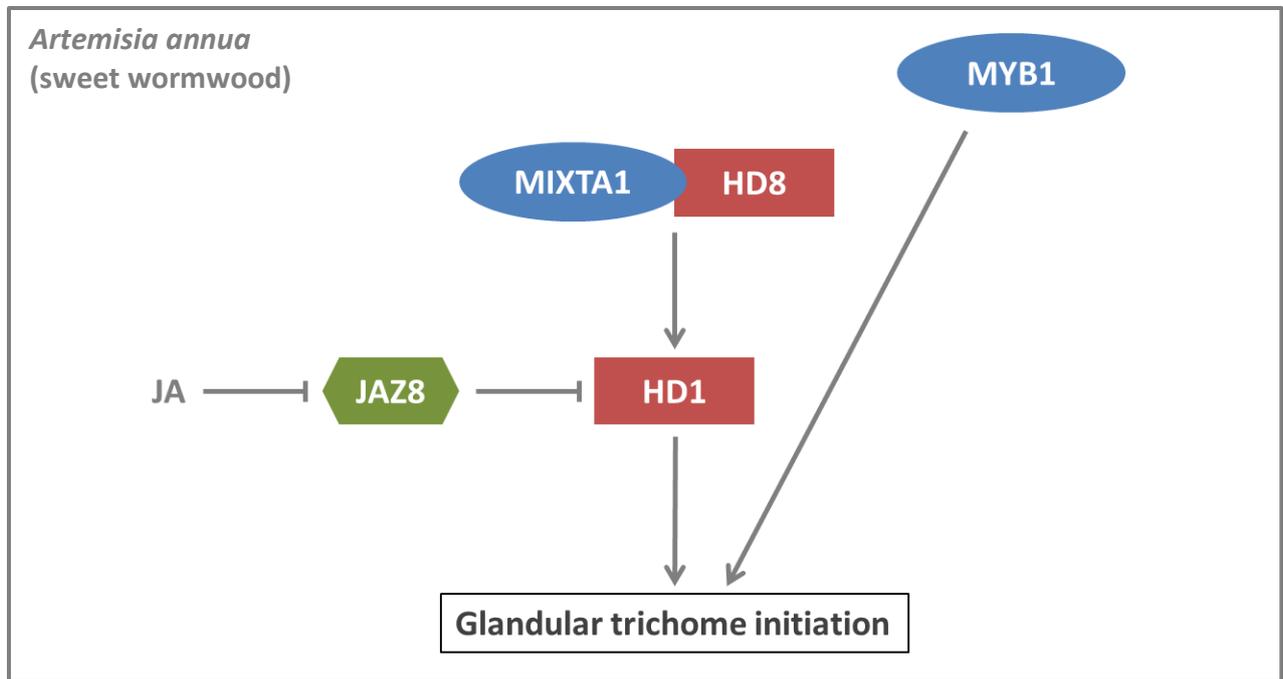


Figure 41: Simplified model of glandular trichome initiation in sweet wormwood (*Artemisia annua*) and tomato (*Solanum lycopersicum*).

In *Artemisia annua*, glandular trichome initiation is positively regulated by the HD-ZIP IV transcription factor **HD1**. The expression of *HD1* is enhanced by a complex formed by the R2R3-MYB transcription factor **MIXTA1** and the HD-ZIP IV transcription factor **HD8**. The R2R3-MYB transcription factor **MYB1** is another positive regulator of glandular trichome initiation. Conversely, the JA signaling repressor **JAZ8** represses HD1 transcriptional activity, thereby inhibiting glandular trichome initiation.

In *Solanum lycopersicum*, glandular trichome initiation is positively regulated by the HD-ZIP IV transcription factor **WOOLLY**. WOOLLY has been shown to interact with the B-type cyclin **CycB2** and with the C2H2 zinc-finger protein **HAIR**, which is itself a positive regulator of glandular trichome initiation. A HAIR-WOOLLY-CycB2 complex may exist, but has not been detected yet. The bHLH transcription factor **MYC1**, the R2R3-MYB transcription factor **MX1** and the HD-ZIP IV transcription factor **CD2** are 3 other positive regulators of glandular trichome initiation. CD2 is the closest tomato homolog of the transcription factor HD8 of *Artemisia annua*. Like HD8, CD2 may interact with a MIXTA-like protein, maybe MX1. MX1 and WOOLLY both induce the expression of *CycB2*. Conversely, the JA signaling repressor **JAZ2** inhibits WOOLLY and *CycB2* expression, thereby inhibiting glandular trichome initiation.

3.3. Regulatory complexes

Non-glandular trichome initiation in *Arabidopsis thaliana* is known to be controlled by a MYB-bHLH-WD40 complex (Pattanaik et al., 2014). An important output of recent efforts in dissecting the genetic network underlying glandular trichome development was the identification of 2 complexes controlling glandular trichome initiation in tomato and *Artemisia annua*. In *Artemisia annua*, the expression of the positive regulator of glandular trichome initiation *AaHD1* was recently shown to be enhanced by a complex formed by two transcription factors: the HD-ZIP IV protein AaHD8 and the R2R3-MYB MIXTA-like protein AaMIXTA1 (Yan et al., 2018) (**Figure 41**). In tomato, the closest homolog of AaHD8 is SICD2 (Yan et al., 2018). SICD2 may interact with a MIXTA-like transcription factor, maybe SIMX1, to form a complex similar to AaHD8-AaMIXTA1, but this remains to be demonstrated. A direct interaction between the C2H2 zinc-finger protein SIH and the HD-ZIP IV transcription factor Wo was recently detected, suggesting that SIH and Wo act as a heterodimer to induce type I glandular trichome formation (Chang et al., 2018) (**Figure 41**). Moreover, Wo and the B-type cyclin SICycB2 physically interact, supporting the hypothesis that these two proteins also act together to induce type I glandular trichome formation (Yang et al., 2011) (**Figure 41**). An H-Wo-CycB2 complex may be involved in the positive regulation of type I glandular trichome initiation in tomato, but has not been detected yet (Chang et al., 2018).

3.4. Genes involved in hormonal signaling

Glandular trichome initiation is known to be regulated by plant hormones in various plant species (Maes and Goossens, 2010). In particular, jasmonates (JA) are able to induce glandular trichome initiation in tomato (Maes and Goossens, 2010) and *Artemisia annua* (Maes et al., 2011). Consistently, proteins involved in JA signaling have recently been shown to impact glandular trichome initiation in both species. In *Artemisia annua*, the transcriptional activity of the positive regulator of glandular trichome initiation AaHD1 is repressed by a direct interaction with AaJAZ8, which belongs to the JAZ family of JA signaling repressors (Yan et al., 2017) (**Figure 41**). In tomato, the overexpression of the JA signaling repressor *SIJAZ2* decreases glandular trichome density, indicating that *SIJAZ2* is a negative regulator of glandular trichome initiation (Yu et al., 2018). The expression of *Wo* and *SICycB2* is strongly repressed in plants overexpressing *SIJAZ2*, suggesting that *SIJAZ2* inhibits glandular trichome development by downregulating the positive regulators *Wo* and *SICycB2* (Yu et al., 2018) (**Figure 41**). The induction of glandular trichome initiation by JA in *Artemisia annua* and tomato probably proceeds through the JA-triggered degradation of the repressors AaJAZ8 and *SIJAZ2* by the proteasome, respectively. Consistently, *SIJAZ2* was shown to directly interact with the F-box protein *SICOI1* in a yeast two-hybrid screen (Thines et al., 2007) and *SICOI* is itself a positive regulator of glandular trichome development (Li et al., 2004). In addition to jasmonate signaling, other hormonal signaling pathways control glandular trichome development. For instance, two genes involved in auxin signaling have been shown to impact glandular trichome development: *SIARF3* (Zhang et al., 2015b) and *SIIAA15* (Deng et al., 2012).

4. Genes controlling later steps of glandular trichome development

4.1. Cytoskeleton regulators

Following the initiation step, glandular trichome morphogenesis starts. Various cellular components are at play to define glandular trichome shape, which is essential for its correct functioning. Indeed, the alteration of type VI glandular trichome morphology caused by the *hairless* mutation in tomato leads to impaired synthesis of defense metabolites and decreased resistance to herbivores (Kang et al., 2016). The mutation responsible for the observed bending and swelling of type VI glandular trichomes has been located in the gene

encoding the SRA1 subunit of the WAVE regulatory complex (Kang et al., 2016). This complex is highly conserved among eukaryotes and controls actin filaments nucleation and polymerization (Kang et al., 2016). Therefore, actin cytoskeleton remodeling seems to play a critical role in glandular trichome morphogenesis. This hypothesis is reinforced by the recent analysis of the tomato *inquieta* mutant (Jeong et al., 2017). Glandular trichomes of this mutant display similar morphological defects as glandular trichomes of the *hairless* mutant. This phenotype has been associated with a mutation in the homolog of the *ARPC2A* gene of *Arabidopsis thaliana*, which is another important actor of actin cytoskeleton polymerization (Jeong et al., 2017).

4.2. Cuticle deposition regulators

Several studies have highlighted a tight link between cuticle deposition and non-glandular trichome development in *Arabidopsis thaliana*, with many genes involved in both processes (Shi et al., 2018). Similarly, the correct accumulation of cuticle may be crucial for glandular trichome development in tomato and *Artemisia annua*. In *Artemisia annua*, downregulation of the AP2/ERF transcription factor gene *AaTAR1* by RNAi leads to an altered cuticular wax deposition and an increase in cuticle permeability (Tan et al., 2015). Interestingly, glandular trichomes of *AaTAR1*-RNAi plants have an abnormal morphology: the top of the gland is swollen and gland cell number is reduced (Tan et al., 2015). In tomato, downregulation of the R2R3-MYB transcription factor gene *SIMX1* by RNAi decreases cuticle deposition along with trichome density, whereas the opposite is observed in lines overexpressing *SIMX1* (Ewas et al., 2016, 2017). Likewise, the *sticky peel* mutant, which carries a mutation in the HD-ZIP IV transcription factor gene *SICD2*, is impaired in cutin accumulation and displays a lower glandular trichome density at the same time (Nadakuduti et al., 2012). These examples highlight a link between cuticle formation and glandular trichome initiation in tomato and *Artemisia annua*, but it is unclear whether cuticle deposition is necessary for proper glandular trichome morphogenesis, or whether these two processes are simply co-regulated. The analysis of glandular trichome morphology in plants harboring mutation in cuticle biosynthesis genes could help to answer this question.

5. Evolution of glandular trichome development regulators

5.1. Evidence supporting the conservation of glandular trichome development regulators in the Solanaceae

Several lines of evidence suggest that multicellular trichome formation, including glandular ones, may be controlled by the same regulatory network in all Solanaceae. Indeed, ectopic expression of *Wo*^v, a strong allele of *Wo*, is able to induce multicellular trichome formation in tobacco and potato, and tobacco homologs of *Wo* and *SlCycB2* were upregulated in *Wo*^v-overexpressing tobacco plants (Yang et al., 2015). The ectopic expression of *SlH* in tobacco also triggers trichome formation (Chang et al., 2018). Conversely, ectopic expression of pepper or tobacco orthologues of *SlH* in tomato plants induces trichome formation (Chang et al., 2018). Moreover, the overexpression of the tobacco gene *NtCycB2* in tomato and tobacco led to a phenotype comparable to the overexpression of *SlCycB2* in both species (Gao et al., 2017). Taken together, these results support the idea that the function of at least *H* and *CycB2*, and probably *Wo*, is conserved among Solanaceae species.

5.2. Some regulators appear to be conserved in distant plant families, while the others are likely to have evolved independently

In *Arabidopsis thaliana*, non-glandular trichome development also involves transcription factors belonging to R2R3-MYB and HD-ZIP IV subfamilies. The R2R3-MYB transcription factor GLABRA 1 (*AtGL1*) interacts with bHLH and WD40 proteins to form a MYB-bHLH-WD40 complex (Pattanaik et al., 2014). This complex is able to induce the expression of the HD-ZIP IV transcription factor GLABRA 2 (*AtGL2*), which positively regulates non-glandular trichome initiation. Besides, single repeat R3-MYB transcription factors repress non-glandular trichome initiation (Pattanaik et al., 2014).

The positive regulator of glandular trichome development *AaMYB1* increases *AtGL1* and *AtGL2* expression and induces non-glandular trichome initiation when ectopically expressed in *Arabidopsis thaliana* (Matías-Hernández et al., 2017). The orthologue of *AaMYB1* in *Arabidopsis thaliana* is *AtMYB61*. Interestingly, non-glandular trichome density is reduced in *myb61* mutants, indicating that *AtMYB61* positively regulates non-glandular trichome initiation in *Arabidopsis thaliana* (Matías-Hernández et al., 2017). Another study shows that a functional orthologue of the R3-MYB gene *AtTRY* may be present in tomato: indeed, *SlTRY*

is able to inhibit trichome initiation when expressed in *Arabidopsis thaliana* (Tominaga-Wada et al., 2013). Taken together, these results suggest that several genes may have a conserved function in the regulation of trichome initiation in *Arabidopsis thaliana*, *Artemisia annua* and *Solanum lycopersicum*.

However, other studies suggest that the network controlling non-glandular trichome initiation in *Arabidopsis thaliana* is different from the one controlling glandular trichome initiation in the Solanaceae. *AtGL1* does not have any impact on glandular trichome development when expressed in tobacco, and the expression of *AmMIXTA*, a gene of *Antirrhinum majus* closely related to *AaMIXTA1* and *SIMX1*, could not rescue the phenotype of the *gl1* mutant of *Arabidopsis thaliana* (Yang and Ye, 2013). Moreover, although *AtGL2* and *Wo* are both HD-ZIP IV transcription factors, the closest homolog of *Wo* in *Arabidopsis thaliana* is not *AtGL2* but *PROTODERMAL FACTOR 2 (AtPDF2)*, a gene involved in shoot epidermal cell differentiation but not in trichome initiation (Yang and Ye, 2013; Yang et al., 2011). Some regulators involved in early steps of trichome development may be shared among a wide range of dicots, whereas other regulators are likely to have evolved several times independently among dicots and monocots. This statement raises the following question: which regulatory networks could have served as starting material for the evolution of trichome development regulators?

5.3. Different scenarios for the evolution of glandular trichome development regulators

MIXTA-like and HD-ZIP IV transcription factors are known to be key regulators of epidermal cell fate in different plant species (Brockington et al., 2013; Takada and Iida, 2014; Yan et al., 2018). For example, several HD-ZIP IV transcription factors are known to be involved in cuticle biosynthesis and stomata patterning (Nadakuduti et al., 2012). Glandular trichomes being specialized epidermal structures, some genes initially involved in the differentiation of other epidermal structures may have acquired a function in glandular trichome development during the course of evolution. In agreement with this hypothesis, several regulators of glandular trichome development characterized in *Artemisia annua* and tomato also have additional functions related to epidermal cell differentiation. *AaMYB1* and *AtMYB61* have been shown to impact not only trichome density, but also stomatal aperture in *Arabidopsis*

thaliana (Matías-Hernández et al., 2017). Similarly, *SICD2* and *SIMX1* in tomato and the AaHD8-AaMIXTA1 complex in *Artemisia annua* seem to be involved in cuticle development in addition to glandular trichome initiation (Nadakuduti et al., 2012; Ewas et al., 2016, 2017; Yan et al., 2018).

The main function of glandular trichomes being the production of specialized metabolites, some genes initially involved in the regulation of specialized metabolism could also have acquired a function in glandular trichome development. Some members of the R2R3-MYB subfamily are specialized metabolism regulators (Chezem and Clay, 2016) and several bHLH transcription factors have been reported to activate terpene biosynthesis in various plant species (Xu et al., 2018). In agreement with this hypothesis, the bHLH transcription factor SIMYC1 differentially regulates mono- and sesquiterpene biosynthesis in leaves and stems of cultivated tomato in addition to its role in glandular trichome initiation (Xu et al., 2018). Lastly, since the production of specialized metabolites by glandular trichomes is a part of plant response to biotic and abiotic stresses, genes initially involved in stress response may have acquired a function in glandular trichome development. Consistently, JAZ repressors have been shown to repress glandular trichome initiation, linking this developmental process to JA signaling, a well-known stress signaling pathway (Yan et al., 2017; Yu et al., 2018).

6. Perspectives

Almost all genes recently shown to be involved in glandular trichome development in tomato and *Artemisia annua* also impact the development of non-glandular trichomes (Deng et al., 2012; Tan et al., 2015; Ewas et al., 2016; Kang et al., 2016; Yan et al., 2017; Matías-Hernández et al., 2017; Vendemiatti et al., 2017; Gao et al., 2017; Shi et al., 2018; Yan et al., 2018; Chang et al., 2018). *SIMYC1* is the only gene identified so far which seems to affect only glandular trichome development (Xu et al., 2018). It would be of high interest to identify other genes specifically controlling glandular trichome development, in particular genes involved in the acquisition of the secretory activity. In tomato, such genes may be found among the targets of SIMYC1, which remain to be identified. Alternatively, a comparative study of type IV and type V trichomes could lead to insights in the genetic regulation of glandular cell differentiation, because these two trichome types are morphologically very similar, except for the apical cell which is glandular in type IV and non-

glandular in type V trichomes (Glas et al., 2012). In *Artemisia annua*, given the fact that HD-ZIP transcription factors act only as dimers, AaHD1 may interact with distinct HD-ZIP transcription factors to induce glandular or non-glandular trichome initiation. Therefore, the identification of AaHD1 interactors may lead to the identification of regulators specific to glandular trichome initiation (Yan et al., 2017).

In tomato, the elucidation of the regulatory network underlying the initiation and differential morphogenesis of all trichome types may be challenging, but this diversity of trichome types could be highly valuable for the study of the fine regulation of trichome differentiation. Laser microdissection could be a powerful tool for comparative analyses of glandular trichome types. It is not easy to isolate protruding organs with this technique. However, it has already been done successfully on *Artemisia annua* to compare glandular and non-glandular trichomes (Soetaert et al., 2013). Besides, the recent identification of quantitative trait loci (QTLs) controlling the shape of type VI glandular trichomes in tomato represents a precious information for the future characterization of more genetic regulators of glandular trichome morphogenesis (Bennewitz et al., 2018).

A better understanding of glandular trichome development will open exciting avenues for the targeted improvement of agronomical traits. For example, tomato lines with more type IV trichomes or bigger type VI trichome secretory cavities could produce more acylsugars and terpenes, respectively, and thus show better resistance to herbivores (Vendemiatti et al., 2017; Bennewitz et al., 2018). The yield of high value-added compounds produced in plant glandular trichomes could also be increased, with benefits for perfume and pharmaceutical industries. Results obtained on *Artemisia annua*, which produces the anti-malarial drug artemisinin, are encouraging. The overexpression of *AaMIXTA1*, *AaHD1* or *AaHD8* significantly enhanced artemisinin production, without any adverse effect on plant growth and fitness (Yan et al., 2017; Shi et al., 2018; Yan et al., 2018). The highest increase was observed with the overexpression of *AaMIXTA1*, which doubled artemisinin content (Shi et al., 2018). Knowing to what extent gene networks controlling glandular trichome development are conserved among the plant kingdom will be critical to develop plant breeding strategies based on glandular trichome phenotype in other plant species. Targeted mutagenesis approaches like CRISPR-Cas9 or TILLING could be used to investigate whether the functions of already identified genes are conserved or not, and to characterize new regulators and actors of glandular trichome development.

References

- Alfieri, M., Vaccaro, M.C., Cappetta, E., Ambrosone, A., De Tommasi, N., and Leone, A.** (2018). Coactivation of MEP-biosynthetic genes and accumulation of abietane diterpenes in *Salvia sclarea* by heterologous expression of WRKY and MYC2 transcription factors. *Sci. Rep.* **8**: 11009.
- Ample Market Research** (2019). Global Sclareol Market 2019 by Manufacturers, Regions, Type and Application, Forecast to 2024.
- Balcke, G.U., Bennewitz, S., Bergau, N., Athmer, B., Henning, A., Majovsky, P., Jiménez-Gómez, J.M., Hoehenwarter, W., and Tissier, A.** (2017). Multi-Omics of Tomato Glandular Trichomes Reveals Distinct Features of Central Carbon Metabolism Supporting High Productivity of Specialized Metabolites. *Plant Cell* **29**: 960–983.
- Banthorpe, D. V., Brown, J.T., and Morris, G.S.** (1990). Accumulation of the anti-fungal diterpene sclareol by cell cultures of *Salvia sclarea* and *Nicotiana glutinosa*. *Phytochemistry* **29**: 2145–2148.
- Barrero, A.F., Altarejos, J., Alvarez-Manzaneda, E.J., Ramos, J.M., and Salido, S.** (1996). Synthesis of (±)-Ambrox from (E)-Nerolidol and β-Ionone via Allylic Alcohol [2, 3] Sigmatropic Rearrangement. *J. Org. Chem.* **61**: 2215–2218.
- Barrero, A.F., Alvarez-Manzaneda, E.J., Altarejos, J., Salido, S., and Ramos, J.M.** (1993). Synthesis of Ambrox® from (–)-sclareol and (+)-cis-abienol. *Tetrahedron* **49**: 10405–10412.
- Barroso, J.G., Pedro, L.G., Figueiredo, A.C., Pais, M.S.S., and Scheffer, J.J.C.** (1992). Seasonal Variation in the Composition of the Essential Oil of *Crithmum maritimum* L. *Flavour Fragr. J.* **7**: 147–150.
- Bathe, U. and Tissier, A.** (2019). Cytochrome P450 enzymes: A driving force of plant diterpene diversity. *Phytochemistry* **161**: 149–162.
- Battilana, J., Emanuelli, F., Gambino, G., Gribaudo, I., Gasperi, F., Boss, P.K., and Grando, M.S.** (2011). Functional effect of grapevine 1-deoxy-D-xylulose 5-phosphate synthase substitution K284N on Muscat flavour formation. *J. Exp. Bot.* **62**: 5497–5508.
- Behnke, K., Ehltng, B., Teuber, M., Bauerfeind, M., Louis, S., Hänsch, R., Polle, A., Bohlmann, J., and Schnitzler, J.-P.** (2007). Transgenic, non-isoprene emitting poplars don't like it hot. *Plant J.* **51**: 485–499.
- Behnke, K., Kleist, E., Uerlings, R., Wildt, J., Rennenberg, H., and Schnitzler, J.-P.** (2009). RNAi-mediated suppression of isoprene biosynthesis in hybrid poplar impacts ozone tolerance. *Tree Physiol.* **29**: 725–736.
- Bennewitz, S., Bergau, N., and Tissier, A.** (2018). QTL Mapping of the Shape of Type VI Glandular Trichomes in Tomato. *Front. Plant Sci.* **9**: 1421.
- Bergau, N., Bennewitz, S., Syrowatka, F., Hause, G., and Tissier, A.** (2015). The development of type VI glandular trichomes in the cultivated tomato *Solanum lycopersicum* and a related wild species *S. habrochaites*. *BMC Plant Biol.* **15**: 289.
- Bick, J.A. and Lange, B.M.** (2003). Metabolic cross talk between cytosolic and plastidial pathways of isoprenoid biosynthesis: Unidirectional transport of intermediates across the chloroplast envelope membrane. *Arch. Biochem. Biophys.* **415**: 146–154.
- Boughton, A.J., Hoover, K., and Felton, G.W.** (2005). Methyl Jasmonate Application Induces Increased Densities of Glandular Trichomes on Tomato, *Lycopersicon esculentum*. *J. Chem. Ecol.* **31**: 2211–2216.

- Brockington, S.F., Alvarez-Fernandez, R., Landis, J.B., Alcorn, K., Walker, R.H., Thomas, M.M., Hileman, L.C., and Glover, B.J.** (2013). Evolutionary Analysis of the MIXTA Gene Family Highlights Potential Targets for the Study of Cellular Differentiation. *Mol. Biol. Evol.* **30**: 526–540.
- Brückner, K. and Tissier, A.** (2013). High-level diterpene production by transient expression in *Nicotiana benthamiana*. *Plant Methods* **9**: 1–10.
- Van Den Brûle, S., Müller, A., Fleming, A.J., and Smart, C.C.** (2002). The ABC transporter SpTUR2 confers resistance to the antifungal diterpene sclareol. *Plant J.* **30**: 649–662.
- Burke, C. and Croteau, R.** (2002). Interaction with the small subunit of geranyl diphosphate synthase modifies the chain length specificity of geranylgeranyl diphosphate synthase to produce geranyl diphosphate. *J. Biol. Chem.* **277**: 3141–9.
- Burke, C.C., Wildung, M.R., and Croteau, R.** (1999). Geranyl diphosphate synthase: cloning, expression, and characterization of this prenyltransferase as a heterodimer. *Proc. Natl. Acad. Sci. U. S. A.* **96**: 13062–7.
- Caissard, J.C. et al.** (2012). Extracellular Localization of the Diterpene Sclareol in Clary Sage (*Salvia sclarea* L., Lamiaceae). *PLoS One* **7**: 1–8.
- Caniard, A., Zerbe, P., Legrand, S., Cohade, A., Valot, N., Magnard, J.-L., Bohlmann, J., and Legendre, L.** (2012). Discovery and functional characterization of two diterpene synthases for sclareol biosynthesis in *Salvia sclarea* (L.) and their relevance for perfume manufacture. *BMC Plant Biol.* **12**: 119.
- Carvalho, F., Duarte, J., Simoes, A.R., Cruz, P.F., and Jones, J.G.** (2013). Noninvasive measurement of murine hepatic acetyl-CoA ¹³C-enrichment following overnight feeding with ¹³C-enriched fructose and glucose. *Biomed Res. Int.* **2013**: 638085.
- Chain, F.E., Leyton, P., Paipa, C., Fortuna, M., and Brandán, S.A.** (2015). FT-IR, FT-Raman, UV-Visible, and NMR spectroscopy and vibrational properties of the labdane-type diterpene 13-epi-sclareol. *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **138**: 303–313.
- Chang, J., Yu, T., Yang, Q., Li, C., Xiong, C., Gao, S., Xie, Q., Zheng, F., Li, H., Tian, Z., Yang, C., and Ye, Z.** (2018). Hair , encoding a single C2H2 zinc-finger protein, regulates multicellular trichome formation in tomato. *Plant J.* **96**: 90–102.
- Chang, S.Y. and Grunwald, C.** (1980). Structural Organization of Tobacco Leaf Polar Cuticular Lipids. *Bot. Gaz.* **141**: 360–365.
- Chen, E. et al.** (2017a). Genome-wide analysis of the HD-ZIP IV transcription factor family in *Gossypium arboreum* and GaHDG11 involved in osmotic tolerance in transgenic *Arabidopsis*. *Mol. Genet. Genomics* **292**: 593–609.
- Chen, F., Tholl, D., Bohlmann, J., and Pichersky, E.** (2011). The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J.* **66**: 212–229.
- Chen, G., Klinkhamer, P.G.L., Escobar-Bravo, R., and Leiss, K.A.** (2018). Type VI glandular trichome density and their derived volatiles are differently induced by jasmonic acid in developing and fully developed tomato leaves: Implications for thrips resistance. *Plant Sci.* **276**: 87–98.
- Chen, H., Jones, A.D., and Howe, G.A.** (2006). Constitutive activation of the jasmonate signaling pathway enhances the production of secondary metabolites in tomato. *FEBS Lett.* **580**: 2540–2546.
- Chen, M. et al.** (2017b). GLANDULAR TRICHOME-SPECIFIC WRKY 1 promotes artemisinin biosynthesis in *Artemisia annua*. *New Phytol.* **214**: 304–316.

- Chezem, W.R. and Clay, N.K.** (2016). Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs. *Phytochemistry* **131**: 26–43.
- Claßen-Bockhoff, R., Speck, T., Tweraser, E., Wester, P., Thimm, S., and Reith, M.** (2004). The staminal lever mechanism in *Salvia L.* (Lamiaceae): A key innovation for adaptive radiation? *Org. Divers. Evol.* **4**: 189–205.
- Colinas, M. and Goossens, A.** (2018). Combinatorial Transcriptional Control of Plant Specialized Metabolism. *Trends Plant Sci.* **23**: 324–336.
- Courdavault, V., Thiersault, M., Courtois, M., Gantet, P., Oudin, A., Doireau, P., St-Pierre, B., and Giglioli-Guivarc’h, N.** (2005). CaaX-prenyltransferases are essential for expression of genes involved in the early stages of monoterpene biosynthetic pathway in *Catharanthus roseus* cells. *Plant Mol. Biol.* **57**: 855–870.
- Craft, J.D., Satyal, P., Setzer, W.N., Craft, J.D., Satyal, P., and Setzer, W.N.** (2017). The Chemotaxonomy of Common Sage (*Salvia officinalis*) Based on the Volatile Constituents. *Medicines* **4**: 47.
- Çulhaoglu, B., Yapara, G., Dirmencic, T., and Topçu, G.** (2013). Bioactive constituents of *Salvia chrysophylla* Stapf. *Nat. Prod. Res.* **27**: 438–447.
- Dalla Costa, L. et al.** (2018). Induction of Terpene Biosynthesis in Berries of Microvine Transformed with VvDXS1 Alleles. *Front. Plant Sci.* **8**: 2244.
- Deng, W., Yang, Y., Ren, Z., Audran-Delalande, C., Mila, I., Wang, X., Song, H., Hu, Y., Bouzayen, M., and Li, Z.** (2012). The tomato SHAA15 is involved in trichome formation and axillary shoot development. *New Phytol.* **194**: 379–390.
- Dodson, C.H., Dressler, R.L., Hills, H.G., Adams, R.M., and Williams, N.H.** (1969). Biologically active compounds in orchid fragrances. *Science* **164**: 1243–9.
- Dong, A.-X. et al.** (2018). High-quality assembly of the reference genome for scarlet sage, *Salvia splendens*, an economically important ornamental plant. *Gigascience* **7**: giy068.
- Dudareva, N., Andersson, S., Orlova, I., Gatto, N., Reichelt, M., Rhodes, D., Boland, W., and Gershenzon, J.** (2005). The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers. *Proc. Natl. Acad. Sci.* **102**: 933–938.
- Dufour, C., Crampon, C., Delbecque, C., Garry, P.P., and Badens, E.** (2017). Purification of sclareol by supercritical CO₂ fractionation process. *J. Supercrit. Fluids* **122**: 35–42.
- Duke, S.O. and Paul, R.N.** (1993). Development and Fine Structure of the Glandular Trichomes of *Artemisia annua L.* *Int. J. Plant Sci.* **154**: 107–118.
- Egea, C., Pérez, M.D.G., and Candela, M.E.** (1996). Capsidiol accumulation in *capsicum annum* stems during the hypersensitive reaction to *Phytophthora capsici*. *J. Plant Physiol.* **149**: 762–764.
- Ewas, M., Gao, Y., Ali, F., Nishawy, E.M., Shahzad, R., Subthain, H., Amar, M., Martin, C., and Luo, J.** (2017). RNA-seq reveals mechanisms of SIMX1 for enhanced carotenoids and terpenoids accumulation along with stress resistance in tomato. *Sci. Bull.* **62**: 476–485.
- Ewas, M., Gao, Y., Wang, S., Liu, X., Zhang, H., Nishawy, E.M.E., Ali, F., Shahzad, R., Ziaf, K., Subthain, H., Martin, C., and Luo, J.** (2016). Manipulation of SIMX1 for enhanced carotenoids accumulation and drought resistance in tomato. *Sci. Bull.* **61**: 1413–1418.
- Fahn, A.** (2000). Structure and Function of Secretory Cells. *Adv. Bot. Res.* **31**: 37–75.
- Figueiredo, A.C., Barroso, J.G., Pedro, L.G., and Scheffer, J.J.C.** (2008). Factors affecting secondary metabolite production in plants: volatile components and essential oils. *Flavour Fragr. J.* **23**: 213–226.

- Foutami, I.J., Mariager, T., Rinnan, R., Barnes, C.J., and Rønsted, N.** (2018). Hundred Fifty Years of Herbarium Collections Provide a Reliable Resource of Volatile Terpenoid Profiles Showing Strong Species Effect in Four Medicinal Species of *Salvia* Across the Mediterranean. *Front. Plant Sci.* **9**: 1877.
- Fowler, D.J., Hamilton, J.T.G., Humphrey, A.J., and O'Hagan, D.** (1999). Plant terpene biosynthesis. The biosynthesis of linalyl acetate in *Mentha citrata*. *Tetrahedron Lett.* **40**: 3803–3806.
- Fujimoto, T., Mizukubo, T., Abe, H., and Seo, S.** (2015). Sclareol Induces Plant Resistance to Root-Knot Nematode Partially Through Ethylene-Dependent Enhancement of Lignin Accumulation. *Mol. Plant-Microbe Interact.* **28**: 398–407.
- Gao, S., Gao, Y., Xiong, C., Yu, G., Chang, J., Yang, Q., Yang, C., and Ye, Z.** (2017). The tomato B-type cyclin gene, *SlCycB2*, plays key roles in reproductive organ development, trichome initiation, terpenoids biosynthesis and *Prodenia litura* defense. *Plant Sci.* **262**: 103–114.
- Gao, Y., Gao, S., Xiong, C., Yu, G., Chang, J., Ye, Z., and Yang, C.** (2015). Comprehensive analysis and expression profile of the homeodomain leucine zipper IV transcription factor family in tomato. *Plant Physiol. Biochem.* **96**: 141–153.
- Glas, J.J., Schimmel, B.C.J., Alba, J.M., Escobar-Bravo, R., Schuurink, R.C., and Kant, M.R.** (2012). Plant glandular trichomes as targets for breeding or engineering of resistance to herbivores. *Int. J. Mol. Sci.* **13**: 17077–17103.
- Goossens, J., Fernández-Calvo, P., Schweizer, F., and Goossens, A.** (2016). Jasmonates: signal transduction components and their roles in environmental stress responses. *Plant Mol. Biol.* **91**: 673–689.
- Gul, S., Ahmad, M., Zafar, M., Bahadur, S., Sultana, S., Ashfaq, S., Ullah, F., Kilic, O., Hassan, F. ul, and Siddiq, Z.** (2019). Foliar epidermal anatomy of Lamiaceae with special emphasis on their trichomes diversity using scanning electron microscopy. *Microsc. Res. Tech.* **82**: 206–223.
- Günnewich, N., Higashi, Y., Feng, X., Choi, K.B., Schmidt, J., and Kutchan, T.M.** (2013). A diterpene synthase from the clary sage *Salvia sclarea* catalyzes the cyclization of geranylgeranyl diphosphate to (8R)-hydroxy-copalyl diphosphate. *Phytochemistry* **91**: 93–99.
- Guo, Z. and Wagner, G.** (1995). Biosynthesis of labdenediol and sclareol in cell-free extracts from trichomes of *Nicotiana glutinosa*. *Planta* **197**: 627–632.
- Han, J., Wang, H., Lundgren, A., and Brodelius, P.E.** (2014). Effects of overexpression of *AaWRKY1* on artemisinin biosynthesis in transgenic *Artemisia annua* plants. *Phytochemistry* **102**: 89–96.
- Hao, D.C., Chen, S.L., Osbourn, A., Kontogianni, V.G., Liu, L.W., and Jordán, M.J.** (2015). Temporal transcriptome changes induced by methyl jasmonate in *Salvia sclarea*. *Gene* **558**: 41–53.
- Hausch, B.J., Lorjaroenphon, Y., and Cadwallader, K.R.** (2015). Flavor Chemistry of Lemon-Lime Carbonated Beverages. *J. Agric. Food Chem.* **63**: 112–119.
- He, J., Zhao, X., Laroche, A., Lu, Z.-X., Liu, H., and Li, Z.** (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Front. Plant Sci.* **5**: 484.
- Hemmerlin, A., Harwood, J.L., and Bach, T.J.** (2012). A raison d'être for two distinct pathways in the early steps of plant isoprenoid biosynthesis? *Prog. Lipid Res.* **51**: 95–148.

References

- Hemmerlin, A., Hoeffler, J.F., Meyer, O., Tritsch, D., Kagan, I.A., Grosdemange-Billiard, C., Rohmer, M., and Bach, T.J.** (2003). Cross-talk between the cytosolic mevalonate and the plastidial methylerythritol phosphate pathways in tobacco bright yellow-2 cells. *J. Biol. Chem.* **278**: 26666–26676.
- Hsiao, Y.-Y., Jeng, M.-F., Tsai, W.-C., Chuang, Y.-C., Li, C.-Y., Wu, T.-S., Kuoh, C.-S., Chen, W.-H., and Chen, H.-H.** (2008). A novel homodimeric geranyl diphosphate synthase from the orchid *Phalaenopsis bellina* lacking a DD(X)₂-4D motif. *Plant J.* **55**: 719–733.
- Huang, H., Liu, B., Liu, L., and Song, S.** (2017). Jasmonate action in plant growth and development. *J. Exp. Bot.* **68**: 1349–1359.
- Huchelmann, A., Boutry, M., and Hachez, C.** (2017). Plant glandular trichomes: natural cell factories of high biotechnological interest. *Plant Physiol.* **175**: 6–22.
- Huchelmann, A., Gastaldo, C., Veinante, M., Zeng, Y., Heintz, D., Tritsch, D., Schaller, H., Rohmer, M., Bach, T.J., and Hemmerlin, A.** (2014). S-Carvone Suppresses Cellulase-Induced Capsidiol Production in *Nicotiana tabacum* by Interfering with Protein Isoprenylation. *Plant Physiol.* **164**: 935–960.
- Iijima, Y., Davidovich-Rikanati, R., Fridman, E., Gang, D.R., Bar, E., Lewinsohn, E., and Pichersky, E.** (2004). The biochemical and molecular basis for the divergent patterns in the biosynthesis of terpenes and phenylpropenes in the peltate glands of three cultivars of basil. *Plant Physiol.* **136**: 3724–36.
- Irwin, R.E., Bronstein, J.L., Manson, J.S., and Richardson, L.** (2010). Nectar Robbing: Ecological and Evolutionary Perspectives. *Annu. Rev. Ecol. Evol. Syst.* **41**: 271–292.
- Jassbi, A.R., Gase, K., Hettenhausen, C., Schmidt, A., and Baldwin, I.T.** (2008). Silencing Geranylgeranyl Diphosphate Synthase in *Nicotiana attenuata* Dramatically Impairs Resistance to Tobacco Hornworm. *Plant Physiol.* **146**: 974–986.
- Jeong, N.-R., Kim, H., Hwang, I.-T., Howe, G.A., and Kang, J.-H.** (2017). Genetic analysis of the tomato *inquieta* mutant links the ARP2/3 complex to trichome development. *J. Plant Biol.* **60**: 582–592.
- Jiang, Z., Kempinski, C., and Chappell, J.** (2016). Extraction and Analysis of Terpenes/Terpenoids. *Curr. Protoc. plant Biol.* **1**: 345–358.
- Jones, J.G.** (2014). Identifying Sources of Hepatic Lipogenic Acetyl-CoA Using Stable Isotope Tracers and NMR. *Adv. Radiol.* **2014**: Article ID 109252.
- Judd, R., Bagley, M.C., Li, M., Zhu, Y., Lei, C., Yuzuak, S., Ekelöf, M., Pu, G., Zhao, X., Muddiman, D.C., and Xie, D.-Y.** (2019). Artemisinin biosynthesis in non-glandular trichome cells of *Artemisia annua*. *Mol. Plant* **12**: 704–714.
- Kai, G., Xu, H., Zhou, C., Liao, P., Xiao, J., Luo, X., You, L., and Zhang, L.** (2011). Metabolic engineering tanshinone biosynthetic pathway in *Salvia miltiorrhiza* hairy root cultures. *Metab. Eng.* **13**: 319–327.
- Kang, J.-H., Campos, M.L., Zemelis-Durfee, S., Al-Haddad, J.M., Jones, A.D., Telewski, F.W., Brandizzi, F., and Howe, G.A.** (2016). Molecular cloning of the tomato *Hairless* gene implicates actin dynamics in trichome-mediated defense and mechanical properties of stem tissue. *J. Exp. Bot.* **67**: 5313–5324.
- Kang, J.H., Shi, F., Jones, A.D., Marks, M.D., and Howe, G.A.** (2010). Distortion of trichome morphology by the *hairless* mutation of tomato affects leaf surface chemistry. *J. Exp. Bot.* **61**: 1053–1064.
- Keene, C.K. and Wagner, G.J.** (1985). Direct demonstration of divatrienediol biosynthesis in glandular heads of tobacco trichomes. *Plant Physiol.* **79**: 1026–32.
- Kharel, Y. and Koyama, T.** (2003). Molecular analysis of cis-prenyl chain elongating enzymes.

- Nat. Prod. Rep. **20**: 111–118.
- Kilian, B. and Graner, A.** (2012). NGS technologies for analyzing germplasm diversity in genebanks. *Brief. Funct. Genomics* **11**: 38–50.
- Kim, S.H., Kim, Y.-H., Ahn, Y.O., Ahn, M.-J., Jeong, J.C., Lee, H.-S., and Kwak, S.-S.** (2013). Downregulation of the lycopene ϵ -cyclase gene increases carotenoid synthesis via the β -branch-specific pathway and enhances salt-stress tolerance in sweetpotato transgenic calli. *Physiol. Plant.* **147**: 432–442.
- Kugler, H.** (1972). Zur Bestäubung von *Salvia sclarea* L. durch Holzbienen (*Xylocopa violacea* L.). *Österreichische Bot. Zeitschrift* **120**: 77–85.
- Kumar, R., Sharma, S., and Pathania, V.** (2013). Effect of shading and plant density on growth, yield and oil composition of clary sage (*Salvia sclarea* L.) in north western Himalaya. *J. Essent. Oil Res.* **25**: 23–32.
- Kutrzeba, L., Dayan, F.E., Howell, J., Feng, J., Giner, J.L., and Zjawiony, J.K.** (2007). Biosynthesis of salvinorin A proceeds via the deoxyxylulose phosphate pathway. *Phytochemistry* **68**: 1872–1881.
- Kuźma, Ł., Bruchajzer, E., and Wysokińska, H.** (2009a). Methyl jasmonate effect on diterpenoid accumulation in *Salvia sclarea* hairy root culture in shake flasks and sprinkle bioreactor. *Enzyme Microb. Technol.* **44**: 406–410.
- Kuźma, Ł., Kalemba, D., Rozalski, M., Rozalska, B., Marzena, W.S., Krajewska, U., and Wysokińska, H.** (2009b). Chemical composition and biological activities of essential oil from *salvia sclarea* plants regenerated in vitro. *Molecules* **14**: 1438–1447.
- Landmann, C., Fink, B., Festner, M., Dregus, M., Engel, K.H., and Schwab, W.** (2007). Cloning and functional characterization of three terpene synthases from lavender (*Lavandula angustifolia*). *Arch. Biochem. Biophys.* **465**: 417–429.
- Lange, B.M.** (2015). The Evolution of Plant Secretory Structures and Emergence of Terpenoid Chemical Diversity. *Annu. Rev. Plant Biol.* **66**: 139–159.
- Lange, B.M. and Turner, G.W.** (2013). Terpenoid biosynthesis in trichomes-current status and future opportunities. *Plant Biotechnol. J.* **11**: 2–22.
- Lattoo, S.K., Dhar, R.S., Dhar, A.K., Sharma, P.R., and Agarwal, S.G.** (2006). Dynamics of essential oil biosynthesis in relation to inflorescence and glandular ontogeny in *Salvia sclarea*. *Flavour Fragr. J.* **21**: 817–821.
- Laville, R., Castel, C., Filippi, J.J., Delbecque, C., Audran, A., Garry, P.P., Legendre, L., and Fernandez, X.** (2012). Amphilectane diterpenes from *Salvia sclarea*: Biosynthetic considerations. *J. Nat. Prod.* **75**: 121–126.
- Lawrence, B.M.** (1994). Production of Clary Sage Oil and Sclareol in North America. *Proc. 4th Nyons Int. Meet.*: 41–58.
- Leffingwell, J.C. and Leffingwell, D.** (2015). Biotechnology - Conquests and Challenges in Flavors & Fragrances. *Leffingwell Reports* **7**: 1–11.
- Legrand, S., Valot, N., Nicolé, F., Moja, S., Baudino, S., Jullien, F., Magnard, J.L., Caissard, J.C., and Legendre, L.** (2010). One-step identification of conserved miRNAs, their targets, potential transcription factors and effector genes of complete secondary metabolism pathways after 454 pyrosequencing of calyx cDNAs from the Labiate *Salvia sclarea* L. *Gene* **450**: 55–62.
- Li, L., Zhao, Y., McCaig, B.C., Wingerd, B.A., Wang, J., Whalon, M.E., Pichersky, E., and Howe, G.A.** (2004). The tomato homolog of CORONATINE-INSENSITIVE1 is required for the maternal control of seed maturation, jasmonate-signaled defense responses, and glandular trichome development. *Plant Cell* **16**: 126–43.

References

- Li, S., Zhang, P., Zhang, M., Fu, C., and Yu, L.** (2013). Functional analysis of a WRKY transcription factor involved in transcriptional activation of the DBAT gene in *Taxus chinensis*. *Plant Biol.* **15**: 19–26.
- Liao, P., Hemmerlin, A., Bach, T.J., and Chye, M.L.** (2016). The potential of the mevalonate pathway for enhanced isoprenoid production. *Biotechnol. Adv.* **34**: 697–713.
- Lipko, A. and Swiezewska, E.** (2016). Isoprenoid generating systems in plants - A handy toolbox how to assess contribution of the mevalonate and methylerythritol phosphate pathways to the biosynthetic process. *Prog. Lipid Res.* **63**: 70–92.
- Liu, W., Chilcott, C.E., Reich, R.C., and Hellmann, G.M.** (2000). Regeneration of *Salvia sclarea* via organogenesis. *Vitr. Cell. Dev. Biol. - Plant* **36**: 201–206.
- Liu, X., Bartholomew, E., Cai, Y., and Ren, H.** (2016). Trichome-Related Mutants Provide a New Perspective on Multicellular Trichome Initiation and Development in Cucumber (*Cucumis sativus* L). *Front. Plant Sci.* **7**: 1187.
- Lu, X., Tang, K., and Li, P.** (2016). Plant Metabolic Engineering Strategies for the Production of Pharmaceutical Terpenoids. *Front. Plant Sci.* **7**: 1–11.
- Lu, X., Zhang, L., Zhang, F., Jiang, W., Shen, Q., Zhang, L., Lv, Z., Wang, G., and Tang, K.** (2013). *AaORA*, a trichome-specific AP2/ERF transcription factor of *Artemisia annua*, is a positive regulator in the artemisinin biosynthetic pathway and in disease resistance to *Botrytis cinerea*. *New Phytol.* **198**: 1191–1202.
- Lv, Z., Zhang, F., Pan, Q., Fu, X., Jiang, W., Shen, Q., Yan, T., Shi, P., Lu, X., Sun, X., and Tang, K.** (2016). Branch Pathway Blocking in *Artemisia annua* is a Useful Method for Obtaining High Yield Artemisinin. *Plant Cell Physiol.* **57**: 588–602.
- Ma, D. et al.** (2016). Genetic basis for glandular trichome formation in cotton. *Nat. Commun.* **7**: 1–9.
- Maes, L. and Goossens, A.** (2010). Hormone-mediated promotion of trichome initiation in plants is conserved but utilizes species- and trichome-specific regulatory mechanisms. *Plant Signal. Behav.* **5**: 205–7.
- Maes, L., Van Nieuwerburgh, F.C.W., Zhang, Y., Reed, D.W., Pollier, J., Vande Castele, S.R.F., Inzé, D., Covello, P.S., Deforce, D.L.D., and Goossens, A.** (2011). Dissection of the phytohormonal regulation of trichome formation and biosynthesis of the antimalarial compound artemisinin in *Artemisia annua* plants. *New Phytol.* **189**: 176–189.
- Mafu, S. and Zerbe, P.** (2018). Plant diterpenoid metabolism for manufacturing the biopharmaceuticals of tomorrow: prospects and challenges. *Phytochem. Rev.* **17**: 113–130.
- Mahzooni-Kachapi, S.S., Mahdavi, M., Jouri, M.H., Akbarzadeh, M., and Roozbeh-nasira'ei, L.** (2014). The effects of altitude on chemical compositions and function of essential oils in *stachys lavandulifolia* vahl. (Iran). *Int. J. Med. Aromat. Plants* **4**: 107–116.
- Mastino, P.M., Marchetti, M., Costa, J., and Usai, M.** (2018). Interpopulation Variability in the Essential Oil Composition of *Cistus creticus* subsp. *eriocephalus* from Sardinia. *Chem. Biodivers.* **15**: e1800151.
- Matías-Hernández, L., Jiang, W., Yang, K., Tang, K., Brodelius, P.E., and Pelaz, S.** (2017). *AaMYB1* and its orthologue *AtMYB61* affect terpene metabolism and trichome development in *Artemisia annua* and *Arabidopsis thaliana*. *Plant J.* **90**: 520–534.
- McDowell, E.T. et al.** (2011). Comparative functional genomic analysis of *Solanum* glandular trichome types. *Plant Physiol.* **155**: 524–39.
- Melito, S., Sias, A., Petretto, G.L., Chessa, M., Pintore, G., and Porceddu, A.** (2013). Genetic

- and Metabolite Diversity of Sardinian Populations of *Helichrysum italicum*. *PLoS One* **8**: e79043.
- Mendoza-Poudereux, I., Kutzner, E., Huber, C., Segura, J., Arrillaga, I., and Eisenreich, W.** (2017). Dynamics of Monoterpene Formation in Spike Lavender Plants. *Metabolites* **7**: 65.
- Mendoza-Poudereux, I., Kutzner, E., Huber, C., Segura, J., Eisenreich, W., and Arrillaga, I.** (2015). Metabolic cross-talk between pathways of terpenoid backbone biosynthesis in spike lavender. *Plant Physiol. Biochem.* **95**: 113–120.
- Mendoza, L., Tapia, L., Wilkens, M., and Urzua, A.** (2002). Antibacterial activity of 13-epi-sclareol, a labdane-type diterpene isolated from *Pseudognaphalium heterotrichum* and *P. cheiranthifolium* (Asteraceae). *Boletín la Soc. Chil. Química* **47**: 91–98.
- Mofidi Tabatabaei, S., Salehi, P., Moridi Farimani, M., Neuburger, M., De Mieri, M., Hamburger, M., and Nejad-Ebrahimi, S.** (2017). A nor-diterpene from *Salvia sahendica* leaves. *Nat. Prod. Res.* **31**: 1758–1765.
- Moore, B.D., Andrew, R.L., Külheim, C., and Foley, W.J.** (2014). Explaining intraspecific diversity in plant secondary metabolites in an ecological context. *New Phytol.* **201**: 733–750.
- Moridi Farimani, M. and Miran, M.** (2014). Labdane diterpenoids from *Salvia reuterana*. *Phytochemistry* **108**: 264–269.
- Nadakuduti, S.S., Pollard, M., Kosma, D.K., Allen, C., Ohlrogge, J.B., and Barry, C.S.** (2012). Pleiotropic phenotypes of the sticky peel mutant provide new insight into the role of CUTIN DEFICIENT2 in epidermal cell function in tomato. *Plant Physiol.* **159**: 945–60.
- Nasermoadeli, S. and Rowshan, V.** (2013). Comparison of *Salvia sclarea* L. Essential Oil Components in Wild and Field Population. *Int. J. Agric. Crop Sci.* **4**: 1997–2000.
- Okada, A., Okada, K., Miyamoto, K., Koga, J., Shibuya, N., Nojiri, H., and Yamane, H.** (2009). OsTGAP1, a bZIP transcription factor, coordinately regulates the inductive production of diterpenoid phytoalexins in rice. *J. Biol. Chem.* **284**: 26510–8.
- Opitz, S., Nes, W.D., and Gershenzon, J.** (2014). Both methylerythritol phosphate and mevalonate pathways contribute to biosynthesis of each of the major isoprenoid classes in young cotton seedlings. *Phytochemistry* **98**: 110–119.
- Orlova, I. et al.** (2009). The small subunit of snapdragon geranyl diphosphate synthase modifies the chain length specificity of tobacco geranylgeranyl diphosphate synthase in planta. *Plant Cell* **21**: 4002–17.
- Otto, L.G., Mondal, P., Brassac, J., Preiss, S., Degenhardt, J., He, S., Reif, J.C., and Sharbel, T.F.** (2017). Use of genotyping-by-sequencing to determine the genetic structure in the medicinal plant chamomile, and to identify flowering time and alpha-bisabolol associated SNP-loci by genome-wide association mapping. *BMC Genomics* **18**: 1–18.
- Ozdemir, C. and Senel, G.** (1999). The morphological, anatomical and karyological properties of *Salvia sclarea* L. *Turk J Bot.* **23**: 7–18.
- Park, J.-E., Lee, K.-E., Jung, E., Kang, S., and Kim, Y.J.** (2016). Sclareol isolated from *Salvia officinalis* improves facial wrinkles via an antiphotaging mechanism. *J. Cosmet. Dermatol.* **15**: 475–483.
- Patil, R.A., Lenka, S.K., Normanly, J., Walker, E.L., and Roberts, S.C.** (2014). Methyl jasmonate represses growth and affects cell cycle progression in cultured *Taxus* cells. *Plant Cell Rep.* **33**: 1479–1492.
- Pattanaik, S., Patra, B., Singh, S.K., and Yuan, L.** (2014). An overview of the gene regulatory network controlling trichome development in the model plant, *Arabidopsis*. *Front. Plant*

- Sci. **5**: 259.
- Pauwels, L., Morreel, K., De Witte, E., Lammertyn, F., Van Montagu, M., Boerjan, W., Inzé, D., and Goossens, A.** (2008). Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured Arabidopsis cells. *Proc. Natl. Acad. Sci. U. S. A.* **105**: 1380–5.
- Peters, R.J.** (2010). Two rings in them all: The labdane-related diterpenoids. *Nat. Prod. Rep.* **27**: 1521.
- Rai, A., Smita, S.S., Singh, A.K., Shanker, K., and Nagegowda, D.A.** (2013). Heteromeric and homomeric geranyl diphosphate synthases from *Catharanthus roseus* and their role in monoterpene indole alkaloid biosynthesis. *Mol. Plant* **6**: 1531–49.
- Rapposelli, E., Melito, S., Barmina, G.G., Foddai, M., Azara, E., and Scarpa, G.M.** (2015). Relationship Between Soil and Essential Oil Profiles in *Salvia desoleana* Populations: Preliminary Results. *Nat. Prod. Commun.* **10**: 1615–8.
- Rasmann, S., Köllner, T.G., Degenhardt, J., Hiltpold, I., Toepfer, S., Kuhlmann, U., Gershenzon, J., and Turlings, T.C.J.** (2005). Recruitment of entomopathogenic nematodes by insect-damaged maize roots. *Nature* **434**: 732–737.
- Reed, J. and Osbourn, A.** (2018). Engineering terpenoid production through transient expression in *Nicotiana benthamiana*. *Plant Cell Rep.* **37**: 1431–1441.
- Rodríguez-Concepcion, M. et al.** (2018). A global perspective on carotenoids: Metabolism, biotechnology, and benefits for nutrition and health. *Prog. Lipid Res.* **70**: 62–93.
- Rodríguez-Concepción, M., Forés, O., Martínez-García, J.F., González, V., Phillips, M.A., Ferrer, A., and Boronat, A.** (2004). Distinct Light-Mediated Pathways Regulate the Biosynthesis and Exchange of Isoprenoid Precursors during Arabidopsis Seedling Development. *Plant Cell* **16**: 144–156.
- Ron, M. et al.** (2014). Hairy root transformation using *Agrobacterium rhizogenes* as a tool for exploring cell type-specific gene expression and function using tomato as a model. *Plant Physiol.* **166**: 455–69.
- Russo, A., Formisano, C., Rigano, D., Senatore, F., Delfino, S., Cardile, V., Rosselli, S., and Bruno, M.** (2013). Chemical composition and anticancer activity of essential oils of Mediterranean sage (*Salvia officinalis* L.) grown in different environmental conditions. *Food Chem. Toxicol.* **55**: 42–47.
- Sallaud, C., Rontein, D., Onillon, S., Jabès, F., Duffé, P., Giacalone, C., Thoraval, S., Escoffier, C., Herbet, G., Leonhardt, N., Causse, M., and Tissier, A.** (2009). A novel pathway for sesquiterpene biosynthesis from Z,Z-farnesyl pyrophosphate in the wild tomato *Solanum habrochaites*. *Plant Cell* **21**: 301–17.
- Sandes, S.S., Zucchi, M.I., Pinheiro, J.B., Bajay, M.M., Batista, C.E.A., Brito, F.A., Arrigoni-Blank, M.F., Alvares-Carvalho, S. V., Silva-Mann, R., and Blank, A.F.** (2016). Molecular characterization of patchouli (*Pogostemon* spp) germplasm. *Genet. Mol. Res.* **15**: gmr.15017458.
- Sashidhara, K. V., Rosaiah, J.N., Kumar, A., Bid, H.K., Konwar, R., and Chattopadhyay, N.** (2007). Cell growth inhibitory action of an unusual labdane diterpene, 13-epi-sclareol in breast and uterine cancers in vitro. *Phyther. Res.* **21**: 1105–1108.
- Sathasivam, R. and Ki, J.-S.** (2018). A Review of the Biological Activities of Microalgal Carotenoids and Their Potential Use in Healthcare and Cosmetic Industries. *Mar. Drugs* **16**: 26.
- Schalk, M.** (2009). Method for producing sclareol. WO2009101126_A1.
- Schalk, M., Pastore, L., Mirata, M.A., Khim, S., Schouwey, M., Deguerry, F., Pineda, V.,**

References

- Rocci, L., and Daviet, L.** (2012). Toward a Biosynthetic Route to Sclareol and Amber Odorants. *J. Am. Chem. Soc.* **134**: 18900–18903.
- Schmelz, E.A., Huffaker, A., Sims, J.W., Christensen, S.A., Lu, X., Okada, K., and Peters, R.J.** (2014). Biosynthesis, elicitation and roles of monocot terpenoid phytoalexins. *Plant J.* **79**: 659–678.
- Schmiderer, C., Grassi, P., Novak, J., Weber, M., and Franz, C.** (2008). Diversity of essential oil glands of clary sage (*Salvia sclarea* L., Lamiaceae). *Plant Biol.* **10**: 433–440.
- Seo, S., Gomi, K., Kaku, H., Abe, H., Seto, H., Nakatsu, S., Neya, M., Kobayashi, M., Nakaho, K., Ichinose, Y., Mitsuhashi, I., and Ohashi, Y.** (2012). Identification of Natural Diterpenes that Inhibit Bacterial Wilt Disease in Tobacco, Tomato and Arabidopsis. *Plant Cell Physiol.* **53**: 1432–1444.
- Sharopov, F.S. and Setzer, W.N.** (2012). The essential oil of *Salvia sclarea* L. from Tajikistan. *Rec. Nat. Prod.* **6**: 75–79.
- Shen, Q. et al.** (2018). The Genome of *Artemisia annua* Provides Insight into the Evolution of Asteraceae Family and Artemisinin Biosynthesis. *Mol. Plant* **11**: 776–788.
- Shen, Q., Lu, X., Yan, T., Fu, X., Lv, Z., Zhang, F., Pan, Q., Wang, G., Sun, X., and Tang, K.** (2016). The jasmonate-responsive AaMYC2 transcription factor positively regulates artemisinin biosynthesis in *Artemisia annua*. *New Phytol.* **210**: 1269–1281.
- Shi, M., Luo, X., Ju, G., Yu, X., Hao, X., Huang, Q., Xiao, J., Cui, L., and Kai, G.** (2014). Increased accumulation of the cardio-cerebrovascular disease treatment drug tanshinone in *Salvia miltiorrhiza* hairy roots by the enzymes 3-hydroxy-3-methylglutaryl CoA reductase and 1-deoxy-D-xylulose 5-phosphate reductoisomerase. *Funct. Integr. Genomics* **14**: 603–615.
- Shi, M., Zhou, W., Zhang, J., Huang, S., Wang, H., and Kai, G.** (2016). Methyl jasmonate induction of tanshinone biosynthesis in *Salvia miltiorrhiza* hairy roots is mediated by JASMONATE ZIM-DOMAIN repressor proteins. *Sci. Rep.* **6**: 20919.
- Shi, P. et al.** (2018). The roles of AaMIXTA1 in regulating the initiation of glandular trichomes and cuticle biosynthesis in *Artemisia annua*. *New Phytol.* **217**: 261–276.
- Simonsen, H., Chen, D.-F., Zhan, X., Pan, X., and King, B.C.** (2014). Heterologous production of patchoulol, beta-santalene, and sclareol in moss cells. *WO2014206412_A1*.
- Soetaert, S.S., Van Neste, C.M., Vandewoestyne, M.L., Head, S.R., Goossens, A., Van Nieuwerburgh, F.C., and Deforce, D.L.** (2013). Differential transcriptome analysis of glandular and filamentous trichomes in *Artemisia annua*. *BMC Plant Biol.* **13**: 220.
- Stukkens, Y., Bultreys, A., Grec, S., Trombik, T., Vanham, D., and Boutry, M.** (2005). NpPDR1, a Pleiotropic Drug Resistance-Type ATP-Binding Cassette Transporter from *Nicotiana plumbaginifolia*, Plays a Major Role in Plant Pathogen Defense. *Plant Physiol.* **108**: 1741–1746.
- Takada, S. and Iida, H.** (2014). Specification of epidermal cell fate in plant shoots. *Front. Plant Sci.* **5**: 49.
- Tamogami, S., Noge, K., Abe, M., Agrawal, G.K., and Rakwal, R.** (2012). Methyl jasmonate is transported to distal leaves via vascular process metabolizing itself into JA-Ile and triggering VOCs emission as defensive metabolites. *Plant Signal. Behav.* **7**: 1378–81.
- Tan, H., Xiao, L., Gao, S., Li, Q., Chen, J., Xiao, Y., Ji, Q., Chen, R., Chen, W., and Zhang, L.** (2015). TRICHOME and ARTEMISININ REGULATOR 1 is required for trichome development and artemisinin biosynthesis in *Artemisia annua*. *Mol. Plant* **8**: 1396–1411.
- Tanaka, Y., Sasaki, N., and Ohmiya, A.** (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* **54**: 733–749.

- Tetali, S.D.** (2018). Terpenes and isoprenoids: a wealth of compounds for global use. *Planta* **249**: 1–8.
- Thines, B., Katsir, L., Melotto, M., Niu, Y., Mandaokar, A., Liu, G., Nomura, K., He, S.Y., Howe, G.A., and Browse, J.** (2007). JAZ repressor proteins are targets of the SCFCO11 complex during jasmonate signalling. *Nature* **448**: 661–665.
- Tholl, D.** (2015). Biosynthesis and Biological Functions of Terpenoids in Plants. *Adv. Biochem. Eng. Biotechnol.* **123**: 127–141.
- Tholl, D., Kish, C.M., Orlova, I., Sherman, D., Gershenzon, J., Pichersky, E., and Dudareva, N.** (2004). Formation of monoterpenes in *Antirrhinum majus* and *Clarkia breweri* flowers involves heterodimeric geranyl diphosphate synthases. *Plant Cell* **16**: 977–92.
- Thomma, B.P.H.J., Eggermont, K., Broekaert, W.F., and Cammue, B.P.A.** (2000). Disease development of several fungi on *Arabidopsis* can be reduced by treatment with methyl jasmonate. *Plant Physiol. Biochem.* **38**: 421–427.
- Tingey, D.T., Manning, M., Grothaus, L.C., and Burns, W.F.** (1980). Influence of light and temperature on monoterpene emission rates from slash pine. *Plant Physiol.* **65**: 797–801.
- Tippmann, S., Chen, Y., Siewers, V., and Nielsen, J.** (2013). From flavors and pharmaceuticals to advanced biofuels: production of isoprenoids in *Saccharomyces cerevisiae*. *Biotechnol. J.* **8**: 1435–44.
- Tissier, A.** (2012). Glandular trichomes: What comes after expressed sequence tags? *Plant J.* **70**: 51–68.
- Tissier, A.** (2018). Plant secretory structures: more than just reaction bags. *Curr. Opin. Biotechnol.* **49**: 73–79.
- Toledo-Ortiz, G., Huq, E., and Rodríguez-Concepción, M.** (2010). Direct regulation of phytoene synthase gene expression and carotenoid biosynthesis by phytochrome-interacting factors. *Proc. Natl. Acad. Sci. U. S. A.* **107**: 11626–31.
- Tominaga-Wada, R., Nukumizu, Y., Sato, S., and Wada, T.** (2013). Control of Plant Trichome and Root-Hair Development by a Tomato (*Solanum lycopersicum*) R3 MYB Transcription Factor. *PLoS One* **8**: e54019.
- Töpfer, N., Fuchs, L.-M., and Aharoni, A.** (2017). The PhytoClust tool for metabolic gene clusters discovery in plant genomes. *Nucleic Acids Res.* **45**: 7049–7063.
- Towler, M.J. and Weathers, P.J.** (2007). Evidence of artemisinin production from IPP stemming from both the mevalonate and the nonmevalonate pathways. *Plant Cell Rep.* **26**: 2129–2136.
- Trapp, S. and Croteau, R.** (2001). Defensive resin biosynthesis in conifers. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**: 689–724.
- Vaccaro, M., Malafrente, N., Alfieri, M., De Tommasi, N., and Leone, A.** (2014). Enhanced biosynthesis of bioactive abietane diterpenes by overexpressing AtDXS or AtDXR genes in *Salvia sclarea* hairy roots. *Plant Cell, Tissue Organ Cult.* **119**: 65–77.
- Vaccaro, M.C., Alfieri, M., Malafrente, N., De Tommasi, N., and Leone, A.** (2017). Increasing the synthesis of bioactive abietane diterpenes in *Salvia sclarea* hairy roots by elicited transcriptional reprogramming. *Plant Cell Rep.* **36**: 375–386.
- Vendemiatti, E., Zsögön, A., Ferreira E Silva, G.F., de Jesus, F.A., Cutri, L., Figueiredo, C.R.F., Tanaka, F.A.O., Nogueira, F.T.S., and Peres, L.E.P.** (2017). Loss of type-IV glandular trichomes is a heterochronic trait in tomato and can be reverted by promoting juvenility. *Plant Sci.* **259**: 35–47.
- Vranová, E., Coman, D., and Gruissem, W.** (2013). Network Analysis of the MVA and MEP

- Pathways for Isoprenoid Synthesis. *Annu. Rev. Plant Biol.* **64**: 665–700.
- Wagner, S., Pflieger, A., Mandl, M., and Böchzelt, H.** (2012). Changes in the Qualitative and Quantitative Composition of Essential Oils of Clary Sage and Roman Chamomile During Steam Distillation in Pilot Plant Scale. *Distill. - Adv. from Model. to Appl.*: 141–158.
- Wang, G. and Dixon, R.A.** (2009). Heterodimeric geranyl(geranyl)diphosphate synthase from hop (*Humulus lupulus*) and the evolution of monoterpene biosynthesis. *Proc. Natl. Acad. Sci.* **106**: 9914–9919.
- Wang, Q., Reddy, V.A., Panicker, D., Mao, H.Z., Kumar, N., Rajan, C., Venkatesh, P.N., Chua, N.H., and Sarojam, R.** (2016). Metabolic engineering of terpene biosynthesis in plants using a trichome-specific transcription factor MsYABBY5 from spearmint (*Mentha spicata*). *Plant Biotechnol. J.* **14**: 1619–1632.
- Wasternack, C. and Strnad, M.** (2019). Jasmonates are signals in the biosynthesis of secondary metabolites — Pathways, transcription factors and applied aspects — A brief review. *N. Biotechnol.* **48**: 1–11.
- Werker, E.** (1993). Function of essential oil-secreting glandular hairs in aromatic plants of Lamiaceae—a review. *Flavour Fragr. J.* **8**: 249–255.
- Werker, E., Ravid, U., and Putievsky, E.** (1985). Glandular hairs and their secretions in the vegetative and reproductive organs of *Salvia sclarea* and *S. dominica*. *Isr. J. Plant Sci.* **34**: 239–252.
- Xiao, L., Tan, H., and Zhang, L.** (2016). *Artemisia annua* glandular secretory trichomes: the biofactory of antimalarial agent artemisinin. *Sci. Bull.* **61**: 26–36.
- Xu, H. et al.** (2016). Analysis of the Genome Sequence of the Medicinal Plant *Salvia miltiorrhiza*. *Mol. Plant* **9**: 949–952.
- Xu, J., van Herwijnen, Z.O., Dräger, D.B., Sui, C., Haring, M.A., and Schuurink, R.C.** (2018). SLMYC1 Regulates Type VI Glandular Trichome Formation and Terpene Biosynthesis in Tomato Glandular Cells. *Plant Cell* **30**: 2988–3005.
- Xu, M., Galhano, R., Wiemann, P., Bueno, E., Tiernan, M., Wu, W., Chung, I.-M., Gershenzon, J., Tudzynski, B., Sesma, A., and Peters, R.J.** (2012). Genetic evidence for natural product-mediated plant-plant allelopathy in rice (*Oryza sativa*). *New Phytol.* **193**: 570–575.
- Xue, S., Dong, M., Liu, X., Xu, S., Pang, J., Zhang, W., Weng, Y., and Ren, H.** (2018). Classification of fruit trichomes in cucumber and effects of plant hormones on type II fruit trichome development. *Planta* **249**: 407–416.
- Yan, T. et al.** (2018). A novel HD-ZIP IV/MIXTA complex promotes glandular trichome initiation and cuticle development in *Artemisia annua*. *New Phytol.* **218**: 567–578.
- Yan, T. et al.** (2017). HOMEODOMAIN PROTEIN 1 is required for jasmonate-mediated glandular trichome initiation in *Artemisia annua*. *New Phytol.* **213**: 1145–1155.
- Yang, C., Gao, Y., Gao, S., Yu, G., Xiong, C., Chang, J., Li, H., and Ye, Z.** (2015). Transcriptome profile analysis of cell proliferation molecular processes during multicellular trichome formation induced by tomato *Wo v* gene in tobacco. *BMC Genomics* **16**: 868.
- Yang, C., Li, H., Zhang, J., Luo, Z., Gong, P., Zhang, C., Li, J., Wang, T., Zhang, Y., Lu, Y., and Ye, Z.** (2011). A regulatory gene induces trichome formation and embryo lethality in tomato. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 11836–41.
- Yang, C. and Ye, Z.** (2013). Trichomes as models for studying plant cell differentiation. *Cell. Mol. Life Sci.* **70**: 1937–1948.
- Yang, S., Cai, Y., Liu, X., Dong, M., Zhang, Y., Chen, S., Zhang, W., Li, Y., Tang, M., Zhai, X., Weng, Y., and Ren, H.** (2018). CsMYB6-CsTRY module regulates fruit trichome initiation

References

- in cucumber. *J. Exp. Bot.* **69**: 1887–1902.
- Yang, S., Tian, H., Sun, B., Liu, Y., Hao, Y., and Lv, Y.** (2016). One-pot synthesis of (–)-Ambrox. *Sci. Rep.* **6**: 32650.
- Yu, X., Chen, G., Tang, B., Zhang, J., Zhou, S., and Hu, Z.** (2018). The Jasmonate ZIM-domain protein gene SIJAZ2 regulates plant morphology and accelerates flower initiation in *Solanum lycopersicum* plants. *Plant Sci.* **267**: 65–73.
- Yu, Z.-X., Li, J.-X., Yang, C.-Q., Hu, W.-L., Wang, L.-J., and Chen, X.-Y.** (2011). The Jasmonate-Responsive AP2/ERF Transcription Factors AaERF1 and AaERF2 Positively Regulate Artemisinin Biosynthesis in *Artemisia annua* L. *Mol. Plant* **5**: 353–365.
- Zaks, A., Davidovich-Rikanati, R., Bar, E., Inbar, M., and Lewinsohn, E.** (2008). Biosynthesis of linalyl acetate and other terpenes in lemon mint (*Mentha aquatica* var. *citrata*, Lamiaceae) glandular trichomes. *Isr. J. Plant Sci.* **56**: 233–244.
- Zerbe, P. and Bohlmann, J.** (2015). Plant diterpene synthases: Exploring modularity and metabolic diversity for bioengineering. *Trends Biotechnol.* **33**: 419–428.
- Zhang, F., Fu, X., Lv, Z., Lu, X., Shen, Q., Zhang, L., Zhu, M., Wang, G., Sun, X., Liao, Z., and Tang, K.** (2015a). A basic leucine zipper transcription factor, AabZIP1, connects abscisic acid signaling with artemisinin biosynthesis in *Artemisia annua*. *Mol. Plant* **8**: 163–175.
- Zhang, X., Yan, F., Tang, Y., Yuan, Y., Deng, W., and Li, Z.** (2015b). Auxin Response Gene SIARF3 Plays Multiple Roles in Tomato Development and is Involved in the Formation of Epidermal Cells and Trichomes. *Plant Cell Physiol.* **56**: 2110–2124.
- Zhou, M. and Memelink, J.** (2016). Jasmonate-responsive transcription factors regulating plant secondary metabolism. *Biotechnol. Adv.* **34**: 441–449.

Titre : Biosynthèse du sclaréol et sa régulation chez la sauge sclarée

Mots-clés : biologie des plantes, terpène, ingénierie métabolique, sclaréol, sauge sclarée, métabolisme spécialisé des plantes

Résumé : Le sclaréol est un diterpène produit par les organes floraux de la sauge sclarée (*Salvia sclarea*, Lamiaceae). Il est utilisé en parfumerie pour l'hémisynthèse de l'ambroxide, une substance caractérisée par une odeur ambrée et une grande capacité de fixation des parfums. L'augmentation de la demande mondiale en sclaréol stimule actuellement les tentatives d'accroître le rendement de la production de sclaréol à partir de la sauge sclarée. L'objectif du travail présenté dans ce manuscrit était d'améliorer notre compréhension de la biosynthèse du sclaréol et de sa régulation chez la sauge sclarée, afin de mettre en évidence des stratégies d'augmentation du contenu en sclaréol de la sauge sclarée. L'analyse de la surface des calices de sauge sclarée par imagerie par spectrométrie de masse suggère que le sclaréol est principalement sécrété par des structures épidermiques spécialisées appelées trichomes glandulaires. De plus, nous avons mis en évidence les contributions respectives des deux voies de biosynthèse des terpènes présentes chez les plantes, les voies MVA et MEP, à la biosynthèse de trois terpènes de la sauge sclarée. Des expériences de marquage au ^{13}C indiquent que le sclaréol et l'acétate de linalyle sont tous deux issus de la voie MEP, alors que le β -caryophyllène semble être d'origine mixte. Nous avons également étudié le rôle potentiel d'une phytohormone, le méthyljasmonate, dans la régulation de la production de sclaréol chez la sauge sclarée. Enfin, nous avons exploré la diversité génétique et phénotypique de populations croates de sauge sclarée sauvage, et montrons que ces populations représentent une ressource génétique distincte par rapport aux populations de référence. L'ensemble de ces résultats met en évidence des pistes prometteuses pour l'amélioration génétique ciblée des performances de la sauge sclarée.

Title: Sclareol biosynthesis in clary sage and its regulation

Keywords: plant biology, terpenoid, metabolic engineering, sclareol, clary sage, plant specialized metabolism

Abstract: Sclareol is a diterpene produced by floral organs of clary sage (*Salvia sclarea*, Lamiaceae). It is used in perfume industry for the hemisynthesis of ambroxide, a high-valued perfume component characterized by an amber scent and a high perfume fixation capacity. The global demand for sclareol currently rises, prompting attempts at increasing the yield of sclareol production from clary sage. The purpose of the work presented in this manuscript was to improve knowledge on sclareol biosynthesis and its regulation in clary sage, in order to highlight strategies aiming at enhancing clary sage sclareol content. The analysis of the surface of clary sage calyces by mass spectrometry imaging suggests that sclareol is mainly secreted by specialized epidermal structures called glandular trichomes. Moreover, we have highlighted the respective contributions of the two terpenoid biosynthesis pathways present in plants, MVA and MEP pathways, to the biosynthesis of three terpenoids of clary sage. ^{13}C -labeling experiments indicate that sclareol and linalyl acetate both originate from the MEP pathway, whereas β -caryophyllene seems to be of mixed origin. We have also investigated the potential role of a phytohormone, methyljasmonate, in the regulation of sclareol production in clary sage. Finally, we have explored the genetic and phenotypic diversity of Croatian wild clary sage populations and show that these populations represent a distinct genetic resource compared to reference populations. Taken together, these results highlight promising avenues for targeted genetic enhancement of clary sage performances.

